



Gene Frequency in Small Populations of Mutant *Drosophila*

Author(s): Peter Buri

Source: *Evolution*, Dec., 1956, Vol. 10, No. 4 (Dec., 1956), pp. 367-402

Published by: Society for the Study of Evolution

Stable URL: <https://www.jstor.org/stable/2406998>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



Society for the Study of Evolution is collaborating with JSTOR to digitize, preserve and extend access to *Evolution*

JSTOR

GENE FREQUENCY IN SMALL POPULATIONS OF MUTANT *DROSOPHILA*¹

PETER BURI

State University of Iowa

Received April 17, 1956

INTRODUCTION

Studies of gene frequency changes in laboratory populations have for the most part been concerned with the systematic effects of selection and with discovery of the nature of the selective differences involved between genetic alternatives. The classical experiments of L'Heritier and Teissier (1934, 1937) with common mutants, the later work of Dobzhansky (Wright and Dobzhansky 1946), and that of Wallace (1948) on naturally occurring genetic variants are examples of such studies applied to rather large populations maintained in population cages of the sort developed by L'Heritier and Teissier (1933).

Such experiments as these were carried out under conditions in which changes in gene frequency due to sampling accidents (random drift) were negligible. In the work of Reed and Reed (1948, 1950), Ludwin (1951) and Merrell (1953) on selective differences between mutants and wild type, non-systematic effects of considerable magnitude and attributable to accidents of sampling occurred; however, these data did not permit a close evaluation of the extent of this influence.

An experimental design in which both random and systematic pressures are brought into play in an orderly fashion and the values of the several theoretical parameters can be ascertained was originated by Kerr (Kerr and Wright, 1954). The method consists in observations on the changes in gene frequency between generations in each of a large number of

small cultures in which the size of the breeding population is controlled. Each initial culture founds a line in which each successive generation is initiated with a random sample of uniform size taken from among the flies of the preceding generation. The form is that of a fairly large experimental population divided into a number of very small completely isolated subunits.

The present study was intended to provide information on the relative selective values of two mutant autosomal alleles and their several zygotic types under several different experimental conditions and to observe the effect of the interaction of selection and random drift on the distribution of gene frequencies among the elements of subdivided populations of the sort observed by Kerr.

MATERIALS AND METHODS

Stocks

The stocks used were obtained from strains of mutants at the brown locus of *Drosophila melanogaster* developed by Slatis (1955). The *bw*⁷⁵ strain was derived by him in June of 1950 by x-irradiation of males from a scarlet stock which was very closely related to an "isogenic" stock of *bw; st*. This *bw; st* stock had been made "isogenic" by a marker-chromosome technique followed by over 40 generations of brother-sister mating. The scarlet stock at the time of irradiation was derived from a series of cultures stemming from an outcross of *bw; st* to *st* at the time homozygosity was initially enforced in the former. This outcross was followed by repeated backcrosses of single heterozygous females (*bw/+; st/st*) to males

¹ Submitted to the Department of Zoology, University of Chicago, in partial fulfillment of the requirements for the degree of Doctor of Philosophy, June 1955.

of the *bw; st* stock in each of the subsequent generations of its inbreeding.

After December of 1950 both mutants were bred as common laboratory stocks until May of 1953 (Slatis, personal communication). At this time *bw⁷⁵* was crossed with the related *bw* stock and a single heterozygous female was backcrossed to the parental *bw* strain (since *st* was homozygous in all strains used in this study, it will usually be dropped from subsequent notation). Such backcrosses were repeated for eight generations after which time strains of *bw⁷⁵* and *bw* were isolated for use as the foundation stocks of these experiments. These two stocks were, then, closely related, by reason of repeated backcross, common ancestry, and the origin of *bw⁷⁵* by mutation in a second chromosome which may be presumed to have been, at the time of irradiation, identical with that of the *bw* strain except in the immediate vicinity of the *bw* locus. The chromosomes of both stocks have been examined by Slatis who has stated that they are free of structural abnormalities (Slatis, personal communication).

The phenotype of the *bw⁷⁵* homozygotes (*bw⁷⁵/bw⁷⁵; st/st*) is a bright red-orange in young flies raised at 25° C. This color darkens somewhat with age and at extreme age may appear a deep red-brown. The heterozygote (*bw⁷⁵/bw; st/st*) is a light orange at this temperature and in newly emerged flies. This, again, darkens with age but is always clearly distinguishable from the *bw⁷⁵* homozygotes. The *bw* homozygote (*bw/bw; st/st*) is white. Routine experience has indicated that the eye colors of homozygous *bw⁷⁵; st* and its heterozygote with *bw; st* are temperature dependent and that the distinction between them may not be as clear at lower temperatures. The material of all experiments in this study was maintained in an incubator regulated at 25° C.

The Random Drift Experiments

Two culture series of 107 and 105 bottles respectively were initiated at a gene frequency of 0.50 by introducing 16

bw⁷⁵/bw flies, 8 male and 8 female, into culture bottles containing about 1 cm. of cornmeal-molasses-agar medium which had been lightly sprinkled with dry baker's yeast. Series I was set up in 35 cc. homeopathic vials with an inside diameter of roughly 2 cm. Series II was cultured in 60 cc. specimen bottles having an inside diameter of about 4 cm. Each successive generation of each bottle was initiated with 16 flies chosen according to the genotypic array in a random sample of 16 flies, 8 male and 8 female, drawn from among the F₁ individuals of the preceding generation of the same culture. Samples were always taken on a particular day after mating, the 12th day in series I and the 14th day in series II. The F₁ individuals appeared on the 10th day in both series. The flies in a bottle to be sampled were etherized, dumped on a paper, and spread out into a line with a brush. The genotypes of the first 8 individuals of each sex to appear in this row characterized the sample.

In order to avoid inbreeding at loci other than brown, an effort was made to reduce or to eliminate the correlation among the parents of any given culture with respect to their origin from among the bottles of the parental generation. Because of the limited number of flies obtainable from the vial cultures of series I after the second day of eclosion it was necessary, for the most part, to use the sample flies themselves as parents of the following generation of the same line. However, the tendency to enforce homozygosis among residual loci was reduced in this series by substituting for one or two flies in a given sample flies of the same phenotype that had been collected as supernumeraries at the time of sampling. In series II the problem could be met more satisfactorily. The eclosion after the twelfth day was sufficiently heavy to make possible the complete separation of sampling and the collection of flies to serve as parents for the next generation. Parental flies were collected on the 12th day from all bottles of the series. Sam-

pling was carried out on the 14th day from among the flies emerging on the 13th day. No flies collected after the 12th day were used as parents in either series and since mating was always carried out on the 14th day the parental flies of each generation of both series were aged two days at the time of mating. Mating was, of course, always according to the genotypic array of the 16 sample flies. All flies used as parents were collected within 12 hours after the cultures had been carefully emptied. Females collected within this period were taken as virgin. More certain criteria of virginity were not feasible because of the large number of parents of nearly uniform age needed to perpetuate a culture series.

The Selection Experiments

To provide data on the mode and magnitude of selection that were independent of the random-drift experiments and would permit a more systematic analysis of the selection parameters than would be possible from the random assortment of population types occurring within the random-drift experiments themselves, a series of experiments was set up designed to present the simplest sort of data on the relative viability and productivity of the several genotypes. Columns 2 and 3 of table 4 present schematically the various types of group mating used to this purpose. Three major categories of population type appear: (I: Comparisons of viability made by examining segregations from group matings between 8 heterozygous females and 8 males of uniform genotype in each case; II: Matings between equal numbers of 2 types of females, 4 each, and 8 individuals of each type of male in all of the 9 possible combinations to test for competitive productivity of females after allowing for possible viability differences among genotypes; III: Reciprocal matings of category II to test for competitive productivity of males.)

These matings were made in 60 cc. specimen jars. The routine of sampling

and collection of flies to serve as parents was the same as that for the series II drift experiments with the exception that the sample consisted of 32 flies equally divided between the sexes instead of 16.

In order to distribute equally among the different categories of parent population any heterogeneity in culture conditions such as might be expected among experiments carried out at different times, an effort was made to make 5 replications of each of the 21 population types every time a group of matings was made. The few irregularities that occurred in carrying out this plan were due either to accidental loss of bottles or inadequate numbers of parental flies of appropriate genotype.

Population Cage Experiments on Selection

A number of small population cages were constructed of plexiglas for the purpose of observing competition between the alleles *bw*⁷⁵ and *bw* under continuous culture and such adverse conditions as might be expected to occur in a confined population facing limited food supply. Selective differences that could easily show no influence on the results of the other experiments, such as longevity, total fecundity, etc. and differences negligible under more favorable conditions of culture might be expected to appear upon prolonged observation of population cages.

The cages had bases measuring roughly 8×13.5 cm. and were about 5 cm. in height. Ventilation was permitted by a 2.2 cm. hole bored in one end and stoppered with a cotton plug that could be renewed periodically. The culture medium was placed on a slant in 35 cc. homeopathic vials which were attached to lucite tubes furnished with pressed cork bushings and set into the long sides of each container, four on each side.

Initially seven populations of 100 heterozygous flies each, 50 male and 50 female, were introduced into as many cages each provided with eight freshly yeasted vials. The females of these parental

groups had been isolated as virgins from the F_1 of a series of matings of bw^{75}/bw^{75} with bw/bw . Before introduction into the cages these flies had been predated in small groups for from one to three days in half-pint bottles containing an abundance of food.

On the 18th day after the initiation of cultures within the cages a single fresh vial was substituted for one of the original vials on each cage. On every third day after this another of the original vials was replaced until all had been removed. After this time the oldest vials remaining were replaced with fresh ones at the same three day interval. Sampling was begun on the 60th day after initiation and was accomplished by counting and classifying the flies that accumulated in a newly attached vial during the first 6 or 8 hours. After classification all flies were returned to the vial which was then re-attached to its cage.

RESULTS

The Population Cage Experiments

Table 1 presents the data from observation of the seven population cages over a period of fifteen weeks. Six observations were made on each cage. At the time of the first sample on the 60th day the frequency of bw^{75} differed strikingly among them. The frequency was in all cases less than 0.50 and ranged from

0.491 to 0.385 with a mean value of $0.446 \pm .013$, significantly below 0.50 ($t = 4.2, p < .001$).

There were no significant differences between the estimates of gene frequency based on the male and female components of each sample as is shown by the comparisons appearing in table 2. Accordingly, the frequency in this discussion is always that of the whole sample.

Despite the diversity in frequency among the cages at the time of the first sample, regression coefficients computed for the series of frequency values obtained from all six samples taken from each cage indicated no significant trends during the period of observation in any case. These coefficients together with their standard errors appear in table 1. Three are positive and four negative, and the regression on sample number of the mean values of the frequency is very small and of no significance.

Considering the group of cage experiments as a whole, there is no indication of a difference in relative selective value between the two alleles apart from the fact that the observed frequencies were, for the most part, less than 0.50 throughout the period of observation. The differentiation between cages, which is fairly consistent throughout sampling, must have been due to shifts in frequency during the earliest period. The significance

TABLE 1. Gene frequency (bw^{75}) in 6 successive samples taken at three-week intervals from each of 7 population cages and the mean frequency among all cages at the sample time

Sample	Cage							\bar{q}
	I	II	III	IV	V	VI	VII	
1	.4907	.4598	.4231	.4673	.4527	.3853	.4420	.4458
2	.4757	.4044	.3985	.4156	.4377	.3833	.4244	.4194
3	.4613	.4405	.4459	.4545	.4638	.3944	.4263	.4410
4	.4987	.4655	.3750	.4296	.5209	.3941	.4459	.4471
5	.5282	.4574	.4087	.3706	.4665	.3425	.3913	.4236
6	.5138	.4326	.4167	.4103	.4771	.4019	.3842	.4338
b_{yx}	+.0089	+.0043	-.0021	-.0127	+.0079	-.0011	-.0105	-.0012
S.E.	.0048	.0057	.0055	.0211	.0068	.0056	.0044	.0030
t	1.85	.75	.38	.60	1.16	.20	2.39	.40
p	.2-.1	.5-.4	.8-.7	.6-.5	.3-.2	.9-.8	.1-.05	$\approx .7$

TABLE 2. Comparison of male and female gene frequency estimates in cage samples. Standard errors computed as $\sqrt{q_i(1 - q_i)(1/N_m + 1/N_f)}$ where q_i is the frequency in the sexes combined.

Cage No.	Sample No.	N ♂	N ♀	q ♂	q ♀	q ♂ - q ♀	S.E.	t
I	1	863	744	.489	.493	-.004	.025	.2
	2	89	137	.500	.460	+.040	.068	.6
	3	211	151	.453	.474	-.021	.053	.4
	4	206	176	.507	.489	+.019	.051	.4
	5	108	105	.472	.586	-.114	.068	1.7
	6	121	132	.484	.542	-.058	.063	.9
II	1	136	125	.445	.476	-.031	.062	.5
	2	86	139	.436	.385	+.051	.067	.8
	3	153	158	.464	.418	+.046	.056	.8
	4	181	109	.472	.454	+.018	.061	.3
	5	92	96	.446	.469	-.023	.073	.3
	6	92	101	.451	.416	+.035	.071	.5
III	1	132	102	.409	.441	-.032	.065	.5
	2	112	159	.393	.403	-.010	.062	.2
	3	155	104	.445	.447	-.002	.063	.3
	4	121	143	.401	.353	+.048	.060	.8
	5	110	109	.400	.417	-.017	.066	.3
	6	84	84	.446	.387	+.059	.076	.8
IV	1	130	115	.477	.457	+.020	.064	.3
	2	81	162	.383	.432	-.049	.067	.7
	3	238	169	.450	.462	-.012	.050	.2
	4	195	146	.464	.384	+.080	.054	1.5
	5	99	98	.369	.372	-.003	.069	.0
	6	85	99	.453	.374	+.079	.073	1.1
V	1	121	143	.471	.437	+.034	.062	.5
	2	111	138	.396	.464	-.068	.063	1.1
	3	173	131	.448	.485	-.037	.058	.6
	4	123	92	.500	.549	-.049	.069	.7
	5	81	83	.469	.464	+.005	.078	.0
	6	116	102	.444	.515	-.071	.068	1.0
VI	1	102	129	.402	.372	+.030	.065	.5
	2	106	194	.426	.361	+.065	.059	1.1
	3	174	186	.394	.395	-.001	.052	.0
	4	141	81	.362	.451	-.089	.068	1.3
	5	102	79	.328	.361	-.033	.071	.5
	6	73	141	.370	.418	-.048	.071	.7
VII	1	162	157	.426	.459	-.033	.056	.6
	2	97	141	.407	.436	-.029	.065	.4
	3	170	115	.409	.452	-.043	.060	.7
	4	141	81	.440	.457	-.017	.069	.2
	5	79	82	.399	.384	+.015	.077	.2
	6	64	139	.383	.385	-.002	.074	.0

of the mean shift suggests a systematic influence effective only during the early history of the cages prior to the time of the first count.

A striking variation in the sex ratio

within samples was apparent. Tests of significance of the departure from the expected proportion were applied to each sample from all cages and appear in table 3. Standard errors were based on a

theoretical value of 0.50. Of a total of 42 cases, 17 (roughly 40%) of the differences have t-values greater than 2.6 ($p = 0.01$) and 9 (about 21%) show t-values greater than 3.3 ($p = 0.001$). A t-test on the total numbers from all cages produced a value of 1.274 ($p = 0.30 - 0.20$) in support of the hypothetical fre-

TABLE 3. *Comparison of the observed proportion of males in population cage samples with a theoretical value of .50. Standard errors are based on the expected proportion.*

Cage No.	Sample No.	N	Males	O - C	S.E.	t
I	1	1607	.537	+.037	.013	3.0**
	2	226	.394	-.106	.033	3.2**
	3	362	.583	+.083	.026	3.2**
	4	382	.539	+.040	.026	1.5
	5	213	.507	+.007	.034	.2
	6	253	.478	-.022	.031	.7
II	1	261	.521	+.021	.031	.7
	2	225	.382	-.118	.033	3.6***
	3	311	.492	-.008	.028	.3
	4	290	.624	+.124	.029	4.3***
	5	188	.489	-.011	.037	.3
	6	193	.477	-.023	.036	.6
III	1	234	.564	+.064	.033	1.9*
	2	271	.413	-.087	.030	2.9**
	3	259	.599	+.099	.031	3.2**
	4	264	.458	-.042	.037	1.1
	5	219	.502	+.002	.034	.1
	6	168	.500	.000	.039	.0
IV	1	245	.531	+.031	.032	1.0
	2	243	.333	-.167	.033	5.1***
	3	407	.585	+.085	.025	3.4***
	4	341	.572	+.072	.027	2.7**
	5	197	.503	+.003	.036	.1
	6	184	.462	-.038	.037	1.0
V	1	264	.458	-.042	.031	1.4
	2	249	.446	-.054	.032	1.7
	3	304	.569	+.069	.029	2.4*
	4	215	.572	+.072	.034	2.1*
	5	164	.494	-.006	.039	.2
	6	218	.532	+.032	.034	1.0
VI	1	231	.442	-.058	.033	1.8
	2	300	.353	-.147	.029	5.1***
	3	360	.483	-.017	.026	.6
	4	222	.635	+.135	.034	4.0***
	5	181	.564	+.064	.037	1.7
	6	214	.341	-.159	.034	4.7***
VII	1	319	.508	+.008	.028	.3
	2	238	.408	-.092	.032	2.9**
	3	285	.597	+.097	.030	3.2**
	4	222	.635	+.135	.034	4.0***
	5	161	.491	-.009	.039	.2
	6	203	.315	-.185	.035	5.3***
Total		11893	.506	+.006	.005	1.2

quency of 0.50 used in calculating standard errors. The signs of the deviations show that they were as nearly equally distributed about the theoretical mean as possible both within the group showing significance beyond the 1% level and among all cases.

This highly significant and wholly un-systematic disturbance may have been due to a periodicity in the rate of emergence reflecting the three day interval of vial exchange. In these stocks routine observations have shown that males develop more rapidly than females. The limited capacity of a single culture vial would probably restrict the period of effective oviposition to within a few hours of the substitution of a fresh vial for an old. If this were the case and if, as seems likely, the longevity of adults was very much decreased by, for example, the scarcity of food or the accumulation of toxic substances, the fluctuation in sex ratio observable throughout the eclosion period of any single bottle culture would affect the sex ratio of the entire cage population.

No good way of estimating the size of the populations within these cages has been found. A total count was made only in the case of the first sampling of cage number 1. Because only those flies that could be induced to leave the cage and enter fresh bottles could be counted, this census cannot be considered complete. A total of 1,607 flies was collected, however, which indicates that cages of these specifications can maintain very large populations. No answer can be made to the important question of how much fluctuation in population size occurred during the 23 weeks these cages were maintained.

The Selection Experiments

The results of the selection experiments appear in table 4. The type of mating, the observed and expected segregations, their differences, standard errors and t-values are presented. Standard errors were based on the distribution of the observed proportion of a single genotype

among all bottles of a given population type and have been applied only to the cases in which there were only two genotypes among the offspring.

The offspring of matings within group I which was designed to test for differences in viability among the several genotypes show no significant departures from expectation. It is therefore inferred that the results of the productivity mating types can be interpreted assuming no differential viability.

It is evident that with the exception of population types II C a, III C a and III C c no significant departures from expectation obtained in the productivity series. In these three cases, however, the deviations were highly significant with t-values of 3.0, 3.3 and 3.3 respectively.

These results taken at face value indicate that there is no difference in the productivity of bw^{75}/bw^{75} and bw^{75}/bw females or of bw^{75}/bw and bw/bw females in "competition" for any single type of male. (Competition is not meant to imply any preconception with regard to the factors operative in this situation.) Similarly, when the competing individuals in these categories are males there are no apparent differences between the alternatives in frequency or effectiveness of insemination of any single female type. However, when bw^{75}/bw^{75} and bw/bw females compete the latter are the more productive when the male type is bw^{75}/bw^{75} but no difference is apparent when the male type is either bw^{75}/bw or bw/bw . When the productivities of bw^{75}/bw^{75} and bw/bw males are compared, the bw^{75}/bw^{75} males leave more offspring when the females are either bw^{75}/bw^{75} or bw/bw but show no significant difference from bw/bw when competing for bw^{75}/bw females.

If these differences be accepted as valid evidence of the situation their expression is obviously highly dependent upon the specific pair of genotypes of the competing individuals as well as on the genotype presented by the opposite sex. No systematic explanation suggests itself upon

TABLE 4. *Observed segregations in selection experiments.* Standard errors are based on the variance of the observed distributions of single F₁ phenotypes and are applied only to two-class segregations.

Mating type	Parental population genotype										S.E.	t				
	Females					Males										
	$\frac{bw^{7/8}}{bw^{7/8}}$	$\frac{bw}{bw}$	$\frac{bw^{7/8}}{bw^{7/8}}$	$\frac{bw}{bw}$	$\frac{bw^{7/8}}{bw^{7/8}}$	$\frac{bw}{bw}$	$\frac{bw^{7/8}}{bw^{7/8}}$	$\frac{bw}{bw}$	$\frac{bw^{7/8}}{bw^{7/8}}$	$\frac{bw}{bw}$						
I-A	8		8		8		8		8		59	.487	+.013	-.013	.009	1.4
I-B	8		8		8		8		8		58	.510	-.005	+.010	.009	1.4
I-C	8		8		8		8		8		60	.504	+.004	-.004	.013	0.3
II-A-a	4	4	4	4	4	4	4	4	4	4	58	.770	-.020	+.020	.011	1.8
II-A-b	4	4	4	4	4	4	4	4	4	4	59	.499	+.001	-.001	.000	.0
II-A-c	4	4	4	4	4	4	4	4	4	4	58	.251	+.001	-.001	.012	.0
II-B-a	4	4	4	4	4	4	4	4	4	4	58	.733	-.017	+.017	.015	1.1
II-B-b	4	4	4	4	4	4	4	4	4	4	58	.497	-.012	-.003	+.015	.7
II-B-c	4	4	4	4	4	4	4	4	4	4	55	.739	-.011	+.011	.015	.7
II-C-a	4	4	4	4	4	4	4	4	4	4	54	.564	+.064	-.064	.021	3.0*
II-C-b	4	4	4	4	4	4	4	4	4	4	(80)	(.541)	(+.041)	(-.041)	(.022)	(1.9)
II-C-c	4	4	4	4	4	4	4	4	4	4	59	.231	-.019	+.008	.020	.0
III-A-a	8		4	4	4	4	4	4	4	4	60	.501	+.001	-.001	(.018)	(0.4)
III-A-b	8		4	4	4	4	4	4	4	4	(81)	(.493)	(-.007)	(+.007)	.015	1.0
III-A-c	8	8	4	4	4	4	4	4	4	4	59	.735	+.015	-.015	.015	.0
III-B-a	8	8	4	4	4	4	4	4	4	4	59	.485	+.005	-.015	.015	0.9
III-B-b	8	8	4	4	4	4	4	4	4	4	60	.263	+.013	-.013	.015	0.6
III-B-c	8	8	4	4	4	4	4	4	4	4	60	.740	-.010	+.010	.015	.0
III-C-a	8	8	4	4	4	4	4	4	4	4	49	.499	.000	-.001	.017	0.6
III-C-b	8	8	4	4	4	4	4	4	4	4	59	.722	+.028	-.028	.016	1.8
III-C-c	8	8	4	4	4	4	4	4	4	4	(80)	(.502)	(+.002)	(-.002)	.023	3.4*
	8	8	4	4	4	4	4	4	4	4	55	.244	-.006	+.010	(.019)	(0.1)
	8	8	4	4	4	4	4	4	4	4	60	.596	+.096	-.096	.030	3.2*
	8	8	4	4	4	4	4	4	4	4	(82)	(.496)	(-.004)	(+.004)	(.032)	(0.2)

examination of these data. There is no self-determined consistency in the behavior of any genotype.

To furnish further evidence on the segregations in categories II C and III C that would either lend credence to or discredit the initial observations, an additional number of replicates, somewhat larger than that in the initial series, was made of categories II C a, II C c, III C a, and III C c. The values obtaining from these observations appear in table 4 in parenthesis below the initially determined values. No evidence of significant deviations appears among them. Although it seems unlikely that the cases of significance in the first set of observations were entirely fortuitous, the fact that no more than five replicates of any population type were made at any given time substantially rules out the possibility that the groups in which significant deviations were observed differed systematically from other groups in environmental conditions. The factors responsible for these shifts remain obscure.

To check against possible heterogeneity among groups of replicates made at dif-

ferent times, an analysis of variance comparing the variability within and between groups of the same population type was carried out. This test is not strictly applicable to data of a binomial character because of the relationship between the mean (np) and variance (npq) but should serve as a fair test in this case since no comparisons have been made between population categories having different theoretical means and these are in no case higher than 0.75 or lower than 0.25 at which values skewness and kurtosis are negligible for such values of n as occur in these cases.

The results of this analysis are presented in table 5. None of the F values was found to be significant beyond the 1% level and only two had probabilities between 0.01 and 0.05. In one of the cases showing probable significance, the variance between groups was smaller than that within groups, which suggests no heterogeneity. The total variances cannot have been appreciably affected by variation in environmental conditions during the period of observations.

The data from the additional replicates

TABLE 5. *Analysis of variance comparing variability within and between groups of replicates in selection experiments. All values on an absolute scale.*

Mating type		D.F.	Mean square	F	Mating type		D.F.	Mean square	F
I-A	within	48	9.1	1.8	I-C	within	47	5.0	1.3
	between	11	16.0			between	11	3.8	
II-A-a	within	46	7.2	1.4	II-A-c	within	46	8.2	1.0
	between	11	5.2			between	11	8.4	
II-B-a	within	46	11.4	1.7	II-B-c	within	44	12.7	1.2
	between	11	19.8			between	10	15.4	
II-C-a	within	43	25.9	1.1	II-C-c	within	48	24.8	1.1
	between	10	23.0			between	11	28.2	
III-A-a	within	48	14.8	2.8*	III-A-c	within	48	14.6	1.0
	between	11	5.2			between	11	14.9	
III-B-a	within	48	15.0	2.3*	III-B-c	within	47	14.2	1.5
	between	11	34.5			between	11	21.6	
III-C-a	within	46	34.3	1.4	III-C-c	within	48	51.0	1.0
	between	11	24.7			between	11	52.2	

TABLE 6. Comparison of total variance between appropriate categories in productivity matings of selection experiments. Variance ratio interpreted on F table of Snedecor. All values are on an absolute scale.

Mating type	d.f.	Mean square	F	Mating type	d.f.	Mean square	F
II-A-a	57	6.8	1.2	III-A-a	58	13.0	1.1
II-A-c	57	8.3		III-A-c	59	14.7	
II-B-a	57	13.1	1.0	III-B-a	59	18.7	1.2
II-B-c	54	13.2		III-B-c	58	15.6	
II-C-a	53	25.4	1.0	III-C-a	57	32.4	1.6*
II-C-c	59	25.4		III-C-c	59	51.2	
II-C-a'	79	38.8	1.4	III-C-a'	79	28.8	1.6*
II-C-c'	80	28.3		III-C-c'	81	45.1	
(Comparison between categories II and III)							
II-A-a	57	6.8	1.9**	II-A-c	57	8.3	1.8*
III-A-a	59	13.0		III-A-c	59	14.7	
II-B-a	57	13.1	1.4	II-B-c	54	13.2	1.2
III-B-a	59	18.7		III-B-c	58	15.6	
II-C-a	53	25.4	1.3	II-C-c	59	25.4	2.0**
III-C-a	57	32.4		III-C-c	59	51.2	
II-C-a'	79	38.8	1.3	II-C-c'	80	28.3	1.6*
III-C-a'	79	28.8		III-C-c'	81	45.1	
(Comparison between categories A and B)							
II-A-a	57	6.8	1.9**	III-A-a	59	13.0	1.4
II-B-a	57	13.1		III-B-a	59	18.7	
II-A-c	57	8.3	1.6*	III-A-c	59	14.7	1.1
II-B-c	54	13.2		III-B-c	58	15.6	

of the population types in categories II C and III C made outside the body of the main experiment could not be included in this analysis since they were collected over a relatively short period of time and were from series of replicates that were grouped unequally and of larger size than the standard of five set for the initial study. The lack of evidence of heterogeneity between groups in the initial experiment indicates that this difference in treatment is not likely to have affected to an appreciable extent the outcome of the later study.

The results of population types presenting three classes of F_1 progeny show only small and unsystematic deviations from

expectation. In the absence of clear evidence of selection from the other categories, no analysis of these types seems necessary.

It is of interest to compare the variability of the observed phenotypic frequencies among the offspring of certain comparable population types both within and between male and female productivity series. The appropriate comparisons are presented in table 6. A rough test of the significance of the variance differences is given by the variance ratio interpreted as F.

Among the comparisons of a and c categories (bw^{75}/bw^{75} vs bw/bw as the one-genotype parental group), the ratios are in all but one case clearly not significantly

different from unity. The ratio is probably significant, $p = 0.05 - 0.01$, in the comparison of III C a' with III C c', and approaches significance at the 5% level in the corresponding case in the initial survey. But, since there is no consistency in the sign of the variance differences in the other comparisons of this type within the male productivity matings, there is not sufficient evidence to assign this affect to the productivity variation of a particular genotype.

With regard to the comparisons between categories II and III (reciprocal mating types), the variances of the observed segregations are in all but one case greater in male productivity population tests than in female. In the comparison of classes II A c and III A c the variance ratio is significant with a probability between 0.05 and 0.01. Further, within the initial series of matings, the comparison of the variance of II C c with III C c yields an F-value significant beyond the 1% level. The variances of the segregations in the groups of the subsidiary study are closely comparable to those of the initial series in the cases of II C c', III C a' and III C c' but are not in such good agreement in the case of II C a'. The comparison of II C a' with III C a' is the only case in which the female productivity variance was found to be higher than the male in comparable categories.

The most interesting differences appear in the comparisons between A and B categories (tests of $bw/bw : bw^{75}/bw$ competitions vs those of the $bw^{75}/bw^{75} : bw^{75}/bw$ competitions). Both such comparisons show significant differences in the case of females. The ratio of the variances of II A a and II B a is 1.913 which is significant at about the 1% level, that of the variances of II A c and II B c is 1.602 and significant at about the 5% level. Neither of the comparisons within male categories is significant; although the III B categories have higher variances than the III A categories in absolute terms, the differences are slight.

The difference in variance between the

A and B categories in females is associated with the presence of the bw^{75}/bw^{75} genotype. These data suggest that this genotype is considerably more variable in its productivity than either of the other female genotypes, a fact that may be of some importance to the discussion in a later section.

It should be pointed out that the observed segregation variances are not directly related to the variation in productivity of the competing parents in the cases of the four C categories. In all such cases at least one of the segregation genotypes is produced by both of the parental types, male or female, under consideration. The variances of the segregations in the C categories would, therefore, be expected to be larger for this reason alone. Since the segregations in the C categories have theoretical means of 0.50 while those of A and B categories with two-class segregations have theoretical means of 0.25 or 0.75, the variances of C categories would be expected to be about 1/3 again as large as those in either A or B in samples of equal size, a variance ratio of 1.33. Tests of significance would accordingly be inappropriate in these cases. All ratios involving C categories were very much greater than 1.33, ranging from 1.737 to 3.717.

RANDOM DRIFT EXPERIMENTS

Selection

To get evidence on the question of selection from within the body of the experimental data of the random drift experiments, tests of significance were applied in each generation to the mean differences in frequency between bottles at the time of mating and at the time of sampling. Standard errors were calculated from the variances of the observed distributions of these differences among bottles. The mean change in frequency may be represented by $\bar{\Delta}q$. These data appear in tables 7 and 8 for random drift series I and II respectively, and are presented graphically in figure 1.

TABLE 7. *Series I. Gene frequency changes by generation; q_t = total frequency including previously fixed cultures, q_m = frequency among unfixed cultures at the time of mating, q_c = frequency among unfixed and newly fixed cultures at the time of sampling. Standard errors are computed from the observed variance of Δq among bottles in each generation.*

Gen.	q_t	N	q_m	q_c	$\overline{(\Delta q)}$ $q_c - q_m$	S.E. $\overline{\Delta q}$	t
0	.500	107	.500	.506	+.006	.008	.8
1	.506	107	.506	.494	-.012	.011	1.1
2	.494	107	.494	.495	+.001	.012	.1
3	.495	107	.495	.490	-.005	.011	.5
4	.490	106	.486	.489	+.003	.010	.3
5	.494	105	.484	.459	-.025	.010	2.5*
6	.469	103	.458	.471	+.013	.010	1.3
7	.482	101	.481	.479	-.002	.010	.2
8	.480	97	.478	.473	-.005	.012	.4
9	.476	96	.468	.477	+.009	.012	.8
10	.484	92	.476	.486	+.010	.012	.8
11	.492	86	.496	.504	+.008	.010	.8
12	.499	78	.466	.462	-.004	.011	.4
13	.495	77	.455	.452	-.003	.012	.3
14	.493	71	.448	.458	+.010	.013	.8
15	.500	66	.463	.471	+.008	.011	.7
16	.505	59	.493	.466	-.027	.011	2.5*
17	.490	55	.481	.463	-.018	.012	1.5
18	.481	52	.451	.432	-.019	.010	1.9
19	.472	—	(.432)	—	—	—	—
0-19	.4899	1672	.4796	.4778	-.0018	.0025	.72
1-19	.4894	1565	.4782	.4759	-.0023	.0027	.85

The data from series I contain little evidence of selection. The t-values indicate decreases in the frequency of bw^{75} between generations 5 and 6 and 16 and 17 that are significant with probabilities of 0.009 and 0.02 respectively. ("generation" is used here to refer to the parental population alone.) Grouping all shifts in frequency between generations, the value of $\overline{\Delta q}$ is not significant, $p = 0.48$. However, since the deviations between generations 0 and 1 were not subject to the influence of variations in productivity they should perhaps be omitted in computing the overall value of $\overline{\Delta q}$. Omitting these deviations the value of $\overline{\Delta q}$ is only slightly different and of no significance, $p = 0.4 - 0.3$. There is no evidence of any persistent trend in the sign of $\overline{\Delta q}$ in the course of the experiment and the array of t-values does not suggest any heterogeneity between data collected at different times.

Frequencies of bw^{75} involve, in the case of mating frequencies, only unfixed cultures, and in the case of collection frequencies, only unfixed and newly fixed populations. It is of interest to note the trend in total gene frequency change considering in each generation previously fixed as well as newly fixed and unfixed classes. These frequencies for both series appear under q_t in tables 7 and 8 and in figure 1.

No systematic trend is apparent in the case of series I. The mean total frequency over all generations is 0.4899 and the ratio of fixation at the two extremes is 30 (bw) : 28 (bw^{75}), approaching equality as would be expected in the absence of selection.

In the case of series II (table 8 and fig. 1) there was a consistent increase in the frequency of bw^{75} during the first 4 generations. The values of $\overline{\Delta q}$ between generations 2 & 3 and 3 & 4 are both

positive and significant at about the 2% and 1% levels respectively. The value of $\overline{\Delta q}$ between generations 17 & 18 is significant at about the 2% level and that taken over the entire sequence of generations is significant at the 5% level. However, omitting the changes between generations 0 and 1 for the same reason as was given in connection with series I, the overall value is somewhat less and of no significance, $t = 1.83, p = 0.10 - 0.05$. The overall shift must be regarded as of doubtful significance.

The total frequency in each generation including previously fixed classes rises during the first four generations to a value of 0.5711 (fig. 1). This frequency subsequently fluctuates somewhat but remains fairly stable. The mean total frequency over the last 16 generations is 0.5674. The gradual drop in frequency in the array of unfixed and newly fixed

populations after the 4th generation is due wholly to a higher rate of fixation for bw^{75} than for bw in this series rather than to consistent negative selection over this period. The net change in gene frequency attributable to selection between generations 5 and 19 is, in fact, positive (+.0083). The ratio of fixation at the two extremes is 13 (bw); 30 (bw^{75}), $t = 3.08, p = 0.01 - 0.001$. This asymmetry is clearly related to the steady increase in the value of $q_{bw^{75}}$ over the first four generations.

When the observed sample frequencies of each series are classified according to the mating frequency rather than by generation and the means of the resulting recipient class distributions are compared with their respective theoretical values, no evidence of selection is found in either case. Table 9 presents these data for series I and II. The standard errors

TABLE 8. *Series II. Gene frequency changes by generation; q_t = total frequency including previously fixed cultures, q_m = frequency among unfixed cultures at the time of mating, q_c = frequency among unfixed and newly fixed cultures at the time of sampling. Standard errors are computed from the observed variance of Δq among bottles in each generation.*

Gen.	q_t	N	q_m	q_c	$\overline{(\Delta q)}$ $q_c - q_m$	S.E. $\overline{\Delta q}$	t
0	.500	105	.500	.509	+.009	.009	1.0
1	.509	105	.509	.522	+.013	.010	1.3
2	.522	105	.522	.546	+.024	.010	2.4*
3	.546	105	.546	.571	+.025	.009	2.8**
4	.571	104†	.570	.569	-.001	.010	.1
5	.569	103	.564	.575	+.011	.010	1.1
6	.579	102	.571	.560	-.011	.008	1.4
7	.569	100	.552	.547	-.005	.010	.5
8	.565	95	.534	.535	+.001	.010	.1
9	.566	93	.536	.541	+.005	.010	.5
10	.570	91	.542	.547	+.005	.009	.6
11	.575	90	.542	.539	-.003	.010	.3
12	.572	88	.529	.523	-.006	.011	.5
13	.568	84	.525	.522	-.003	.011	.3
14	.566	79	.492	.477	-.015	.010	1.5
15	.555	76	.483	.486	+.003	.010	.3
16	.557	74	.473	.469	-.004	.010	.4
17	.555	69	.488	.511	+.023	.010	2.3*
18	.570	61	.480	.484	+.004	.011	.4
19	.572	—	(.484)	—	—	—	—
0-19	.5578	1729	.5273	.5317	+.0044	.0022	2.00*
1-19	.5608	1624	.5290	.5332	+.0042	.0023	1.83

† Accidental loss.

were computed as $\sqrt{\sigma_o^2/N}$. Although a standard error of this sort takes no account of the expected asymmetry of the distributions at the extremes of frequency, examination of the array of *t*-values together with the signs of the differences between observed and theoretical means indicates that there can be little doubt as to the points of significance. In each series only three of the differences are significant, none beyond the 1% level. There are no consistent trends in sign in either case and it can therefore be safely assumed that any changes in gene frequency attributable

to selection in either series are wholly negligible in this context.

The Variance of Recipient Class Gene Frequencies

In the absence of appreciable selection and assuming random mating and Poisson variability in productivity of parents (cf. Crow, 1955) the variances of the recipient class frequencies grouped about each of the 31 donor class frequencies as means should be binomial in character and equal to $q(1-q)/2N_e$, where $2N_e$ equals the effective sample size in terms

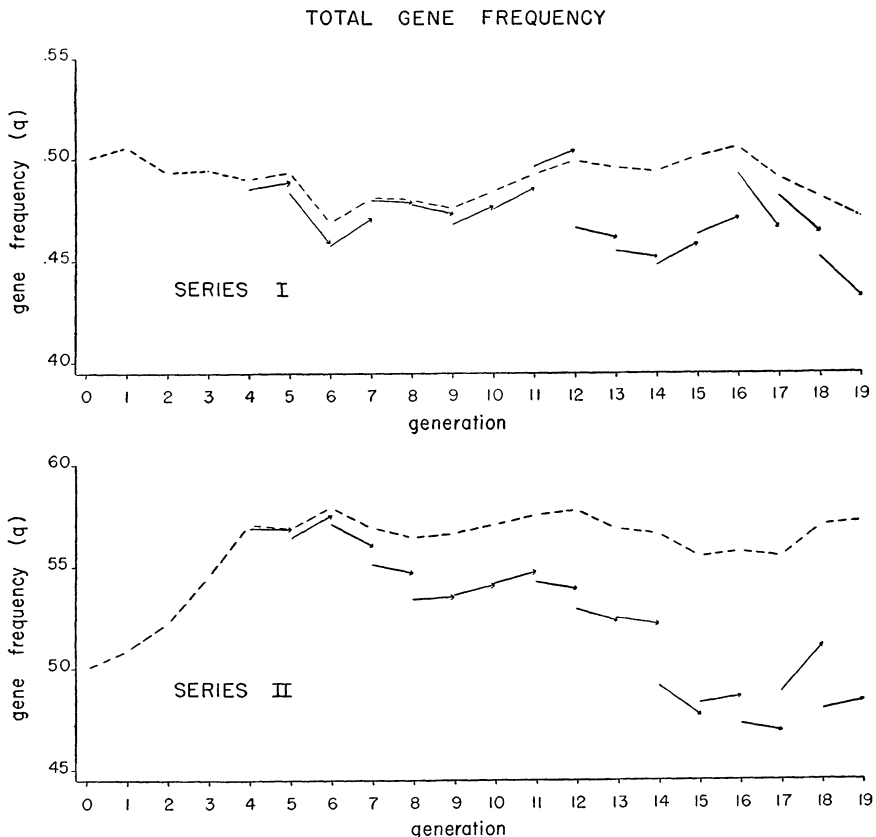


FIG. 1. Total gene frequency by generation for series I and II. Frequency including previously fixed as well as newly fixed cultures is indicated by the dashed line. The arrows represent the changes in gene frequency between generations among unfixed bottles. The break in this line at any generation is brought about by frequency changes due to unequal fixation for the two alleles.

TABLE 9. *Series I and II. Comparison of observed with theoretical recipient class means; $\bar{q}_{th} = 1/32$ class value. Standard errors are based on the variance of q_0 about q_{th} .*

Class value	Series I					Series II				
	n	\bar{q}_0	$\bar{q}_0 - \bar{q}_{th}$	S.E.	t	n	\bar{q}_0	$\bar{q}_0 - \bar{q}_{th}$	S.E.	t
1	25	.018	-.014	.006	2.3*	18	.040	+.009	.011	.8
2	25	.081	+.019	.012	1.6	24	.066	+.004	.008	.5
3	43	.088	-.006	.010	.6	20	.072	-.022	.013	1.7
4	38	.119	-.006	.012	.5	22	.124	-.001	.015	.1
5	31	.137	-.019	.013	1.5	20	.158	+.002	.017	.1
6	46	.166	-.021	.011	1.9	29	.193	+.005	.018	.3
7	45	.251	+.032	.018	1.8	44	.212	-.006	.011	.5
8	47	.242	-.008	.011	.7	38	.273	+.023	.013	1.8
9	64	.284	+.003	.013	.2	46	.295	+.014	.014	1.0
10	52	.311	+.002	.016	.1	53	.315	+.003	.014	.2
11	60	.338	-.006	.015	.4	55	.339	-.005	.012	.4
12	69	.382	-.007	.011	.6	61	.385	+.010	.013	.8
13	79	.415	+.009	.014	.6	76	.411	+.005	.012	.4
14	78	.435	-.002	.012	.2	90	.447	+.009	.011	.8
15	95	.478	+.009	.012	.8	94	.493	+.024	.010	2.4*
16	109	.494	-.006	.011	.5	89	.511	+.011	.012	.9
17	83	.540	+.008	.013	.6	102	.533	+.002	.010	.2
18	86	.537	-.025	.014	1.8	88	.557	-.006	.011	.5
19	66	.565	-.029	.014	2.1*	93	.580	-.014	.011	1.3
20	63	.622	-.003	.013	.2	66	.612	-.013	.012	1.1
21	46	.649	-.007	.020	.3	52	.665	+.009	.015	.6
22	45	.687	-.001	.017	.1	59	.702	+.014	.014	1.0
23	35	.674	-.045	.023	2.0*	55	.722	+.003	.014	.2
24	39	.776	+.026	.016	1.6	62	.749	-.002	.011	.2
25	33	.771	-.010	.019	.5	33	.791	+.010	.017	.6
26	35	.835	+.022	.017	1.3	45	.811	-.001	.013	.1
27	27	.854	+.011	.015	.7	38	.849	+.005	.013	.4
28	32	.869	-.006	.014	.4	44	.893	+.018	.009	2.0*
29	22	.915	+.009	.016	.6	39	.911	+.005	.009	.6
30	25	.929	-.009	.015	.6	42	.936	-.002	.006	.3
31	22	.977	+.009	.023	.4	27	.982	+.013	.006	2.2*

of genes, theoretically = 32 in the present case. Estimates of $2N_e$ can be made for each class of mating frequency as $q(1-q)/\sigma_o^2$, where σ_o^2 is the observed recipient class variance.

The magnitude of the effect of random sampling in bringing about gene frequency changes is inversely proportional to $2N_e$, and directly proportional to $q(1-q)$. Under circumstances in which the effective population number fluctuates in time an overall estimate of $2N_e$ may be made as the harmonic mean of the values at successive intervals (Wright 1939). In the present case, in which each class estimate is itself the harmonic mean of the values by generation, the 31 class estimates may be combined in this

way if it is assumed that $2N_e$ is itself independent of q :

$$2\bar{N}_e = N/\sum[1/2N_e]$$

Values of 17.86 and 22.99 were obtained for series I and II respectively.

These estimates of the average value of $2N_e$ for each of the two series can be used as a means of judging the homogeneity of the array of observed variances with respect to this parameter. Tables 10 and 11 present a comparison of these variances with theoretical values computed for each class as $q(1-q)/2N_e$. Standard errors are based on the theoretical variances and take into account the degree of kurtosis expected at each

value of q :

$$S.E. = \sqrt{q(1 - q)[4\bar{N}_e q(1 - q) + 1 - 6q(1 - q)]/n(2\bar{N}_e)^3}$$

(adapted from Fisher, 1925), where \bar{N}_e = the effective sample number and n = the number of samples.

The data of series I (table 10) are markedly heterogeneous; 6 of the 31 t -values are greater than 1.96. In the case of series II there is no evidence of significant discrepancy (table 11). Graphical comparisons (figs. 2 and 3) indicate that the data of series I are sensibly in poor accord with expectation as contrasted with series II. In particu-

lar, the observed distribution is decidedly skewed, indicating a tendency toward smaller values of $1/2N_e$ at lower values of q . It may be noted that there is a suggestion of positive skewness in the graph of the recipient class variance data of series II. Equations of the linear regression of class estimates of $1/2N_e$ on q were computed for both series, weighting each value by the total frequency in the recipient class distribution upon which each is based:

I: $1/2N_e = \sigma^2_{\delta q}/q(1 - q) = .0451 + .0228q, S.E._b = .0014, t = 16.3$

II: $1/2N_e = \sigma^2_{\delta q}/q(1 - q) = .0427 - .0020q, S.E._b = .0009, t = 2.2$

TABLE 10. *Series I. Comparison of observed recipient class variances with theoretical values based on a common estimate of $2N_e = 18$. Class value = $32q_{th}$. All values are presented on a relative scale.*

Class value	Obs.	Th.	(Obs. - Th.)	S.E.	t
1	.0010	.0017	-.0007	.0006	1.2
2	.0038	.0033	+.0005	.0011	.5
3	.0043	.0047	-.0004	.0011	.4
4	.0058	.0061	-.0003	.0015	.2
5	.0051	.0073	-.0022	.0019	1.2
6	.0053	.0085	-.0032	.0018	1.8
7	.0139	.0095	+.0044	.0020	2.2*
8	.0059	.0104	-.0045	.0021	2.1*
9	.0103	.0112	-.0009	.0020	.5
10	.0126	.0119	+.0007	.0023	.3
11	.0132	.0125	+.0007	.0022	.3
12	.0079	.0130	-.0051	.0022	2.3*
13	.0160	.0134	+.0026	.0021	1.2
14	.0112	.0137	-.0025	.0021	1.2
15	.0146	.0138	+.0008	.0020	.4
16	.0131	.0139	-.0008	.0018	.4
17	.0150	.0138	+.0012	.0021	.6
18	.0164	.0137	+.0027	.0020	1.4
19	.0121	.0134	-.0013	.0023	.6
20	.0109	.0130	-.0021	.0023	.9
21	.0176	.0125	+.0051	.0026	2.0*
22	.0137	.0119	+.0018	.0025	.7
23	.0186	.0112	+.0074	.0026	2.8**
24	.0102	.0104	-.0002	.0023	.1
25	.0113	.0095	+.0018	.0023	.8
26	.0097	.0085	+.0012	.0020	.6
27	.0063	.0073	-.0010	.0020	.5
28	.0061	.0061	+.0000	.0016	.0
29	.0054	.0047	+.0007	.0015	.5
30	.0054	.0033	+.0021	.0011	1.9*
31	.0012	.0017	-.0005	.0007	.7

TABLE 11. *Series II. Comparison of observed recipient class variances with theoretical values based on a common estimate of $2N_e = 23$. Class value = $32q_{th}$. All values are presented on a relative scale.*

Class value	Obs.	Th.	(Obs. - Th.)	S.E.	t
1	.0022	.0013	+.0009	.0006	1.5
2	.0017	.0025	-.0008	.0008	1.0
3	.0031	.0037	-.0006	.0012	.5
4	.0048	.0048	-.0000	.0015	.0
5	.0059	.0057	+.0002	.0018	.1
6	.0097	.0066	+.0031	.0018	1.7
7	.0056	.0074	-.0018	.0016	1.1
8	.0062	.0082	-.0020	.0019	1.0
9	.0086	.0088	-.0002	.0018	.1
10	.0106	.0093	+.0013	.0018	.7
11	.0074	.0098	-.0024	.0018	1.3
12	.0099	.0102	-.0003	.0018	.2
13	.0115	.0105	+.0010	.0017	.5
14	.0114	.0107	+.0007	.0016	.4
15	.0095	.0108	-.0013	.0015	.9
16	.0133	.0109	+.0024	.0016	1.5
17	.0094	.0108	-.0014	.0015	.9
18	.0099	.0107	-.0012	.0016	.8
19	.0120	.0105	+.0015	.0015	1.0
20	.0100	.0102	-.0002	.0017	.1
21	.0112	.0098	+.0014	.0019	.7
22	.0108	.0093	+.0015	.0017	.9
23	.0103	.0088	+.0015	.0017	.9
24	.0080	.0082	-.0002	.0015	.1
25	.0090	.0074	+.0016	.0018	.9
26	.0071	.0066	+.0005	.0014	.4
27	.0066	.0057	+.0009	.0013	.7
28	.0034	.0048	-.0014	.0010	1.4
29	.0034	.0037	-.0003	.0009	.3
30	.0017	.0025	-.0008	.0006	1.3
31	.0010	.0013	-.0003	.0005	.6

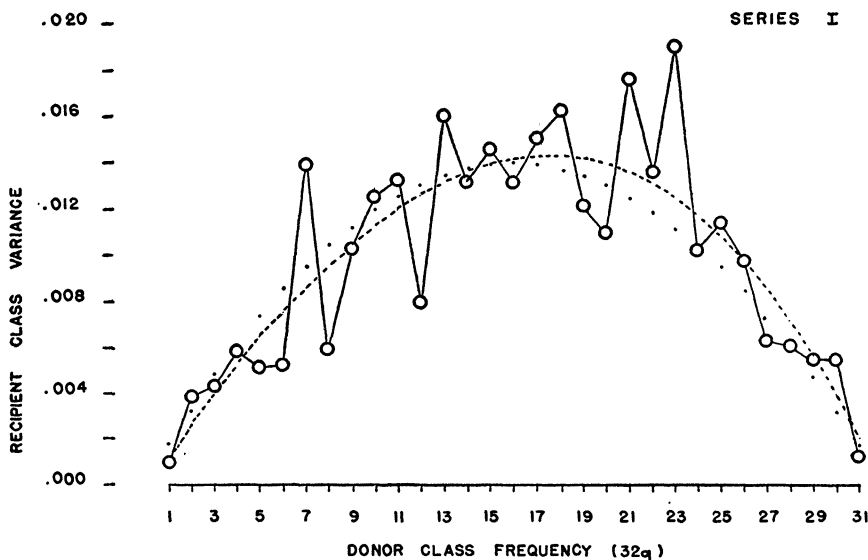


FIG. 2. Comparison of the observed recipient class variances of series I with theoretical values based upon the mean estimates of $1/2N_e$ (dots) and with values based upon the regression equation representing the observed dependence of $1/2N_e$ on q (dashed line).

The effect is highly significant in the case of series I (fig. 4). Considering the graphical presentation of the series II data (fig. 5) no significance can be ascribed to the slight negative trend in this case.

Theoretical values of the variance may be computed for series I by transforming

the above regression formula:

$$\sigma^2_{\delta q} = q(1-q)(1-.506q)/22.18.$$

Comparison of the observed class variances with these values results in considerably better agreement with the observations than is the case using a common estimate (table 12); standard errors were

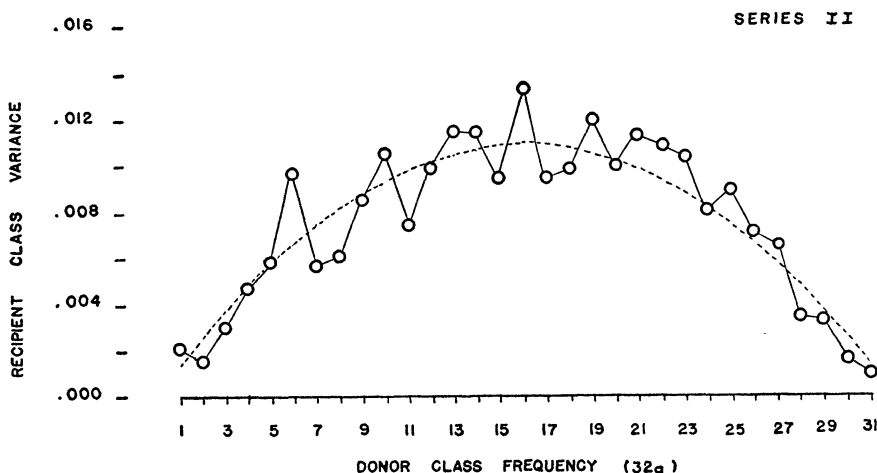


FIG. 3. Comparison of the observed recipient class variances of series II with theoretical values based upon the mean class estimates of $1/2N_e$ (dashed line).

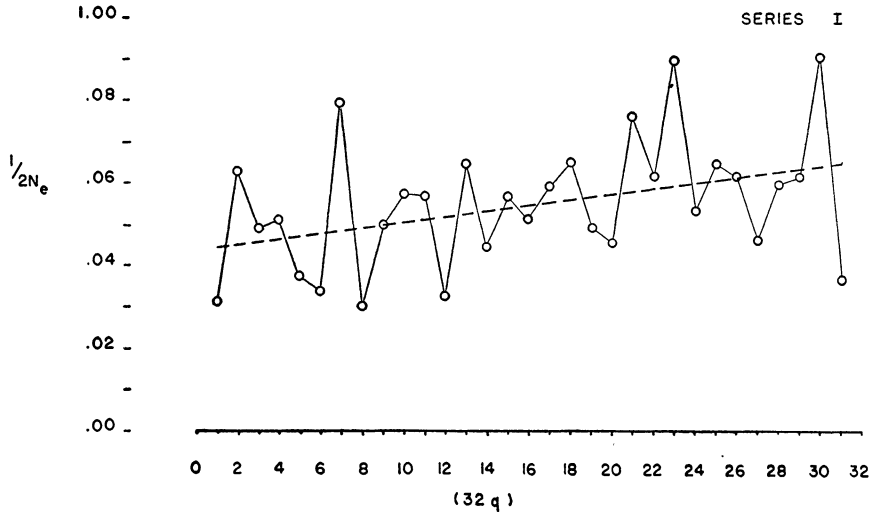


FIG. 4. Observed regression of $1/2N_e$ on q in series I.

computed from the theoretical $2N_e$ values as before. Taking $2N_e$ as a function of q , only three of the points of significance remain, two with probabilities less than .01. The class estimates of series I seem somewhat more variable than would be expected assuming the hypothesis is adequate; however, examination of figure 2 indicates clearly that it is unlikely a refinement of the theory would bring the points of major discrepancy into line. There is apparently some real heterogeneity of an unspecified sort among the frequency classes in this case.

The evident dependence of $2N_e$ on q indicates that the distributions by class are not to be considered purely binomial but rather compounded of binomial elements with variation in productivity associated with genotype.

The Distribution of Gene Frequencies Among Bottles

The histories of the random drift series I and II presented in terms of the distribution of gene frequency by generation for each series appears in tables 14 and 15

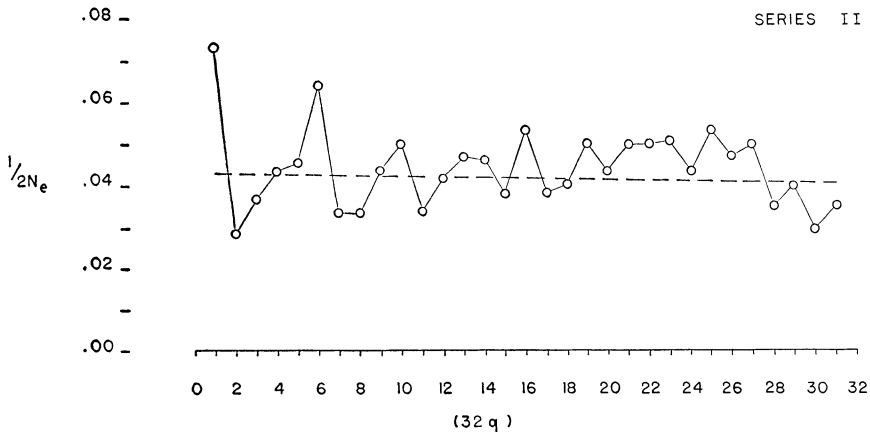


FIG. 5. Observed regression of $1/2N_e$ on q in series II.

TABLE 12. *Series I. Comparison of observed recipient class variances with theoretical values obtained from the calculated linear regression of class estimates of $1/2N_e$ on frequency. Class value = $32q_{th}$. All values are presented on a relative scale.*

Class value	Obs.	Th.	Diff.	S.E.(th)	t
1	.0010	.0014	-.0004	.0006	.7
2	.0038	.0027	+.0011	.0009	1.2
3	.0043	.0040	+.0003	.0014	.2
4	.0058	.0052	+.0006	.0016	.4
5	.0051	.0064	-.0013	.0021	.6
6	.0053	.0075	-.0022	.0020	1.1
7	.0139	.0086	+.0053	.0018	2.9**
8	.0059	.0095	-.0036	.0022	1.6
9	.0103	.0104	-.0001	.0021	.0
10	.0126	.0112	+.0014	.0021	.7
11	.0132	.0119	+.0013	.0022	.6
12	.0079	.0126	-.0047	.0022	2.1*
13	.0160	.0131	+.0029	.0021	1.4
14	.0112	.0136	-.0024	.0020	1.2
15	.0146	.0139	+.0007	.0020	.4
16	.0131	.0141	+.0010	.0021	.5
17	.0150	.0142	+.0008	.0019	.4
18	.0164	.0143	+.0021	.0021	1.0
19	.0121	.0141	-.0020	.0020	1.0
20	.0109	.0139	-.0030	.0024	1.3
21	.0176	.0135	+.0041	.0026	1.6
22	.0137	.0131	+.0006	.0024	.3
23	.0186	.0124	+.0062	.0023	2.7**
24	.0102	.0117	-.0015	.0021	.7
25	.0113	.0108	-.0005	.0026	.2
26	.0097	.0097	-.0000	.0021	.0
27	.0063	.0085	-.0022	.0020	1.1
28	.0061	.0071	-.0010	.0016	.6
29	.0054	.0056	-.0002	.0014	.1
30	.0054	.0039	+.0015	.0010	1.5
31	.0012	.0020	-.0008	.0008	1.0

and figures 6 and 7. The class values, ranging from 0-32, are gene frequencies on the absolute scale determined by the controlled sample size of 16 flies per bottle (32 genes).

The distributions of series I and II are evidently very similar. Certain differences are, however, apparent. The spread in the range of observed frequencies following the initiation of each series is somewhat more rapid in series I. The diagrams in the early generations of series II are positively skewed reflecting the consistent positive value of $\bar{\Delta}q$ over the first 4 generations as already noted. The stable form of the frequency distribu-

tion, which is theoretically approximately uniform in the absence of systematic pressures and assuming a constant value of $1/2N_e$ for all classes (Fisher 1930, Wright 1931), is sensibly approached in generation 11 in series I while in series II certainly not before generation 17.

After stability of form has been reached, the fixation rate is theoretically uniform ($1/2N_e$) (Wright, 1931). A chi-square applied to the observed total number of bottles fixed over the last 9 generations of series I, basing the calculated values on the ratio of the number of bottles observed to reach fixation to the total number exposed to fixation over this period (.0676) gives a value of $\chi^2 = 6.08$, $n = 8$, $p = 0.50 - 0.70$. The stable form in series II cannot have been represented for a number of generations sufficient to warrant consideration of the observed fixation rate.

In the case of series I there is a strong suggestion of a negative regression of frequency on class value. The grouped frequency distribution over the last 9 generations (fig. 8) bears this out. A trend line fitted to these data, omitting the newly fixed classes and ignoring the obvious non-linearity, has a coefficient, $b = -.024$, which is significant with a probability of about 0.02, $t = 2.38$, $n = 29$. Despite the non-random distribution of signs about the fitted line, this method gives a rough estimate of the significance of this tendency. The effect is clearly related to the observed dependence of $1/2N_e$ on q in this case.

The grouped frequency distribution over the last three generations of series II appears in figure 9. There is here no suggestion of slope. The stable distribution may be taken to be uniform.

Considering the probable significance of the positive value of $\bar{\Delta}q$ obtained from the whole body of the data of series II and the asymmetry of the distributions in the earlier generations it is at first sight surprising that there should have been no definite slope to the terminal distributions. However, the marked

TABLE 13. Series I. Distributions of numbers of bw⁷⁵ genes among the cultures of each generation and grouped for generations (11-19) and (1-19).

Generation	Number of bw ⁷⁵ genes																																Total fixed bw ⁷⁵	Total newly fixed	Total frequency	Fixation rate	
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31					32
0	107																																				
1	1	0	0	0	0	0	1	0	0	0	0	2	5	8	10	13	22	9	14	14	7	1	1														
2		2	0	1	1	2	3	5	4	8	7	7	7	2	15	9	8	13	7	5	3	1	2	4	1	0	2	2	3	0	1						
3			2	0	1	1	2	3	5	3	4	9	8	8	9	8	6	7	4	7	2	5	2	2	2	2	2	2	3	0	0	1					
4				2	0	2	1	5	4	6	8	6	2	5	8	9	3	3	7	4	3	3	6	4	2	2	2	2	2	0	1	2	0	1			
5				1	1	2	3	5	3	6	5	1	6	4	4	5	10	7	5	3	4	3	4	6	7	0	4	0	3	3	2	0	1	0	1	3	2
6				1	2	1	2	2	3	6	2	10	7	2	4	4	7	4	2	1	4	6	7	0	4	0	3	3	2	0	1	0	3	2	3	2	
7					2	1	0	0	1	5	7	5	2	4	3	5	12	5	5	4	5	1	1	3	1	2	5	1	2	7	1	1	0	0	3	2	
8						2	0	1	1	2	3	4	2	10	3	4	5	8	6	3	2	2	5	3	1	3	1	0	0	3	2	3	2	5	4	10	0
9							0	1	2	5	3	2	1	3	5	3	4	2	5	4	2	1	4	4	3	3	1	3	3	2	3	1	1	3	1	6	1
10								2	2	4	3	5	0	3	1	4	2	3	2	3	2	4	5	2	1	5	1	2	2	1	2	2	2	2	2	8	4
11									4	1	3	6	3	5	2	1	2	1	3	2	1	3	2	2	3	5	2	2	0	2	2	2	2	2	2	10	6
12										1	0	4	6	2	2	4	2	2	1	3	0	6	3	5	1	4	2	4	3	0	1	2	1	0	1	7	8
13											0	4	3	4	0	3	3	0	2	1	7	5	4	3	4	2	1	2	2	2	2	1	1	1	1	18	1
14												3	6	1	5	2	0	3	1	2	4	1	3	3	1	6	3	5	2	2	0	0	1	2	3	21	6
15													3	4	2	3	3	0	0	1	5	2	5	3	1	4	4	1	1	6	0	2	2	3	5	71	0.0704
16														5	0	2	4	1	2	1	2	1	2	2	3	2	3	1	2	2	1	3	2	23	5		
17															3	0	1	4	1	4	1	0	0	3	3	2	3	2	4	3	1	3	2	25	7		
18																1	3	1	4	2	0	3	1	1	4	2	3	1	3	0	1	1	2	4	59	0.0678	
19																	3	2	0	3	1	1	1	1	5	2	0	3	1	2	0	3	3	52	0.0577		
1-19	30	27	25	46	39	33	46	50	49	67	53	61	70	84	80	95	109	85	86	68	65	49	47	36	41	35	36	27	32	22	25	26	28	1672			
11-19	23	21	13	31	25	19	21	17	15	20	15	16	20	33	20	23	28	24	24	18	23	18	17	14	19	15	18	9	14	14	12	17	20	636			

TABLE 14. Series II. Distributions of numbers of bw^{78} genes among the cultures of each generation and grouped for generations (17-19) and (1-19).

Generation	Number of bw^{78} genes																																Total newly fixed	Total frequency	Fixation rate		
	Total fixed bw^{78}																																				
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31				32	
0	105																																				
1																																		105			
2																																			105		
3																																			105		
4																																			105		
5																																			104*		0.0096
6																																			103		0.0097
7																																			102		0.0196
8																																			100		0.0500
9																																			95		0.0211
10																																			93		0.0215
11																																			91		0.0110
12																																			90		0.0222
13																																			88		0.0455
14																																			84		0.0595
15																																			79		0.0380
16																																			76		0.0263
17																																			74		0.0676
18																																			69		0.1159
19																																			61		0.0000
1-19																																			1729		
17-19																																			204		

* Accidental loss.

asymmetry of generations 1 to 4, reflecting the period of initial increase in the frequency of bw^{75} , is largely dissipated by the 15th generation due wholly to a drop in gene frequency among unfixed classes brought about by the asymmetry in fixation, as has already been pointed out.

The theoretical distribution in the case of series I in which a constant value of $1/2N_e$ cannot be assumed may be ob-

tained from the Fokker-Plank equation (Wright, personal communication):

$$\frac{1}{2} \frac{\partial^2}{\partial q^2} [\sigma_{\%q}^2 \phi(q, t)] - \frac{\partial}{\partial q} (\Delta q \phi(q, t)) = \frac{\partial \phi(q, t)}{\partial t}$$

Considering the distribution only after the stable form and, accordingly, a con-

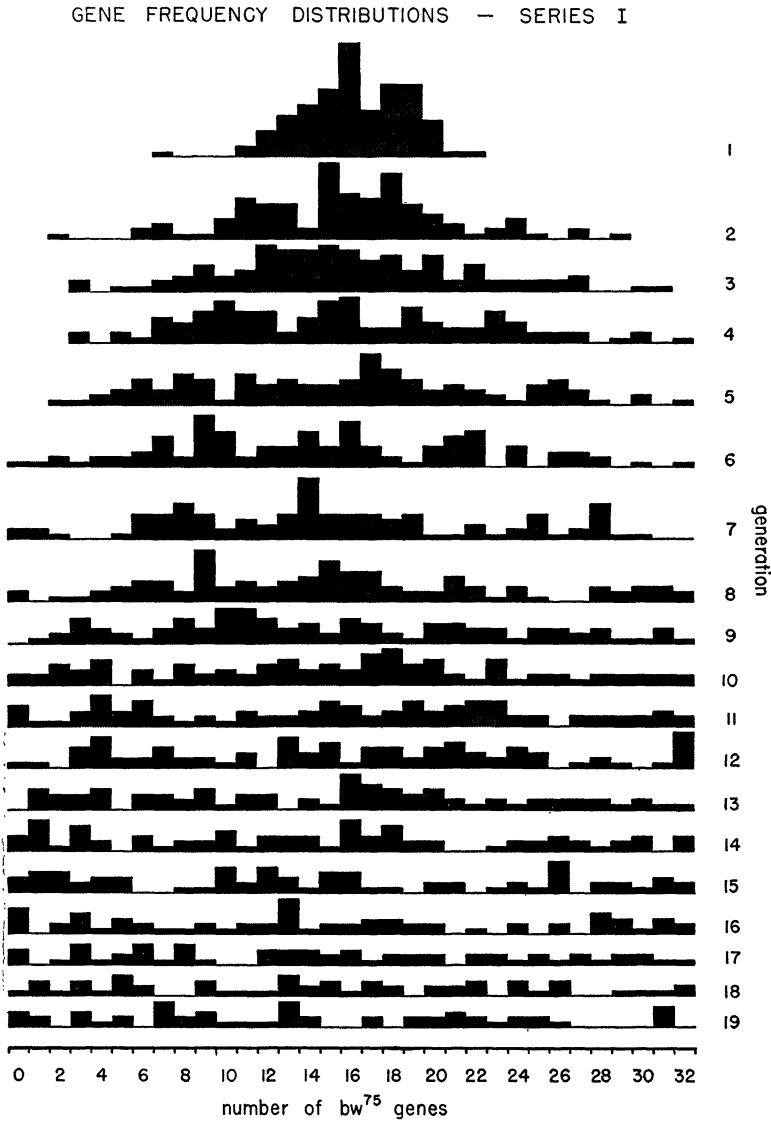


FIG. 6. Gene frequency distributions by generation in series I. Graphical presentation of the data of table 13.

GENE FREQUENCY DISTRIBUTIONS - SERIES II

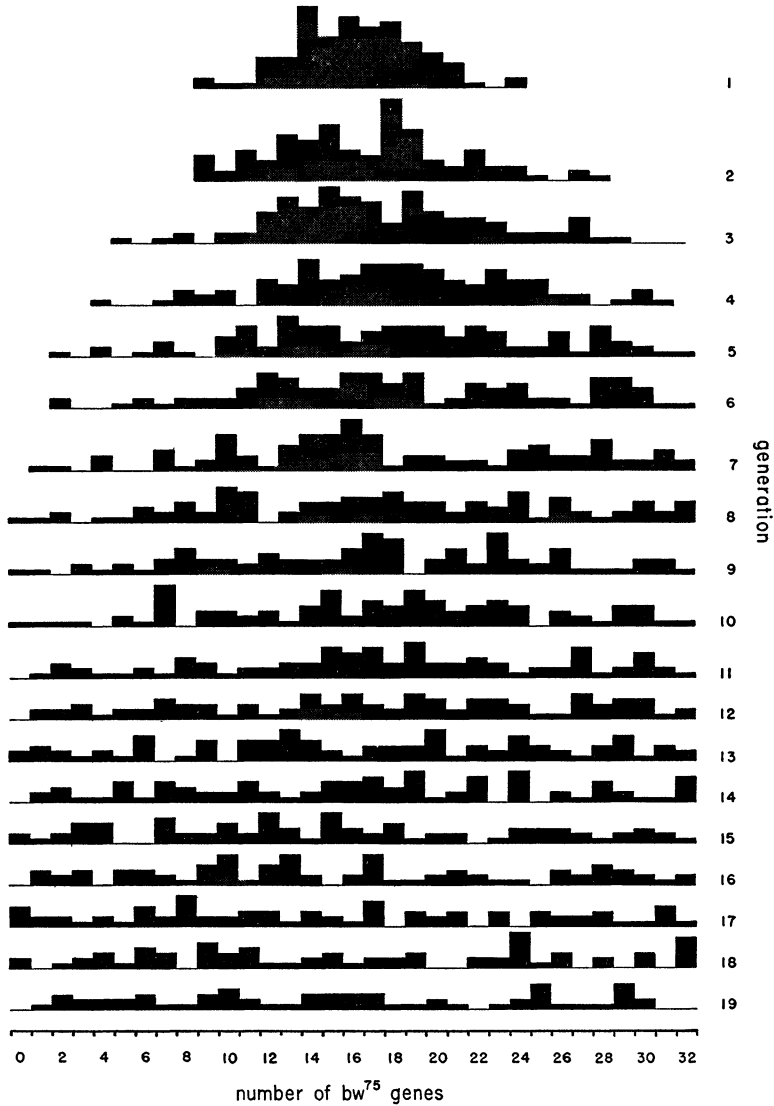


FIG. 7. Gene frequency distributions by generation in series II. Graphical presentation of the data of table 14.

stant rate of decay has been reached, this reduces to the ordinary differential equation:

$$\frac{1}{2} \frac{d^2}{dq^2} (\sigma^2_{\delta q} \phi(q)) - \frac{d}{dq} (\Delta q \phi(q)) + k \phi(q) = 0$$

(Wright, 1945). In the present case in which systematic forces may be neglected,

the second term drops out ($\Delta q = 0$) and the equation to be solved for $\phi(q)$ is:

$$\frac{1}{2} \frac{d^2}{dq^2} (\sigma^2_{\delta q} \phi(q)) + k \phi(q) = 0$$

Let

$$\phi(q) = C(1 + C_1q + C_2q^2 + C_3q^3 \dots C_nq^n)$$

and

$$\sigma^2_{\delta q} = q(1 - q)(1 + bq)/2N_0,$$

TABLE 15. *Computed frequency distribution for series I.*

32q	$\phi(q)/C$	%	32q	$\phi(q)/C$	%	32q	$\phi(q)/C$	%
0	.7214	3.01	11	.7885	3.28	22	.6406	2.67
1	.9772	4.07	12	.7729	3.22	23	.6293	2.62
2	.9552	3.98	13	.7578	3.16	24	.6183	2.58
3	.9340	3.89	14	.7432	3.10	25	.6077	2.53
4	.9136	3.81	15	.7290	3.04	26	.5975	2.49
5	.8939	3.72	16	.7152	2.98	27	.5875	2.45
6	.8749	3.64	17	.7018	2.92	28	.5777	2.41
7	.8565	3.57	18	.6889	2.87	29	.5682	2.37
8	.8387	3.49	19	.6763	2.82	30	.5590	2.33
9	.8214	3.42	20	.6640	2.77	31	.5499	2.29
10	.8047	3.35	21	.6521	2.72	32	.5877	2.45

where N_0 = the effective population number at $q = 0$. Then,

$$\frac{d^2}{dq^2} [q(1-q)(1+bq) \times (1 + C_1q + C_2q^2 + C_3q^3 \dots)] = -4N_0k(1 + C_1q + C_2q^2 + C_3q^3 \dots)$$

Arranging by powers of q and differentiating,

$$2[C_1 - (1-b)] + 6[C_2 - C_1(1-b) - b]q + 12[C_3 - C_2(1-b) - C_1b]q^2 \dots + n(n+1)[C_n - C_{n-1}(1-b) - C_{n-2}b]q^{n-1} = -4N_0k(1 + C_1q + C_2q^2 + C_3q^3 \dots C_{n-1}q^{n-1})$$

Equating the coefficients of like powers of q ,

$$C_1 = (1-b) - 2N_0k$$

$$C_2 = b + (1-b)C_1 - [2N_0kC_1/3]$$

$$C_3 = bC_1 + (1-b)C_2 - [2N_0kC_2/6]$$

$$C_n = bC_{n-2} + (1-b)C_{n-1} - [4N_0kC_{n-1}/n(n+1)]$$

The system of coefficients must be evaluated by finding by repeated trial the value of $2N_0k$ that will make the series convergent, which requires that the coefficients be kept alternating in sign. It has been found (Wright, personal communication) that, with $b = .506$, $2N_0k = 1.2375$ is correct to 4 decimals and permits the accurate computation of

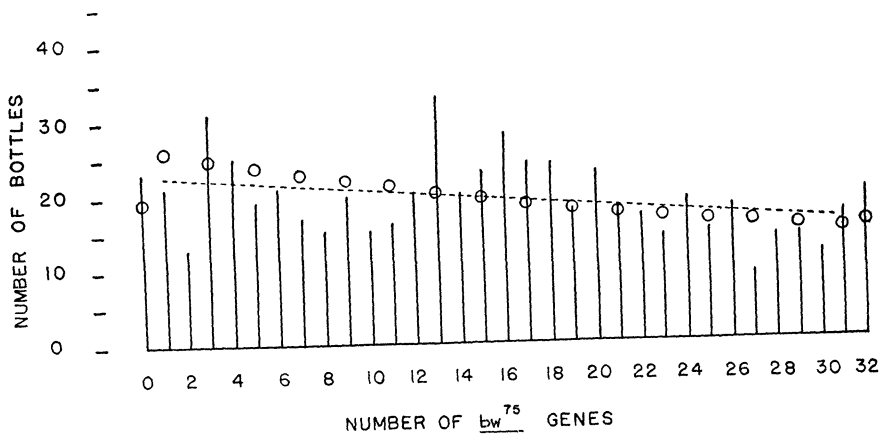


FIG. 8. Cumulative frequency distribution over the last 9 generations of series I. The dashed line represents the equation of linear trend. The open circles indicate theoretical values obtained from the frequency function, $\phi(q)$, based on the regression of $1/2N_0$ on q .

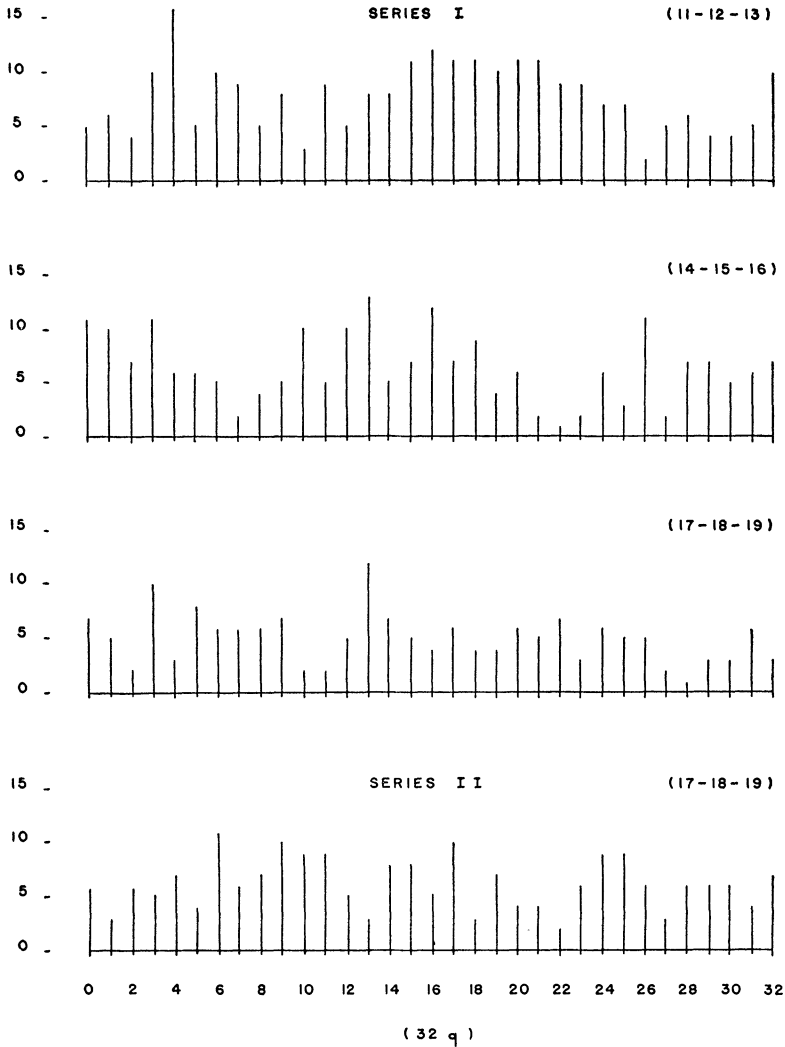


FIG. 9. Grouped frequency distribution from the data of series I over the last 3 consecutive groups of 3 generations together with the distribution over the last 3 generations of series II.

the coefficients to four places up to C_{15} . The equation for the distribution of gene frequencies may be written with sufficient accuracy:

$$\begin{aligned} \phi(q) = & C(1 - .7435q + .4454q^2 - .2481q^3 \\ & + .1335q^4 - .0706q^5 - .0369q^6 \\ & - .0191q^7 + .0099q^8 - .0051q^9 \\ & + .0026q^{10} - .0013q^{11} + .0007q^{12} \\ & - .0004q^{13} + .0002q^{14} - .0001q^{15}) \end{aligned}$$

The terminal frequencies may be found to be $\phi(0) = C$ and $\phi(1) = .5410C$ respectively.

The total fixation rate at the two extremes is equal to the rate of decay of the unfixed classes. For the present value of $2N_0k$, $k = .0558$, taking $2N_0 = 22.18$ from the calculated regression of $1/2N_e$ on q . This value is .996 of $(1/2N_e)$ and thus only slightly less than the expected value in the case of uniform class

frequencies computed on the basis of the average class value, $[\sigma^2_{\delta q}/q(1-q)]$.

This method of determining the frequency distribution is in error near the extremes. Fisher (1930) has given the correction factors in the case of uniform uncorrected frequencies. Similar corrections would be required in the present case for an exact result but the uncorrected ordinates at multiples of the theoretical class range (1/32) may be taken as a sufficient approximation.

It has been shown (Wright, 1931; Wright and Kerr, 1954) that good approximations of the rate of fixation of each allele may be taken to be half the proportional frequency in the uncorrected subterminal classes. In the present case in which $2N_e$ varies, the *effective* class ranges must be taken as $1/2N_0 = 1/22.18$ adjacent to $q=0$, and as $1/2N_1 = 1/14.73$ adjacent to $q=1$. The corresponding ordinates may be taken as $\phi(0) = C$ and $\phi(1) = .5410C$ from the formula as given. The approximate theoretical rates of fixation are then:

$$k_0 = 1/2 f_{0(e)} = 1/2[C/22.18] \\ = .7214C/32$$

$$k_1 = 1/2 f_{1(e)} = 1/2[.5410C/14.73] \\ = .5877C/32$$

The corresponding theoretical relative frequencies for the unfixed classes, on a scale with class ranges of 1/32, may be estimated as $\phi(q)/32$. Table 15 gives the relative frequencies in the form $\phi(q)/C$ and the same values on a percentage scale. It should be noted that the estimated total proportion fixed per generation (.0546) is in fair agreement with the rate (.0558) obtained in solving the general equation (which applies strictly when $2N \rightarrow \infty$).

It remains to consider the observed distributions in the light of the theoretical in both cases. Figure 8 presents the calculated distribution for series I in comparison with the observed frequencies summed over the last 9 generations.

The theoretical distribution has been adjusted to a total frequency of 636.

There is a rather close agreement between this curve and the simple linear trend line although the former is based solely on the observed recipient class variances. The calculated curve results in no improvement with respect to the distribution of the signs associated with the observed deviations. There is an obvious clumping of high class frequency in the middle range and a dearth of frequency in the lower frequency classes. A chi-square test (table 16) for goodness of fit gives a value of 42.28 with 31 degrees of freedom, $p = 0.10$. The signs of the differences bring out again the discrepancies already mentioned. Grouping the 33 classes into 6 alternating groups of 5 and 6 frequency classes, $\chi^2 = 14.15$, with 5 degrees of freedom, $p = 0.01 - 0.02$; using 3 groups of 11 classes each, $\chi^2 = 11.57$, $n = 2$, $p = 0.001 - 0.01$. The decrease in the probability as the groups are made larger is entirely due to the non-random distribution of signs.

There are several suggestions that can be made regarding the rather poor accord with expectation in this case. It has been shown theoretically (Kimura, 1955) that in cases such as are considered here it takes about $2N_e$ generations to reach substantial equilibrium with respect to form. The average value of $2N_e$ in series I was about 18 while at the upper and lower extremes the values were about 15 and 22 respectively. From this it appears that the upper half of the distribution should not be found in good accord with expectation much before the 15th generation and the lower half not before the end of the experiment.

Comparison of successively smaller terminal groupings shows that the preponderance of negative deviations in the upper half of the range of frequency classes is alleviated somewhat in the grouping of generations 15-19, and progressively improves in the grouping of generations 16-19 and 17-19. However, the distribution as a whole undergoes

TABLE 16. *Chi-square comparing the observed class frequencies over generations 1-19 in series I with values obtained from the frequency distribution, $\phi(q)$, calculated on the bases of the observed regression of $1/2N_e$ on q . Tests are presented for the ungrouped distributions, for groupings into alternating groups of 5 and 6 classes, and for 3 groups of 11 classes each.*

32q	o	c	o-c	$\frac{(o-c)^2}{c}$	o	c	o-c	$\frac{(o-c)^2}{c}$	o	c	(o-c)	$\frac{(o-c)^2}{c}$
0	23	19.1	+ 3.9	.80								
1	21	25.9	- 4.9	.93								
2	13	25.3	-12.3	5.98								
3	31	24.7	+ 6.3	1.61								
4	25	24.2	+ 0.8	.03	113	119.2	- 6.2	.32				
5	19	23.7	- 4.7	.93								
6	21	23.2	- 2.2	.21								
7	17	22.7	- 5.7	1.43								
8	15	22.2	- 7.2	2.34								
9	20	21.8	- 1.8	.15								
10	15	21.3	- 6.3	1.86	107	134.9	-27.9	5.77	220	254.1	-34.1	4.57
11	16	20.9	- 4.9	1.15								
12	20	20.5	- 0.5	.01								
13	33	20.1	+12.9	8.28								
14	20	19.7	+ 0.3	.00								
15	23	19.3	+ 3.7	.71	112	100.5	+11.5	1.32				
16	28	18.9	+ 9.1	4.38								
17	24	18.6	+ 5.4	1.57								
18	24	18.2	+ 5.8	1.85								
19	18	17.9	+ 0.1	.00								
20	23	17.6	+ 5.4	1.66								
21	18	17.3	+ 0.7	.03	135	108.5	+26.5	6.47	247	209.0	+38.0	6.91
22	17	17.0	0.0	.00								
23	14	16.7	- 2.7	.44								
24	19	16.4	+ 2.6	.41								
25	15	16.1	- 1.1	.08								
26	18	15.8	+ 2.2	.31	83	82.0	+ 1.0	.01				
27	9	15.6	- 6.6	2.79								
28	14	15.3	- 1.3	.11								
29	14	15.0	- 1.0	.07								
30	12	14.8	- 2.8	.53								
31	17	14.6	+ 2.4	.39								
32	20	15.6	+ 4.4	1.24	86	90.9	- 4.9	.26	169	172.9	- 3.9	.09
636	636.0	0.0	42.28		636	636.0	0.0	14.15	636	636.0	0.0	11.57

some distortion during this period. Figure 9 presents a breakdown of the frequency distributions of generations 11-19 into 3 successive groups of 3. It is evident that the "rarefaction" in the gene frequency range 4-13 of (11-12-13) is enhanced in the next group (14-15-16) wherein another "rarefaction" in the region of frequency 19-26 appears. The terminal grouping (17-18-19) shows a

restoration of these secondary imbalances. Such fluctuations have doubtless, in the present case, delayed the attainment of satisfactory agreement with the theoretical distribution. Further comment on this point will be made in the discussion.

Grouping the last 3 generations of series II a χ^2 for goodness of fit on the hypothesis of uniformity gives a value

of 28.29 with 32 degrees of freedom $p = 0.50$. Though the estimated value of $2N_e$ in this case is about 23, there is only a suggestion (fig. 9) of higher frequencies in the middle range and the deviations are well distributed about the mean value with respect to sign.

Total Variance and the Proportion of Heterozygosis

In a subdivided diploid population, with random breeding within the subunits but complete isolation between them, a decrease in the proportion of heterozygosis and an increase in the total variance of gene frequency among subpopulations are expected in each generation (Wright, 1931). Both can be calculated as functions of $2N_e$. It is of interest to compare the rates of change in each of these factors with theoretical rates based on the values of $2N_e$ obtained from consideration of the variances of recipient class frequencies in the last section. In the following computations values of $2N_e$ of 18 and 23 were taken as sufficiently accurate estimates for series I and II respectively. This should be adequate even in the case of series I despite the fact of regression since $2N_e$

is close to the computed value of $2N_e$ for class 16. Selection is assumed to have been absent throughout.

The expected proportion (P) of heterozygotes relative to that in a very large undivided population under random mating and at the same gene frequency may be calculated by generation using the following formula:

$$P = P' - 1/2N_e(2P' - P''),$$

where P' and P'' are the proportions in the first and second preceding generations respectively (Wright, 1931). The percentage of heterozygotes expected in any given generation (n) is given by $P_n 2q_0(1 - q_0)$, where q_0 is the initial gene frequency. $P = (1 - F)$, where F , the inbreeding coefficient, measures the correlation between uniting gametes. Since all cultures were started with heterozygous flies, the gametes uniting to form the foundation population show perfect negative correlation, $P_0 = 2$. $P_1 = 1$, because only segregation is involved. Since all populations in generation 1 are identical, there will be no overall correlation between uniting gametes if mating is random within cultures, thus P_2 will also be 1. Subsequently the value of P

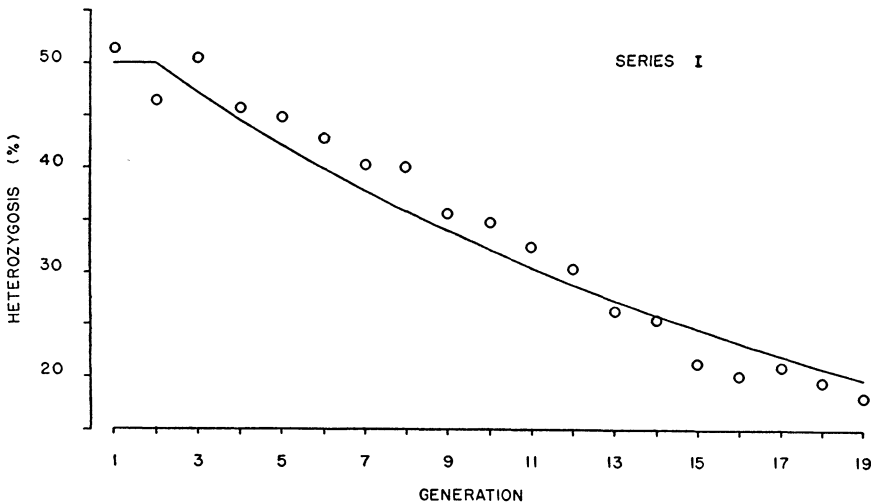


FIG. 10. Series I. Comparison of observed heterozygosis in percent (open circles) with expected values based on a common estimate of $2N_e = 18$ (smooth line).

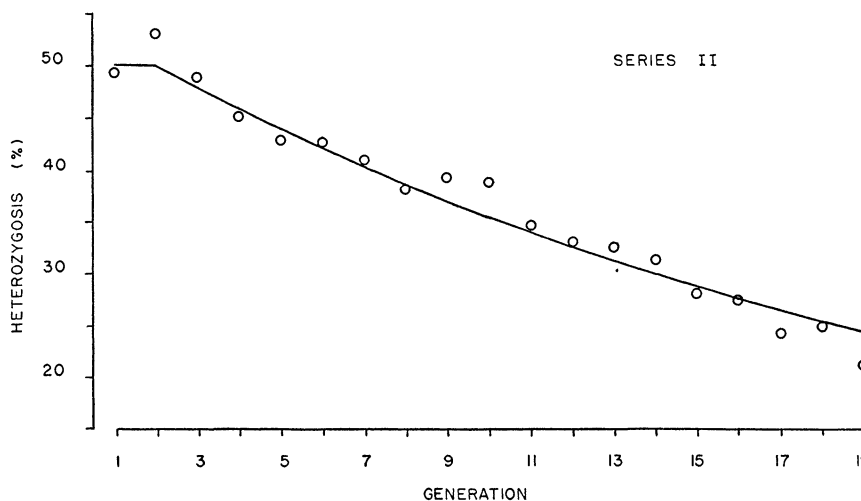


FIG. 11. Series II. Comparison of observed heterozygosis in percent (open curves) with expected values based on a common estimate of $2N_e = 23$ (smooth line).

decreases due to accidents in the sampling of $2N$ gametes.

A comparison of the observed with the calculated values of the percentage of

heterozygosis appears in figures 10 and 11 and table 17. It is evident from examination of the figure that the agreement with expectation is somewhat better

TABLE 17. *Proportion of heterozygosis and total gene frequency by generation.* Theoretical values were calculated assuming $2N_e$ values of 18 and 23 for series I and II respectively. The expected proportion of heterozygotes in % was computed as $P_n 2q_0(1 - q_0)$, where $q_0 = 0.50$ and is compared by generation with the observed values. The computed total variances are compared in each generation with the observed variance both including (A) and excluding (B) previously fixed classes.

Gen.	Heterozygosis						Total variance					
	Series I			Series II			Series I			Series II		
	P	Exp. (%)	Obs. (%)	P	Exp. (%)	Obs. (%)	σ^2_{th}	σ^2_A	σ^2_B	σ^2_{th}	σ^2_A	σ^2_B
0	2.00	100.0	100.0	2.00	100.0	100.0	.000	.000	.000	.000	.000	.000
1	1.00	50.0	51.4	1.00	50.0	49.2	.014	.006	.006	.011	.009	.009
2	1.00	50.0	46.4	1.00	50.0	52.9	.027	.026	.026	.021	.018	.018
3	.94	47.0	50.4	.96	48.0	48.5	.039	.031	.031	.031	.024	.024
4	.90	45.0	45.6	.92	46.0	45.1	.051	.042	.042	.041	.030	.030
5	.85	42.5	44.8	.88	44.0	42.7	.062	.050	.048	.050	.045	.045
6	.80	40.0	42.8	.84	42.0	42.6	.073	.055	.051	.059	.052	.050
7	.76	38.0	40.3	.81	40.5	40.8	.082	.062	.054	.067	.060	.058
8	.72	36.0	40.2	.77	38.5	38.0	.092	.072	.062	.075	.070	.065
9	.68	34.0	35.8	.74	37.0	39.1	.101	.083	.066	.083	.072	.059
10	.65	32.5	34.8	.71	35.5	38.6	.109	.090	.071	.090	.076	.059
11	.61	30.5	32.5	.68	34.0	34.6	.117	.105	.081	.097	.084	.064
12	.58	29.0	30.5	.65	32.5	33.1	.124	.112	.077	.103	.094	.074
13	.55	27.5	26.3	.63	31.5	32.5	.131	.123	.073	.110	.101	.079
14	.52	26.0	25.5	.60	30.0	31.3	.138	.136	.089	.116	.105	.075
15	.49	24.5	21.6	.57	28.5	28.0	.144	.140	.096	.122	.117	.078
16	.47	23.5	20.2	.55	27.5	27.5	.150	.155	.095	.127	.122	.079
17	.44	22.0	21.0	.53	26.5	24.3	.155	.160	.084	.133	.133	.088
18	.42	21.0	19.7	.51	25.5	24.9	.161	.165	.082	.138	.139	.089
19	.40	20.0	18.3	.48	24.0	21.0	.166	.170	.081	.143	.144	.077

in the case of series II than for series I. Since the process is stochastic the appropriate comparisons are between the apparent slopes rather than between the observed differences in each generation. There is no inherent tendency to correct for shifts due to sampling accidents. The slopes of the observed and theoretical curves are quite comparable in series II. Following an exceptionally high observed value in the second generation there is a suggestion of a greater rate of decrease than expected for about four generations. However, the rate over generations 4 to 14 is in good accord with expectation. After this time it appears to increase slightly again.

In series I the agreement is perhaps better than appears at first sight. The observed rate after the rather erratic fluctuations of the first three generations is in good accord with expectation up to generation 12. The curves, although not coincident, are roughly parallel during this period. After this the rate appears to increase somewhat, reflecting sharp decreases in observed heterozygosis between generations 12 & 13 and 14 & 15.

It should be pointed out again that the theoretical curves are based upon the assumption of a uniform value of $2N_e$ throughout the course of the experiment. Some heterogeneity in the average value of this parameter among generations in finite populations would be expected and would result in the sort of loose "fit" apparent in both cases.

The expected total variance of gene frequency among populations, including those fixed, is given for any generation by the formula:

$$\sigma_q^2 = q_0(1 - q_0)[1 - (1 - 1/2N_e)^n]$$

(Wright, 1942), where q_0 is the initial frequency and n the number of generations. The total variance of q among bottles including both fixed and unfixed cultures reaches a maximum when all bottles are fixed for one or the other of the two alternatives. The theoretical variance for a given generation excluding the previously fixed cultures rises to a maximum value equal to the variance of the stable form of the frequency distribution. The hypothesis that the stable form in these cases is approximated by

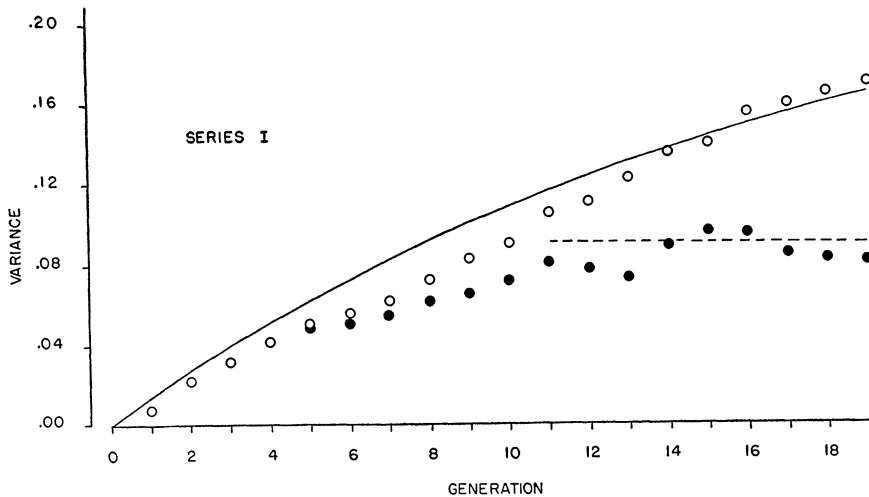


FIG. 12. Theoretical variances of the total frequency distribution by generation including fixed classes and based on a common estimate of $2N_e = 18$ for series I are represented by the smooth curves. Open circles show the observed variance of the distribution including previously fixed classes. Closed circles indicate the observed total variance excluding fixed classes. The asymptote ($= 0.091$) indicates approximately the theoretical maximum value of this variance. All values are on a relative scale.

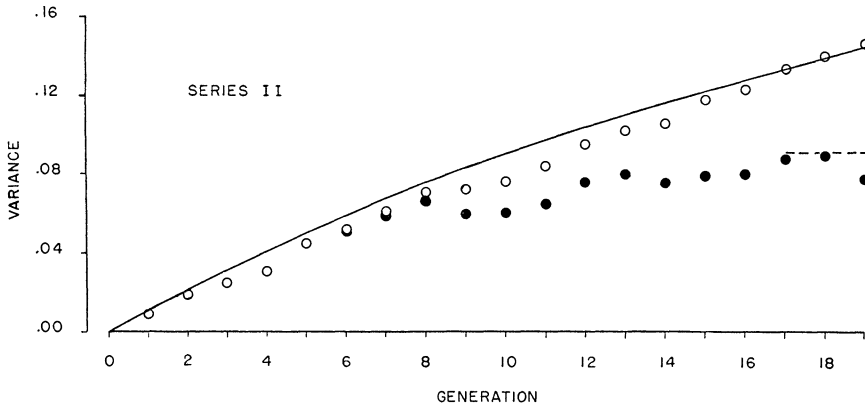


FIG. 13. Theoretical variances of the total frequency distribution by generation including fixed classes and based on a common estimate of $2N_e = 23$ for series II are represented by the smooth curve. Open circles show the observed variance of the distribution including previously fixed classes. Closed circles indicate the observed total variance excluding fixed classes. The asymptote ($= 0.091$) indicates approximately the theoretical maximum value of this variance. All values are on a relative scale.

assuming uniform frequencies for all classes has been accepted. This, again, should apply to the data of series I in the present case in spite of the non-uniformity of the theoretical distribution because of the closeness of the calculated distribution to linearity and the practical identity of the observed and theoretical mean frequencies. In computing the theoretical maximum variance excluding previously fixed classes no account has been taken of the slight falling off in frequency toward the extremes.

Table 17 and figures 12 and 13 present the observed and expected values of the variance on a relative scale by generation both excluding and including bottles previously fixed. The value of the theoretical maximum variance excluding previously fixed populations is in both cases ≈ 0.091 .

The agreement of the data with expectation is fairly good in each case. Again, there is a closer similarity between observed and expected in series II. A greater stability is, of course, conferred on the observed values of the variance including fixed cultures in both cases by the fact that the proportion of freely varying components is gradually reduced as fixation proceeds.

The asymptote is approached by the observed variance excluding previously fixed classes in the 11th and 17th generations. This is in agreement with the apparent equilibrium of form of the frequency distributions in these generations of series I and II respectively.

DISCUSSION

The Random Drift Experiments

The evidence of heterogeneity among class estimates of N_e in series I apart from the significance of the regression of $1/2N_e$ estimates on class value is clearly related to the poor accord with expectation of the form of the terminal distributions of gene frequency in this experiment after stability of form had been approached.

The fact of a tendency toward higher recipient class variances in higher donor frequency classes can itself explain both the negative slope of the grouped frequency distribution over the terminal generations in series I and the equality of the fixation rate at the two extremes despite this slope. Both effects are due to a more rapid loss of cultures from the upper frequency range, either to fixation for bw^{75} or to classes of lower frequency.

But with regard to the secondary irregularities in the apparent form of the distributions after generation 11 in this case, it is of interest to consider the effect of variability in the value of $2N_e$ among the bottles of each generation on the stability of the total distribution. In theoretical considerations upon which it may be concluded that a stable form is approached in the distribution of gene frequencies among elements in a subdivided population, a constant value is assumed for the effective population size. In the absence of systematic pressures, each frequency class (q) contributes to the entire range of $2N + 1$ frequency classes according to the expansion of the binomial $[q(A) + (1 - q)(a)]^{2N}$ and to any particular class (q_1) in proportion to the $2Nq_1$ th term of this expansion and to the frequency of donor class gene frequency (q) (Wright, 1931). Stability of form is reached when the rate of fixation is equal to $1/2N$ and the relative frequencies of the unfixed classes are constant and decaying at the rate of $1/2N$ per generation.

Variation in the effective value of $2N$ among the isolated components of a population would result in variation in the probability of a contribution being made by a given donor population at a frequency (q) to any particular recipient frequency class (q_1). Stability in the form of the distribution is dependent upon a proportional distribution of such probabilities from each donor class and upon the magnitude of the total frequency. Variability in the value of $2N$ and particularly gross heterogeneity in this factor would disturb the probability distribution among classes. Where the number of subdivisions is sufficiently large that each frequency class within the total range of unfixed classes from $1/2N$ to $(2N - 1)/2N$ is represented by a relatively large number of populations, no appreciable effect upon stability would be expected as the result of such heterogeneity. However, in cases such as are considered here, in which the mean frequency among un-

fixed classes during the last nine generations is less than three (series I), such a situation could bring about considerable distortion. These distortions would tend to compensate for one another within a series of generations and the grouped frequency distributions should more closely approach the theoretical form the larger the number of generations observed. But since a "condensation" in one part of the frequency range would not be expected to be immediately dissipated by a "rarefaction" in the same region, such compensation might easily require more than nine generations.

The somewhat better agreement of the observations with the theory in series II than in series I may be construed as rather direct evidence of greater heterogeneity in the average value of $2N_e$ among generations in the latter case. The differences between the results of series I and II are directly related to the size of the culture container in each case. The specific factor through which these differences are mediated is effective population size, but the immediate causes of reduction in the value of $2N_e$ as compared with the theoretical value in both cases, and the lower mean value, apparently greater variability, and functional relationship to q of $2N_e$ in series I, are not known. Variability in productivity extending to complete sterility is certainly a component. It seems probable that these effects are related to competition for oviposition space. In the vial cultures of series I the medium surface area was roughly three square centimeters as compared with about twelve for the series II containers. It is well known that a small environment can be quickly saturated with young larvae by a very few ovipositing females and the work of Sang (1949a, 1949b), although under quite different experimental conditions, showed that there is a strong relationship between larval mortality (as well as the duration of the larval period) and the degree of crowding of larvae. It was also found by Sang that successive batches of larvae

take longer to develop in more severely crowded cultures. It seems likely that the first females to oviposit would have a distinct advantage with respect to their contribution to the F_1 population, which would be the greater the more severe competition in the developmental phases became. However, there is no direct evidence in support of this conjecture under the conditions of these experiments.

The results of the present experiments may be compared with certain aspects of those of Kerr and Wright (1954a and b) on forked vs. type and spineless vs. arista-pedia and Wright and Kerr (1954) on Bar vs. type. Selection was not altogether negligible in the case of forked, Bar was at a great selective disadvantage, and the heterozygote ss/ss^a was strongly favored, so the stable forms in these several cases were diverse. The agreement between observed and theoretical distributions of gene frequency was excellent in the two latter cases, comparable with that of series II of the present data, although based upon a much larger total frequency in one instance. None of the sort of disturbance which occurred in series I was apparent. The distribution could not be constructed in the case of forked because individuals were not completely classifiable according to genotype.

The cultures of Kerr's studies were maintained in vials of unspecified size with breeding populations of eight individuals as compared with sixteen in the present case. It seems likely that the degree of crowding was more closely comparable with that of the series II cultures than with series I, in accord with the evident association of crowding with excessive irregularity in the "stable" form of the frequency distribution in the present data.

With regard to estimates of effective population size, these authors found values relative to theoretical of 83% (forked), 72% (Bar), and 67% (spineless) best fitted the data. The last two values are in agreement with the estimate for series II (72%). The much lower average

value obtaining in series I (56%) is, again, associated with severe crowding.

Crow (1954, 1955) has estimated the ratio N_e/N by quite different methods from those employed in population studies of the present type. He finds in experiments on *Drosophila melanogaster*, by a method involving enumeration of adult progeny, that for females $N_e/N = 0.71$ and for males 0.48. This suggests that in situations influenced equally by both sexes an intermediate value of about 0.60 might be expected, not far from the mean value for series I (0.56) of this study.

Selection

The alleles bw^{75} and bw have such a small difference in selective value as to escape detection in most aspects of these experiments. Changes in gene frequency attributable to a selective difference did occur, however, in three instances. But among these there is no agreement even as to sign and the effect in both the cage experiments and the series I random drift experiment was confined to the earliest generations. The data from the selection experiments were similarly inconclusive.

The association of genotypes having indeterminate differences in selective value with striking phenotypic differences seems to be unique among documented cases. Studies concerned with only a single character of selective value such as fecundity or only a part of the life cycle need not be considered in this context. The lowest known selective difference in the literature that has behaved with relative consistency under various conditions of continuous population culture is that between forked and type (Ludwin, 1951; Merrell, 1953; Kerr and Wright, 1954). However, even in this case considerable differences have been found under certain conditions and with a particular strain (author, unpublished). Ludwin (1951) found no evidence of significant changes in the frequency of raspberry over a period of more than six months in an array of cultures with an initial frequency of

0.50; however, Merrell (1953) reported a considerable differential, although under, again, quite different conditions of culture.

The experiments with bw^{75} vs. bw have so far included several different population densities ranging from the extreme of crowding in population cages to the conditions of the series II random drift experiments in which the number of parental flies per square centimeter of medium surface was approximately 1.25. All material was maintained at 25° C. and at high humidity. It may be that under different environmental conditions or against another "background genotype" definite differences in selective value can be found.

SUMMARY

1) Competition between the alleles bw^{75} and bw in population cages gave no evidence of a trend in gene frequency change in any of seven cages during the period of observation, although the mean frequency over all cages at the time of the first sampling was significantly below 0.50, the frequency in the foundation flies.

2) Experiments designed to ascertain the relative viabilities and productivities of the several genotypes (bw^{75}/bw^{75} , bw^{75}/bw & bw/bw), male and female, showed significant departures from the expected segregation ratios in 3 of 21 population types. These differences could not be systematically interpreted and were not reproduced in a subsidiary study.

3) Study was made of the changes in the frequency of bw^{75} between generations in cultures perpetuated for 19 generations with random samples of 16 individuals. The sample size was sufficiently small to permit considerable random drift. Two series of about 100 cultures each were set up at an initial frequency of 0.50. Series I was conducted in 35 cc. homeopathic vials, series II in 60 cc. specimen jars. The total gene frequency in series I did not change systematically during the course of the experiment. In series II the gene frequency rose for four genera-

tions after which no consistent changes occurred.

4) The fixation rate after generation 11 in series I was not significantly different from uniformity by chi-square. Series II was not carried for a number of generations sufficient to permit a satisfactory estimate of fixation rate.

5) No evidence consistent with an hypothesis of selection appeared in either series when the sample frequencies were grouped according to donor frequency classes. From this standpoint gene frequency changes could safely be attributed wholly to accidents of sampling.

6) The distribution of gene frequency by generation among the bottles of series I and II approached stability of form in generations 11 and 17 respectively. The grouped frequency distribution over the last 9 generations in series I was in poor accord with an hypothesis of uniformity, based on the apparent absence of selection, showing both an excessive irregularity and a negative slope of probable significance. The grouped distribution over the last 3 generations of series II was in good accord with the expected uniformity.

7) Estimates of the average effective population size made from the observed recipient class variances gave values of N_e equal to 9.0 and 11.5 for series I and II respectively, 56% and 72% of the sample size of 16 individuals.

8) A highly significant regression of $1/2N_e$ on donor class frequency was observed in series I indicating higher variances at higher frequencies than could be accounted for on the assumption of a common estimate of effective population size. This effect explains the slope of the frequency distributions of series I after stability of form was approximated. The theoretical distribution on the assumption of the observed regression is in agreement with respect to general slope but indicates that stability of form had not been fully reached as early as preliminary considerations suggested. The lack of accord between the observed and theoretical distri-

butions of series I is discussed from several standpoints.

9) The observed fixation rate in series I was higher than expected assuming an effective population size of 9.0 but the difference was not significant.

10) Curves representing the expected increase in the total variance of gene frequency and decrease in the percentage heterozygosis over the period of observation were calculated for each series using the mean estimates of effective population size based upon observed recipient class variances. The agreement between the observed rates and rates indicated by the calculated curves was good in the case of series II but showed some inconsistency in series I.

11) The differences between series I and II with respect to effective population size estimates, dependence of $1/2N_e$ on q , agreement with expectation of the form of the distribution with expectation of the form of the distribution of gene frequencies after stability had been approached, the homogeneity of the observed recipient class variances, rates of increase in total variance and decrease in heterozygosis are related to the size of the culture container. The data of series I are more heterogeneous in all respects.

ACKNOWLEDGMENTS

The author is indebted to Dr. Herman M. Slatis who furnished mutant material and examined the salivary gland chromosomes of the foundation stocks, and especially to Professor Sewall Wright for his generous assistance and encouragement throughout the preparation of this study. The experiments were aided by a grant from the Wallace C. and Clara A. Abbott Memorial Fund of the University of Chicago.

LITERATURE CITED

- CROW, J. F. 1954. Breeding structure of populations. II. Effective population number. Chapter 43 in *Statistics and Mathematics in Biology*, ed. Kempthorne *et al.*, Iowa State College Press.
- , and N. E. MORTON. 1955. Measurement of gene frequency drift in small populations. *EVOL.*, 9: 202-216.
- FISHER, R. A. 1925. *Statistical Methods for Research Workers*. Edinburgh, London, Oliver and Boyd.
- . 1930. *The Genetical Theory of Natural Selection*. Oxford, The Clarendon Press.
- KERR, WARWICK E., AND SEWALL WRIGHT. 1954a. Experimental studies of the distribution of gene frequencies in very small populations of *Drosophila melanogaster*. I. Forked. *EVOL.*, 8: 172-177.
- , AND —. 1954b. Experimental studies of the distribution of gene frequencies in very small populations of *Drosophila melanogaster*. III. Spineless. *EVOL.*, 8: 293-302.
- KIMURA, MOTOO. 1955. Solution of a process of random genetic drift with a continuous model. *P.N.A.S.*, 41: 144-150.
- L'HERITIER, PH., AND G. TEISSIER. 1933. Étude d'une population des Drosophiles en équilibre. *C. R. Acad. Sci. (Paris)*, 197: 1765-1767.
- , AND —. 1934. Une expérience de sélection naturelle courbe d'élimination des gene "Bar" dans une population des Drosophiles en équilibre. *C. R. Soc. Biol.*, 117: 1049.
- , AND —. 1937. Elimination des formes mutant dans les populations de Drosophiles. *C. R. Soc. Biol.*, 124: 880.
- LUDWIN, I. 1951. Natural selection in *Drosophila melanogaster* under laboratory conditions. *EVOL.*, 5: 231-242.
- MERRELL, D. J. 1953. Gene frequency changes in small laboratory populations of *Drosophila melanogaster*. *EVOL.*, 7: 95-100.
- REED, S. C., AND E. W. REED. 1948. Natural selection in laboratory populations of *Drosophila*. *EVOL.*, 2: 176-186.
- , AND —. 1950. Natural selection in laboratory populations of *Drosophila*. II. Competition between a white-eye gene and its wild type allele. *EVOL.*, 4: 34-42.
- SANG, J. H. 1949. The ecological determinants of population growth in a *Drosophila* culture. III. Larval and pupal survival. *Physiol. Zool.*, 22: 183-202.
- , 1949. The ecological determinants of population growth in a *Drosophila* culture. IV. The significance of successive batches of larvae. *Physiol. Zool.*, 22: 202-210.
- SLATIS, H. M. 1955. Position effects at the brown locus in *Drosophila melanogaster*. *Gen.*, 40: 5-23.
- WALLACE, B. 1948. Studies on "sex-ratio" in *Drosophila pseudoobscura*. I. Selection and "sex-ratio." *EVOL.*, 2: 189-217.
- WRIGHT, SEWALL. 1931. Evolution in mendelian populations. *Gen.*, 16: 97-159.

- , 1939. Statistical genetics in relation to evolution. *Actualités scientifiques et industrielles* 802. Exposés de Biométrie et de statistique biologique. Hermann et Cie, Paris.
- , 1942. Statistical genetics and evolution. *Bull. Am. Math. Soc.*, **48**: 223–246.
- , 1945. The differential equation of the distribution of gene frequencies. *P.N.A.S.*, **31**: 382–389.
- , 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*. *Gen.*, **31**: 125–156.
- , AND W. E. KERR. 1954. Experimental studies of the distribution of gene frequencies in very small populations. II. *Bar. Evol.*, **8**: 225–240.