GENETIC DRIFT AND NATURAL SELECTION IN EXPERIMENTAL POPULATIONS OF DROSOPHILA PSEUDOOBSCURA

THEODOSIUS DOBZHANSKY AND N. P. SPASSKY*

DEPARTMENT OF ZOOLOGY, COLUMBIA UNIVERSITY

Communicated December 15, 1961

An interesting indeterminacy of outcomes is observed in experiments with certain laboratory populations of Drosophila. Natural populations of *Drosophila pseudoobscura* and of many other species are polymorphic with respect to the gene arrangements in their chromosomes. The polymorphism is balanced; the heterokaryotypes with two chromosomes of a pair differing in gene arrangement are, in most environments, superior in Darwinian fitness to the corresponding homokaryotypes. The relative frequencies of the different karyotypes in experimental populations change from generation to generation until quasistable equilibria are attained. The results of the experiments depend, however, on whether the populations are uniform or mixed in geographic origins, i.e., on the chromosomes having been derived from the same natural population or from populations of different localities. In the former populations, the results are generally determinate; if the environment is well controlled, the selection rates and the equilibrium points are constant and reproducible. Not so in populations of geographically mixed origins; replicate experiments often give disparate results, and the selection rates and the equilibrium points are unpredictable for a given population.[1–4]

A Working Hypothesis.—The indeterminacy of the outcomes of the selection processes in populations of geographically mixed origins may be due to our experimental populations containing far too few flies to have more than a fraction of the possible genotypes actually realized. Supposing that the geographic populations, or races, crossed differ in 50 to 100 genes, between 3^{50} and 3^{100} genotypes will be potentially possible in the progenies; on the other hand, the populations fluctuate in size roughly between 1,000 and 4,000 adult individuals. Now, the relative fitness of the homo- and heterokaryotypes depends upon the genetic system as a whole. Which of the possible recombination genotypes arise first in a given population, and which appear later or not at all, is a matter of chance. Natural selection in the experimental populations will, however, work with the genotypes which will happen to be available. The dissimilar results found in experimental populations of geographically mixed origins arise, then, owing to the divergent action of natural selection; the divergence is due, in turn, to the operation of the sampling processes which bring some of the potentially possible genotypes to realization and leave others unrealized.

This working hypothesis permits some predictions to be made. In populations large enough to have all the potentially possible genotypes actually appear, the results should always be determinate. Experimental populations of 3^{50} or 3^{100} individuals obviously cannot be obtained. There must, however, be some relation between the population size and the degree of the indeterminacy, the variance of the results presumably increasing as the population size decreases. To test this prediction, Dobzhansky and Pavlovsky[5] have crossed 12 strains of *D. pseudoobscura*

from Texas having the PP gene arrangement in their third chromosomes with 10 strains of the same species from California having the AR gene arrangement. The F_2 generation hybrids served as the foundation stock for 20 experimental populations. Of these, 10 populations were started with 4,000 flies in each, and the other 10 populations with 20 flies per population cage. Because of the high fertility of the flies, these initially "large" and "small" populations became about equal in size a generation later, and continued to be equal from then on. All the populations were treated as much alike as possible. Nevertheless, 17 months later the frequencies of PP and AR chromosomes were found to vary, the initially large populations containing from 20 to 35 per cent PP, and the initially small ones from 16 to 47 per cent. The variance in the small populations was 4.4 times greater than in the large ones—a statistically significant difference.

Another experimentally verifiable deduction from the working hypothesis is as follows. If the sizes of the foundation stocks of the populations are kept constant, the variances of the outcomes should be greater if these foundation stocks come from a source with a high genetic variability than if they come from a genetically more uniform source. In a sense, this is verified by the observation that the selectional events are determinate in experimental populations of geographically uniform origins. Such populations must have less genetic variance than the geographically mixed populations.[6] There is, however, another method to test the validity of the same deduction. As stated above, in the experiments of Dobzhansky and Pavlovsky, 22 different strains entered in the foundation stocks of the experimental populations. We now wish to report experiments in which such populations are compared with populations derived from only two strains, hybrids of single homozygous lines from two geographic regions.

The Experiments.—In March, 1957, we started 10 experimental populations, five of each of the two following kinds. The population of the first kind was derived from 20 foundation strains; 10 strains of *D. pseudoobscura* with PP third chromosomes from Austin, Texas, were crossed to 10 strains with AR chromosomes from Mather, California (these were the same strains as used by Dobzhansky and Pavlovsky[5]). An F_2 progeny was raised, and from it five groups of 20 flies each (10 ♀ ♀ + 10 ♂ ♂) were taken in a way which insured that all 20 ancestral strains would be represented in each group. The flies were allowed to reproduce in ordinary culture bottles, and in the next generations, 500–1,000 flies were taken from the cultures of each group and placed in population cages Nos. 1–5. The populations of the second kind were derived from a single pair of flies, one with PP chromosomes from Texas and the other AR from California. An F_2 generation was raised, and five groups of flies, 10 ♀ ♀ + 10 ♂ ♂ in each, were taken at random, allowed to reproduce in ordinary cultures, and five sets of 500–1,000 flies were placed in the population cages Nos. 6–10. We shall refer to the populations of the first kind as "multichromosomal" and of the second kind as "bichromosomal."

All the populations were allowed to breed freely for about four months. At the end of this first "cycle," the composition of each population was determined by examination of the chromosomes in the salivary glands of 150 larvae (300 chromosomes). A sample of 20 flies (10 ♀ ♀ + 10 ♂ ♂) was taken from each population and made to serve as the foundation stock of a new, second "cycle," population. After another four-month breeding period, the process was repeated, and popula-

tions of the third cycle started. The multichromosomal populations were carried through nine such cycles. For reasons to be explained below, the original bichromosomal populations, Nos. 6–10, were carried through four cycles, whereupon another series of bichromosomal populations, Nos. 11–15, were started and carried through five cycles.

Table 1 shows the zygotic and gametic frequencies of the different karyotypes in all the experimental populations. Let us consider the multichromosomal populations, Nos. 1–5, first. The average frequencies of AR and PP chromosomes in the foundation stocks of these populations must have been 50 per cent; the frequencies of AR in all the samples without exception were found to be above, and those of PP below, 50 per cent. This is as expected; in the experiments of Dobzhansky and Pavlovsky[5] the fitness of homozygous AR/AR was found to be above that of PP/PP; therefore, natural selection drives the populations towards equilibria at which AR is more frequent than PP. However, the frequencies of AR and PP vary from population to population, and the variance increases with time in the consecutive cycles. In fact, the heterogeneity chi-square for the first cycle is 6.01, which for four degrees of freedom has a probability between 0.2 and 0.1; from the second cycle on, the chi-squares correspond to probability values below 0.001. The increase of the variance is represented graphically in Figure 1. The straight line is fitted to the observed points by least squares, and it appears that during the early cycles the variance grows linearly with time. The ninth cycle

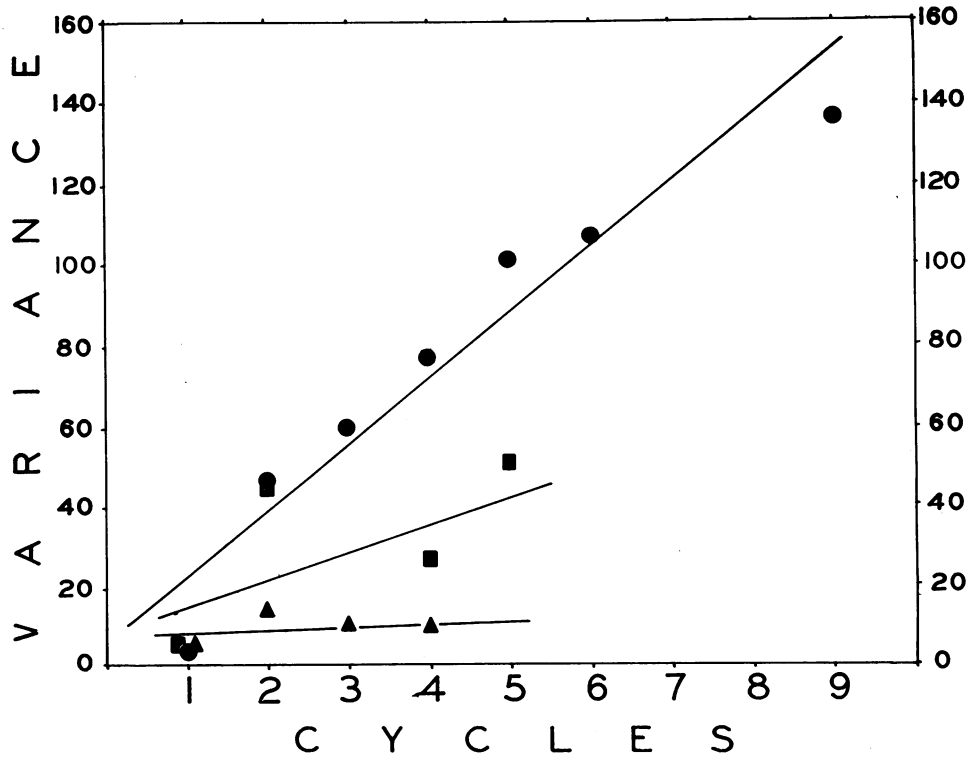

FIG. 1.—The increasing variance of the chromosome frequencies in multichromosomal populations (circles) and in bichromosomal populations Nos. 6–10 (triangles) and Nos. 11–15 (squares).

gives, however, a variance lower than expected, which probably means that the increase slows down and perhaps stops after several cycles. This is unfortunately not certain, since the chromosome frequencies were not scored during the seventh and eighth cycles.

The bichromosomal populations Nos. 6–10 behaved differently from the multichromosomal ones. As shown in Table 1, the frequencies of AR chromosomes not only did not increase but dropped, and those of PP chromosomes increased slightly from the initial value of 50 per cent. The mean frequencies of AR for the four cycles were 49.6, 46.0, 36.6, and 37.7 per cent respectively. Furthermore, the replicate populations diverged relatively little. The heterogeneity chi-squares were 5.74, 11.33, 9.60, and 9.65 for the 1st to 4th cycles (a chi-square 9.49 corresponds, for four degrees of freedom, to a probability of 0.05). Figure 1 shows that the variance failed to increase anywhere near as sharply as it did in multichromosomal populations; in fact, the straight line fitted to the four observed points is not very different from horizontal.

There is one further difference between the multichromosomal and the bichromosomal populations which is worth noting. Table 1 shows the observed numbers of heterozygous AR/PP and of homozygous AR/AR and PP/PP individuals. We have tested, for each sample, the fit of these observed numbers to the expectation on the basis of the Hardy-Weinberg equilibrium condition. Among the 35 samples in the multichromosomal populations, 21 contained more and 14 fewer heterokaryotypes than expected. Among the 20 samples in the bichromosomal populations Nos. 6–10, 18 contained more and only two fewer than the expected number of the heterokaryotypes. We are obliged to Howard Levene for his advice to test the statistical significance of the observed excesses in the incidence of the heterokaryotypes by means of the chi statistics. The normal deviates, testing the significance of the observed frequencies of the heterokaryotypes AR/PP in excess of those expected if the Hardy-Weinberg equilibrium were realized, are shown in Table 2. Since we are interested only in excessive, not in deficient, frequencies of the heterokaryotypes, the normal deviates greater than 1.645 have probabilities less than 0.05 of chance occurrence. It can be seen that bichromosomal populations contain too many AR/PP, and too few AR/AR and PP/PP, karyotypes, while in the multichromosomal populations this is not observed.

It is evident that in the biochromosomal populations Nos. 6–10 the homokaryotypes were grossly deficient in viability, while in the multichromosomal populations they were much less deficient. This is in no wise surprising, since many individual chromosomes found in natural populations are deficient in vigor or even lethal in double dose. The homokaryotypes in bichromosomal populations are genic as well as chromosomal homozygotes, while in the multichromosomal ones they are genic heterozygotes, and this permits some hybrid vigor to manifest itself. The behavior of bichromosomal populations under selection depends, therefore, upon which particular natural chromosomes happen to be included. To make sure that the relative homogeneity of the results in the bichromosomal populations Nos. 6–10 was not accidental, we started five new bichromosomal populations, Nos. 11–15, and carried them through five cycles. The data in Table 1 show that in these populations AR chromosomes came, as in the multichromosomal ones, to outnumber PP chromosomes. From the second cycle on, the five replicate populations were

TABLE 1

FREQUENCIES OF KARYOTYPES AND GAMETIC FREQUENCIES OF THE CHROMOSOMES WITH DIFFERENT GENE ARRANGEMENTS

AR, Arrowhead chromosomes
PP, Pike's Peak chromosomes

Population	AR/AR	AR/PP	PP/PP	AR (%)	PP (%)
		1st Cycle, October, 1957			
1	83	61	6	75.7	24.3
2	93	49	8	78.3	21.7
3	78	63	9	73.0	27.0
4	71	69	10	70.3	29.7
5	89	47	14	75.0	25.0
		2nd Cycle, March, 1958			
1	78	60	12	72.0	28.0
2	102	41	2	85.0	15.0
3	103	43	4	83.0	17.0
4	76	62	12	71.3	28.7
5	70	68	12	69.3	30.7
		3rd Cycle, August, 1958			
1	74	61	5	69.7	30.3
2	103	44	3	83.3	16.7
3	114	32	4	86.7	13.3
4	93	48	9	78.0	22.0
5	72	61	17	68.3	31.7
		4th Cycle, January, 1959			
1	119	27	4	88.3	11.7
2	130	20	..	93.3	6.7
3	78	60	12	72.0	28.0
4	97	49	4	81.0	19.0
5	84	54	12	74.0	26.0
		5th Cycle, May, 1959			
1	90	56	4	78.7	21.3
2	126	24	..	92.0	8.0
3	67	69	14	67.7	32.3
4	61	78	11	66.7	33.3
5	92	55	3	79.7	20.3
		6th Cycle, September, 1959			
1	95	51	4	80.3	19.7
2	80	57	13	72.3	27.7
3	54	80	16	62.7	37.3
4	127	21	2	91.7	8.3
5	91	52	7	78.0	22.0
		9th Cycle, September, 1960			
1	103	40	7	82.0	18.0
2	76	60	14	70.7	29.3
3	58	73	19	63.0	37.0
4	131	19	0	93.7	6.3
5	105	43	2	84.3	15.7
		1st Cycle, October, 1957			
6	28	82	40	46.0	54.0
7	30	91	29	50.3	49.7
8	35	87	28	52.3	47.7
9	29	80	41	46.0	54.0
10	39	82	29	53.3	46.7
		2nd Cycle, March, 1958			
6	34	77	39	48.3	51.7
7	42	73	35	52.3	47.7
8	28	85	37	47.0	53.0
9	22	78	50	40.7	59.3
10	15	95	40	41.7	58.3
		3rd Cycle, August, 1958			
6	15	66	69	32.0	68.0
7	21	74	55	38.7	61.3
8	15	68	67	32.7	67.3
9	14	84	52	37.3	62.7
10	23	81	46	42.3	57.7

Population	AR/AR	AR/PP	PP/PP	AR (%)	PP (%)
		4th Cycle, Janary, 1959			
6	18	75	57	37.0	63.0
7	17	70	63	34.7	65.3
8	22	80	48	41.3	58.7
9	22	86	42	43.3	56.7
10	21	65	64	35.7	64.3
		1st Cycle, May, 1959			
11	64	73	13	67.0	33.0
12	51	80	19	60.7	39.3
13	50	79	21	59.7	40.3
14	48	74	28	56.7	43.3
15	52	75	23	59.7	40.3
		2nd Cycle, October, 1959			
11	59	74	17	64.0	36.0
12	82	58	10	74.0	26.0
13	55	76	19	62.0	38.0
14	68	68	14	68.0	32.0
15	38	84	28	53.3	46.7
		4th Cycle, June, 1960			
11	43	78	29	54.7	45.3
12	75	46	29	65.3	34.7
13	43	84	23	56.7	43.3
14	55	78	17	62.7	37.3
15	35	99	16	56.3	43.7
		5th Cycle, September, 1960			
11	72	68	10	70.7	29.3
12	86	55	9	75.7	24.3
13	53	79	18	61.7	38.3
14	46	87	17	59.7	40.3
15	41	93	16	58.3	41.7

clearly and significantly divergent. The increase of the variance in the successive cycles can be seen in Figure 1; it is much less rapid than in the multichromosomal populations. Out of the 20 samples taken in the populations Nos. 11–15, the heterokaryotypes were more frequent than the Hardy-Weinberg equilibrium demands in 18 samples; Table 2 shows that the excesses of the heterokaryotypes were statistically significant.

TABLE 2

Cycles	Multichromosomal Nos. 1–5	Bichromosomal Nos. 6–10	Bichromosomal Nos. 11–15
1	+0.37	+3.55*	+1.95*
2	+0.80	+2.86*	+1.74*
3	−1.02	+2.10*	+1.75*
4	−0.44	+1.72*	..
5	+3.03*	..	+3.77*
6	+0.48	..	..
9	+0.37	..	..
Totals	+1.36	+5.12*	+4.61*

* Excesses significant at the five per cent level or better.

Conclusions.—Our experiments have shown that divergence of replicate populations is observed both in multichromosomal and in bichromosomal populations of geographically mixed origins. In the former, the divergence is, however, much more striking than in the latter. This is consistent with the prediction derived from our working hypothesis.

It is nevertheless necessary to analyze the experimental situation further. In the experiments of Dobzhansky and Pavlovsky,[5] the replicate populations which started with 20 "founders" diverged more than did those started with 4,000 progenitors, all taken from the same F_2 generation progeny of an interracial cross. The cause of the divergence was the sampling error involved in taking segments of the gene pool of the segregating progeny of interracial hybrids to start the experimental populations. These were the evolutionary "unique events" in Wright's terminology.[7] In the experiments described in the present article, the founders were always 20 in number, but reductions of the population sizes down to this level occurred at four-month intervals, at the start of each cycle. The effective population size, Wright's N, was therefore moderately low, probably between 40 and 100.

The question that arises is whether the divergence of the frequencies of the AR and PP chromosomes observed in replicate populations can be accounted for by accidental inclusion of more AR or more PP chromosomes in the samples of 20 founders. This is most unlikely. The relative frequencies of AR and PP chromosomes in a population are under the control of a strong natural selection driving them towards certain equilibrium values. Although four months, the length of a "cycle," is insufficient to have the populations attain the equilibrium frequencies of the chromosomes, it is clear that the disturbances in the frequencies of AR and PP due to sampling errors would be greatly reduced or extinguished by the force of the selection.

The divergence of the replicate populations arises not from failures to reach equilibria but from shifts in the position of the equilibrium values themselves. This is shown by the different behavior of the multichromosomal and bichromosomal populations. The errors of sampling should be the same in both types of the populations, since the number of founders is kept constant. Yet the multichromosomal populations became far more heterogeneous than the bichromosomal ones. Even though the selection pressure is stronger in the latter because of the deficient vigor of the homokaryotypes, the decisive factor is that the genetic variance available in the bichromosomal populations is more limited than that in the multichromosomal ones. The "unique events" are here antecedent to the taking of the samples of 20 founders; as indicated above, the founders of the bichromosomal populations are descended from a single pair of parents, while those of the multichromosomal ones came from crosses involving 20 different strains of flies. Furthermore, recombination in the bichromosomal populations could take place only in the chromosomes other than the third, since recombination in the PP/AR inversion heterozygotes is effectively suppressed in the chromosomes carrying the inversions.

The question may be asked whether in the present experiments, and in those of Dobzhansky and Pavlovsky,[5] we have been observing the operation of natural selection, of random genetic drift, or of the founder principle? The changes in the relative frequencies of AR and PP chromosomes are clearly selectional in all the populations studied. But if this were the whole story, these chromosomes would eventually have reached uniform equilibrium values everywhere. This is not the case; we observe selection working differently not only in populations started with different kinds of founders but also in replicate populations which we have treated as similarly as possible. Here, random events are clearly involved, providing different genetic backgrounds on which the selection working on the AR and PP chro-

mosomes works differently in different populations. The founder principle (Mayr[8, 9]) is a special case of Wright's random genetic drift, when the latter takes place as a unique accident of sampling.[7]

The five multichromosomal populations in our experiments were becoming increasingly different with time (Fig. 1). In this process of progressive drifting apart, one detail is worth noting, which can be seen most clearly in Figure 2. The rela-

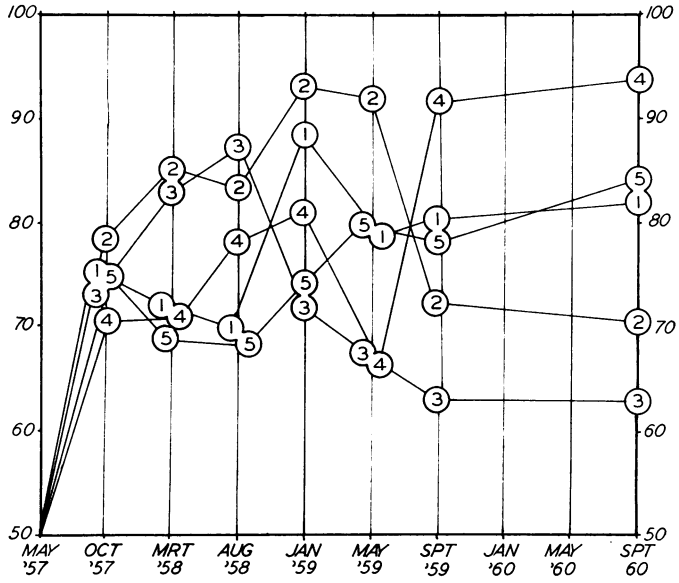

FIG. 2.—Percentages of AR chromosomes in five multichromosomal populations.

tive positions of the populations with respect to the frequencies of AR and PP were changing apparently haphazardly during the early cycles. Thus, population No. 2 had the highest frequency of AR in the 1st, 2nd, 4th, and 5th cycles, but it turned out to be the second from the bottom in the 6th cycle. The lowest frequency of AR was found in population No. 4 in the 1st cycle, No. 5 in the 2nd and 3rd, No. 3 in the 4th, again No. 4 in the 5th, and No. 3 in the 6th. Between the 6th and the 9th cycles, the rank order of the populations seems, however, to have remained about the same, although the lack of data for the chromosome frequencies during the 7th and the 8th cycles renders this conclusion uncertain. Our working hypothesis would, however, explain this relative stabilization of the populations. The drifting apart of the populations is due to the inclusion of only small segments of the field of potential variability in the groups of 20 founders in each cycle; the potential variability should then gradually dwindle as the cycles go on. The founders of the populations of the 1st cycle came from the F_2 generation of an interracial cross; after several cycles, we may be dealing with populations each of which has a genetic variance of about the same order of magnitude as a race or a geographic population normally has in its natural state.

The implications of the findings described in the present article for the general theory of evolution will be discussed elsewhere. It will be sufficient to note here

that the evolutionary events observed in our experimental populations appear to involve changes not in frequencies of single genes but reconstructions of gene systems. These are not microevolutionary but mesoevolutionary changes. Having drifted apart from each other, the replicate multichromosomal populations are genetically perhaps as distinct as natural geographic races. This is not a mere speculation but a testable working hypothesis. Experiments to test it have been undertaken by Angela Solima, and we hope that their results will be described in due course.

Summary.—Interactions of natural selection with random genetic drift have been observed in experimental populations of *Drosophila pseudoobscura*. The evolutionary changes observed in such populations may be determinate and reproducible, or more or less indeterminate, depending upon the genetic constitution of the populations. The changes are reproducible in populations of uniform geographic origin, but variable in populations descended from hybrids between different geographic races. The degree of the variability depends on the number and the genetic constitution of the "founders" of the experimental populations.

This paper is dedicated to Leslie C. Dunn, a friend and a colleague for a quarter of a century, in recognition of his long and distinguished career. Thanks are due to Olga Pavlovsky, who prepared the foundation stocks of most populations, and to Leigh Van Valen who made the chromosome determinations during the absence of one of the authors (Th. D.). Howard Levene has contributed invaluable advice and criticism.

* The work reported in this article has been carried under Contract No. AT-(30-1)-1151, U.S. Atomic Energy Commission.

[1] Dobzhansky, Th., and H. Levene, *Amer. Natur.*, **85,** 247–264 (1951).

[2] Dobzhansky, Th., and O. Pavlovsky, *Evolution*, **7,** 198–210 (1953).

[3] Dobzhansky, Th., "Evolution and environment," in *Evolution after Darwin*, ed. S. Tax (University of Chicago Press, 1960), vol. 2, pp. 403–428.

[4] Dobzhansky, Th., *Australian J. Science*, **23,** 71–74 (1960).

[5] Dobzhansky, Th., and O. Pavlovsky, *Evolution*, **11,** 311–319 (1957).

[6] Prout, T., unpublished data.

[7] Wright, S., in *Population Genetics: The Nature and Cause of Genetic Variability in Populations*, Cold Spring Harbor Symposia on Quantitative Biology, vol. 7 (1955), pp. 16–24.

[8] Mayr, E., *Systematics and the Origin of Species* (New York: Columbia University Press, 1942).

[9] Mayr, E., in *Evolution as a Process*, ed. J. Huxley *et al.* (New York: Macmillan, 1954), pp. 157–180.