



Experimental studies of the distribution of gene frequencies in very small populations of *Drosophila melanogaster*: I. Forked¹

W.E. Kerr and S. Wright

Universidade de São Paulo and The University of Chicago
Received February 23, 1954²

INTRODUCTION

For quantitative study of the sort of random drift due to inbreeding it is desirable to experiment with segregating genotypes that can be classified without risk of error and with populations that are all of a very small definitely known size.

With respect to the first condition, it is much the most satisfactory if *all* segregating genotypes can be distinguished accurately. This, however, limits severely the number of loci that can be studied. Something can be done where only dominants and recessives can be accurately classified.

Most pairs of alleles with such conspicuous difference in effect as to satisfy the first condition, turn out to be subject to such enormous differences in selection that the accumulation of random deviations, implied by the term random drift, is largely prevented. Selection pressure and the effects of random processes are roughly comparable in magnitude at a given gene frequency if the change in frequency in a generation which the former tends to bring about (Δq) and the variance increment due in a generation to the latter ($\sigma_{\delta q}^2$) are of the same order (Wright, 1931, 1948). Where the ratio ($\Delta q/\sigma_{\delta q}^2$) is as high as 10 in absolute value, there is not much accumulation of random deviations and hence little random drift. The size of experimental populations in studies of the latter, should be small enough to meet this condition. They should be constant for convenient comparison with theory.

The present experiments (all with *Drosophila melanogaster*) were designed to meet these conditions. Three series were performed. About 120 lines were started in each. In the first series, the sex-linked mutation forked (f) competed with its type allele. Four females (1 f/f, 2 f/+, 1 +/+) and 4 males (2 f/0, 2 +/0) were put in each vial. The second series involved the sex-linked semidominant mutation Bar (B) and its type allele. Each initial vial contained 4 B/+ females and 4 males (2 B/0,

¹Experimental data by W. E. Kerr, under a fellowship of the Rockefeller Foundation, mathematical analysis by S. Wright. Acknowledgment is made to Dr. Th. Dobzhansky for material and for his hospitality during the conduct of these experiments. Acknowledgment is also made to Dr. J. Crow for material and suggestions. Analysis was aided by a grant from the Wallace C. and Clara A. Abbott Memorial Fund of The University of Chicago.

²Editor's Note: This paper is the first of a series of three. The others will be published in successive issues of *Evolution* [*Evolution* 8: 172-177, June 1954]. Reprinted by permission of The Society for the Study of Evolution and the authors from *Evolution* 8: 172-177, 1954.

2 +/0). The third series was with the autosomal alleles aris-tapedia, ss^a , and spineless ss ; which produce a heterozygote that is close to type. Each initial vial contained 4 ss^a/ss females and 4 ss^a/ss males. The alleles were thus equally frequent at the beginning of each experiment.

The cultures were allowed to develop until about 2 to 4 days after the offspring began to emerge. The flies hatched up to this time were discarded (in an evening). The flies which appeared next morning (if enough had emerged) were etherized and from among them 4 females and 4 males were taken at random and served as progenitors of the following generation. The etherized flies were put on a porcelain plate and the first 4 males and the first 4 females that happened to be closest to the right end of the plate were the flies taken. It was often, however, necessary to wait to the second and sometimes to the third day to obtain 4 of each sex. This procedure was repeated in every following generation in every line. The first series was carried 16 generations, the second 10 and the third 9 generations. In general, all 8 parents in each culture lived until their progeny started to emerge. In some instances one or more died a few days after the culture was started. No new flies were substituted. Lines were discontinued if fixation was attained. A few were discarded on account of mite infection and other accidents. These latter strains have not been included in the statistics presented here. The present paper will be confined to the experiments with the forked.

FORKED

Table 1 shows the most important results. Among the 96 lines with initial gene frequency (.50 f + .50 f⁺), 26 were still unfixed at the end of 16 generations. Wild type had become fixed in 41 and forked in 29. There is here only an in-significant suggestion that forked was at a selective disadvantage. There seems, however, to have been a shift toward increasing disadvantage of forked as the experiment proceeded. In the first 8 generations, wild type became fixed in 17 lines, forked in 23, while in the last 8 generations wild type became fixed in 24 lines, forked in only 6, a difference with a probability from accidents of sampling of less than .01. Nevertheless it is clear that the selection against forked must have been slight to have permitted as much fixation as occurred against its pressure.

Table 1. The amount of fixation of forked and its type allele in 96 lines, each consisting of 4 ♂'s and 4 ♀'s in each generation and carried 16 generations unless fixed earlier. The amount of fixation from generation 4 to 16 is compared with the expected amount at a constant rate of 8.9% per generation.

	Type newly fixed	Not fixed	Forked newly fixed	Total	Observed (fixed)		Calc. No. (c)	(o-c)	(o-c) ² c
					No. (o)	%			
1	1	94	1	96	2	2.1			
2	0	92	2	94	2	2.1			
3	1	87	4	92	5	5.4			
4	5	79	3	87	8	9.2	7.7	+0.3	.01
5	3	70	6	79	9	11.4	7.0	+2.0	.57
6	1	66	3	70	4	5.7	6.2	-2.2	.78
7	5	59	2	66	7	10.6	5.9	+1.1	.21
8	1	56	2	59	3	5.1	5.3	-2.3	1.00
9	3	52	1	56	4	7.1	5.0	-1.0	.20
10	4	47	1	52	5	9.6	4.6	+0.4	.03
11	5	39	3	47	8	17.0	4.2	+3.8	3.44
12	2	37	0	39	2	5.1	3.5	-1.5	.64
13	3	34	0	37	3	8.1	3.3	-0.3	.03
14	3	30	1	34	4	11.8	3.0	+1.0	.33
15	1	29	0	30	1	3.3	2.7	-1.7	1.07
16	3	26	0	29	3	10.3	2.6	+0.4	.06
1-3	2	273	7	282	9	3.2			
4-8	15	330	16	361	31	8.6	32.1	-1.1	
9-16	24	294	6	324	30	9.3	28.9	+1.1	
4-16	39	624	22	685	61	8.9	61.0		8.37

This result is in agreement with those of previous experiments. In Ludwin's (1951) experiments, the initial gene frequency of .50 in cultures which contained on the average about 44 males, 51 females but with enormous variations, fell to .30 in about 2 or 3 months and was still about the same at 6 months. In experiments by Merrill (1953) in populations that rarely exceeded 100 adults and in many cases were down to less than 10 flies, the frequency of forked similarly fell from .50 at the beginning to .30 by 99 days and averaged .33 in counts from 125 days to 270 days. There was no indication of increasing selection against forked in those experiments. Both authors found evidence of important differences in gene frequency among individual cultures which they attributed to random drift.

From inspection of Table 1, it appears that the percentage of fixation (including both that of type and forked) rose to generation 4 but did not change consistently thereafter. The average rate for generations 4 to 16 was 8.91% (standard error 1.09%). Assuming theoretical constancy at this figure, the difference between observed and calculated numbers for unfixed and for newly fixed lines yields $\chi^2 = 8.4$, 12 degrees of freedom, probability .70 - .80 of being exceeded by accidents of sampling.

The theoretical rate of fixation for a neutral sex-linked gene after a steady rate has been attained has been given as approximately $(2 N_M + N_F) / (9 N_M N_F)$ in which N_M is the effective number of males and N_F that of females (Wright, 1933). This approximation only applies, however, if N_M and N_F are moderately large. The panmictic index $P (= 1 - F$ where F is the inbreeding coefficient) measures the amount of heterozygosity relative to that in a random bred stock. The exact recurrence formula derived by the method of path coefficients was given as follows in the paper cited. Primes refer to preceding generations.

$$P = P' - C_1 (2 P' - P'') + C_2 (2 P'' - P''')$$

where

$$C_1 = (N_F + 1) / (8 N_F),$$

$$C_2 = (N_M - 1) (N_F - 1) / (8 N_M N_F)$$

In the present case

$$C_1 = 5/32, C_2 = 9/128.$$

$$P = .6875 P' + .296875 P'' - .0703125 P'''$$

The correlation $(1 - P_0)$ between the gametes that united to produce the foundation females was zero and the correlation $(1 - P_1)$ between the gametes that unite to produce their daughters is also zero. If there were no differential fecundity among the foundation flies there would be no correlation between mating males and females and hence $(1 - P_3) = 0$. From this point, however, the inbreeding coefficient rises. For calculation of P_3 , $P' (= P_2)$, $P'' (= P_1)$ and $P''' (= P_0)$ are all assigned the value 1. Table 2 shows the values of P , the ratio of successive values P/P' and the percentage change in P per generation $(100 (P - P') / (P'))$ for 20 generations. It may be seen that P/P' oscillates about the value .9260855. This is approximated to 4 places by the 11th generation and to 7 places by the 19th. Values of P beyond this point can thus be calculated to 6 places by the formula $P = .9260855 P'$.

Table 2. The theoretical values of P (= 1 - F), P/P' and the proportional rate of change of P in a population consisting of 4 males and 4 females under sex linked inheritance and no selection.

Generation	P (= 1 - F)	P/P'	$\Delta P/P'$
0	1		
1	1	1	0
2	1	1	0
3	.914062500	.914062500	.085937
4	.854980469	.935363248	.064637
5	.788848877	.922651341	.077349
6	.731885910	.927789759	.072210
7	.677245259	.925342665	.074657
8	.627418808	.925427759	.073572
9	.580946889	.925931581	.074068
10	.538047138	.926155468	.073845
11	.498260630	.926053861	.073946
12	.461439099	.926099858	.073900
13	.427329066	.926079014	.073921
14	.395744515	.926088456	.073912
15	.366492734	.926084178	.073916
16	.339403833	.926086117	.073914
17	.314316879	.926085236	.073915
18	.291084347	.926085636	.073914
19	.269568980	.926085455	.073915
20	.249643934	.926085538	.0739145

The rate of decrease of heterozygosis after it reaches stability can also be derived at once by equating P/P' , P'/P'' and P''/P''' and expressing in terms of $\chi = \Delta P/P'$ (Wright, 1933).

$$\chi^3 + \chi^2 (2 + 2 C_1) + \chi (1 + 3 C_1 - 2 C_2) + (C_1 - C_2) = 0$$

The solution is .0739145, the limit about which $\Delta P/P'$ oscillates in the successive generations.

The rate of fixation of lines approaches this same value although somewhat more slowly. Actually it has been noted that a practically constant rate 8.91% is attained by F_4 . As this is 20.5% larger than the theoretical rate, it is implied that the actual variance due to random processes in each generation was 20.5% greater than expected. As the difference is only 1.5 times its standard error, it is not certainly significant. Taking it at face value, the excess might conceivably be due to fluctuating selection but as there is very little average selection it more probably means that the effective size of population is about 83% of the theoretical value. In an autosomal diploid population the rate of decrease of heterozygosis is $\Delta P/P' = 1 / (2N)$ (Wright, 1931). If effective size is defined as $P'/2 \Delta P$, the theoretical effective size with sex linkage and 4 females and 4 males per generation is 6.7646 while the effective size of the experimental population was 5.61. The difference if real can easily be accounted for if 1 or 2 of the 8 flies fail completely to reproduce in each generation.

The distribution of gene frequencies, during the period of constant rate of fixation, must have practically reached equilibrium of form. The actual distribution was not deter-

mined because of the lack of visible distinction between $+/+$ and $+/f$. It is of some interest, however, to consider what it must have been. It may suffice to give the distribution for an effective population of 6 and ignore the indicated slight selective differential (Wright, 1931). The standard is considered a population of 6 monoecious diploid individuals with completely random union of gametes, as it has been shown that the distribution is nearly the same for given N_e , irrespective of the system of mating (Wright, 1931).

Let $f(q)$ be the frequency of gene frequency q . The class with this gene frequency contributes to the classes of the next generation according to the expansion of $[(1 - q) (f^+ + q (f))^{12} f(q)$ and thus to the class with gene frequency q_1 by

$$\frac{12!}{(2Nq)! [2N(1 - q)]!} q^{2Nq_1} (1 - q)^{2N(1 - q_1)} f(q).$$

The total frequency of q is the sum of all contributions from the 11 values of q from 1/12 to 11/12. This must be 11/12 of its value in the preceding generation if there is equilibrium of the form of the distribution. The equation can be solved algebraically as in a number of simple examples given in the reference cited³ but it is probably simplest in this case to start from a rough approximation and iterate, rating up each generation by 12/11, until stability is reached. The resulting distribution is shown in Table 3 and Figure 1.

Table 3. The distribution of unfixed classes after stability of form has been reached in a population in which $2N = 12$, no selection; and the distribution in the following generation, including newly fixed classes (each 1/24 of the total unfixed classes of the preceding generation). The frequency in each of the unfixed classes is 11/12 of its value in the preceding generation, thus maintaining stability of form.

	Unfixed classes	Following generation
0		.04167
1	.07881	.07224
2	.08985	.08236
3	.09317	.08541
4	.09475	.08685
5	.09554	.08758
6	.09576	.08778
7	.09554	.08758
8	.09475	.08685
9	.09317	.08541
10	.08985	.08236
11	.07881	.07224
12		.04167
	1.00000	1.00000

³We will note that one of these is incorrect as published, viz., that for irreversible mutation in populations of 3 monoecious individuals (p. 118). The correct percentages are 42.40% for $q_a = 1/6$, 21.07% for $q_a = 2/6$, 15.54% for $q_a = 3/6$, 11.99% for $q_a = 4/6$, 9.00% for $q_a = 5/6$, total 100% for unfixed classes. The rate of fixation of A is 17.05% and of a, 4.53% of the unfixed classes*.

*The last sentence should read: "The rate of fixation of A is 16.31% and of a, 4.34% of the unfixed classes."

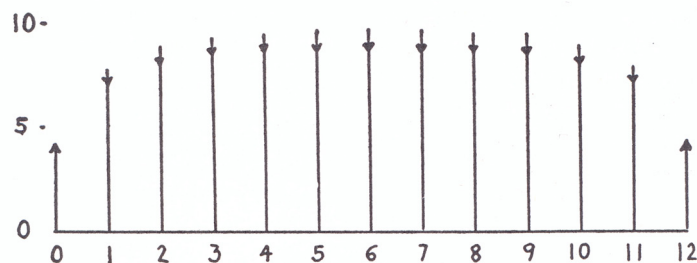


Figure 1. The theoretical distribution of gene frequencies (including newly fixed classes) after attainment of stability of form in a population in which $2N = 12$ and selection is absent. Rough estimates for the unfixed classes are given by the ordinates of the curve, $y = 1/12$, unit area (base 12) derived from indefinitely large $2N$. The frequencies ($1/24$) of the newly fixed classes are given exactly by half the terminal ordinates of this curve.

While this solution applies to a situation with a slightly lower rate of fixation ($1/24$ for each allele) than that observed, the form of the distribution apart from the newly fixed classes is substantially correct for 4 ♀'s, 4 ♂'s (12 representative of the locus) for any effective size of parental population not too remote from 6. Indeed with an indefinitely large population in each generation, the form of the distribution is not very different, being approximately uniform for all gene frequencies (Wright, 1931).

A situation that arose in one line (No. 109) in this series is interesting. In the 5th generation the flies taken to be parents of the 6th generation were found to be segregating for an eye color found to behave as an allele of white. This gene drifted in frequency until it became fixed in the 12th generation. The frequencies of f and this gene (apparently w^a) were as shown in Table 4.

Table 4. The history of a line in which an eye color mutation appeared in generation 5 and drifted into fixation in generation 12.

	F ₄		F ₅		F ₆		F ₇		F ₈		F ₉		F ₁₀		F ₁₁		F ₁₂	
	f	w ^a	f	w ^a	f	w ^a	f	w ^a	f	w ^a	f	w ^a	f	w ^a	f	w ^a	f	w ^a
♀ recessive	2	-	1	1	0	0	1	3	0	0	0	2	0	4	0	3	0	4
♀ dominant	2	-	3	3	4	4	3	1	4	4	4	2	4	0	4	1	4	0
♂ recessive	1	-	1	3	0	4	1	0	0	4	0	2	0	3	0	4	0	4
♂ dominant	3	-	3	1	4	0	3	4	4	0	4	2	4	1	4	0	4	0

Thus this line had probably become fixed for the type allele of forked by the 9th generation and after the 12th generation for the new eye color mutant.

DISCUSSION

Populations of 4 males and 4 females per generation are so exceedingly small that experiments such as the present may seem to have no implications for evolution in nature. It must be borne in mind, however, that changes in the underlying multifactorial genetic struc-

ture of species probably occur so slowly that an appreciable change in a thousand generations must be considered as an explosively rapid process.

Study in the laboratory of the factors that can contribute to such change is practicable only by stepping up the rates by at least one hundred fold. Thus the interaction between a weak selective advantage of one isoallele over another and a slight random drift, due to inbreeding, can be simulated by using alleles with selective differentials of ten percent or more instead of perhaps only one tenth of a percent or even one hundredth of a percent in populations of only one percent or even one tenth of one percent of the size of a typical natural deme.

In the case of forked, the selective differential is clearly much less than ten percent so that the results of the present paper illustrate random drift from inbreeding in an almost pure form. More complicated situations will be considered in the later papers of this series.

SUMMARY

Ninety-six lines of flies (*Drosophila melanogaster*) were started, each from 4 females (1 f/f, 2 f/+, 1 +/+) and 4 males (2 f/0, 2 +/0), and continued to fixation or to the sixteenth generation by random selection of 4 females and 4 males as parents of each new generation. Type (f⁺) became fixed in 41 lines, forked (f) in 29 lines, and 26 lines were still unfixed at the end. The amount of selection against forked was thus slight, although there was evidence that it was greater in the later generations than at first.

The rate of fixation (of both alleles combined) reached approximate constancy by the fourth generation at 8.9% per generation. This would imply an effective size of population 83 per cent of that expected under sex linkage with 4 females and 4 males per generation but the reduction is of doubtful significance.

In one line, an eye color mutation, probably apricot, appeared in F₇ and became fixed in F₁₂, three generations after fixation of the type allele of forked.

REFERENCES

- Ludwin I (1951). Natural selection in *Drosophila melanogaster* under laboratory conditions. *Evolution* 5: 231-242.
Merrell DJ (1953). Gene frequency changes in small laboratory populations of *Drosophila melanogaster*. *Evolution* 7: 95-101.
Wright S (1931). Evolution in Mendelian populations. *Genetics* 16: 97-159.
Wright S (1933). Inbreeding and homozygosis. *Proc. Natl. Acad. Sci. U. S. A.* 19: 411-420.
Wright S (1948). On the roles of directed and random changes in gene frequency in the genetics of populations. *Evolution* 2: 279-294.