

alleles in this series may yet be found. For example, the genetic explanation of the black pigment present in fawn or brindle breeds or of such dark pigment as occasionally appears around the eyes or muzzles of Shorthorns is still uncertain. If such other alleles are found, a complete description of some other breed may include the statement that in it the frequency of  $B$  is .40, of  $B_1$  is .45, and of  $b$  is .15. Naturally it will be difficult to get such complete information except for a few genes which individually have conspicuous effects. For a long time to come it is likely that the practical description of a breed will consist mainly of its averages, or "type" in traits which do not lend themselves readily to

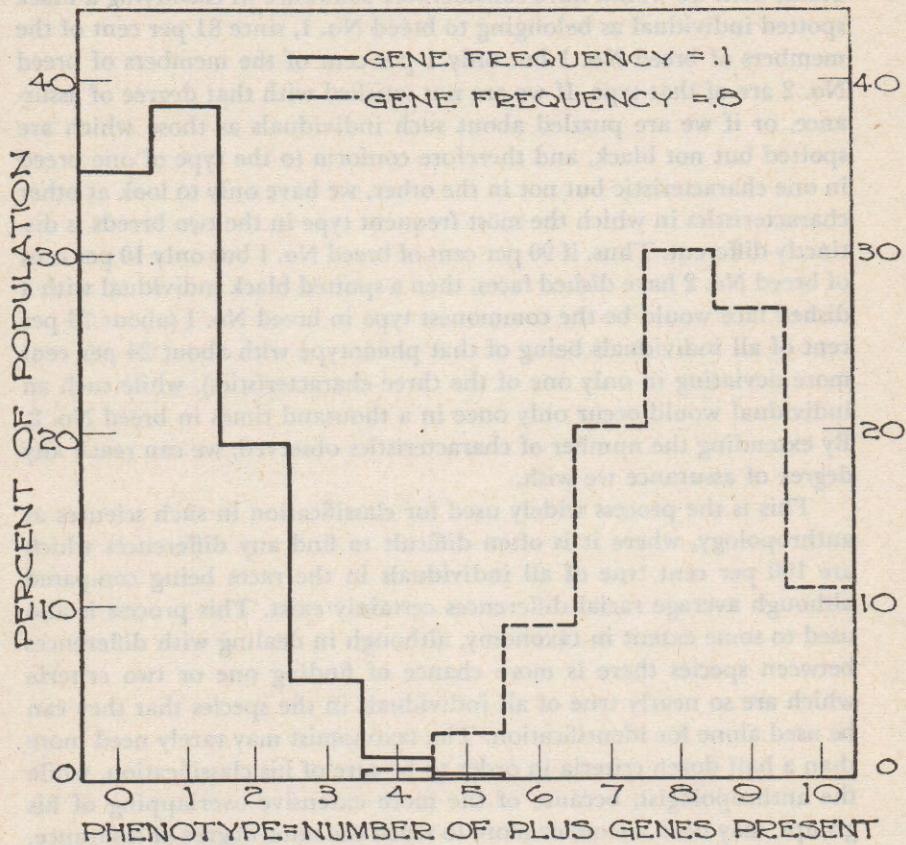


FIG. 12. The distribution of genotypes expected in two random breeding populations each of which is heterozygous for five pairs of genes with equal effects. The genes lack dominance and combine their effects additively. In one population the frequency of the plus gene in each of the five pairs is .1, while in the other it is .8. This will illustrate how two breeds could both have exactly the same *kinds* of genes and yet overlap so little that there would be practically no mistakes in classifying individuals.

numerical averaging, with a few sketchy semi-quantitative comments about its variability in those features.

Two breeds may overlap in every observable characteristic, and yet it may be possible to identify with certainty the breed to which every individual belongs. If 90 per cent of the individuals in breed No. 1 but only 10 per cent of those in breed No. 2 are spotted, we cannot with much assurance classify the breed of an individual by examining it for spotting alone. There are too many exceptions to the "type." But if 90 per cent of the individuals in breed No. 1 are black in their colored areas, while only 10 per cent of the individuals in breed No. 2 are black, then we would have considerable assurance in classifying a black spotted individual as belonging to breed No. 1, since 81 per cent of the members of breed No. 1 but only 1 per cent of the members of breed No. 2 are of that type. If we are not satisfied with that degree of assurance, or if we are puzzled about such individuals as those which are spotted but not black, and therefore conform to the type of one breed in one characteristic but not in the other, we have only to look at other characteristics in which the most frequent type in the two breeds is distinctly different. Thus, if 90 per cent of breed No. 1 but only 10 per cent of breed No. 2 have dished faces, then a spotted black individual with a dished face would be the commonest type in breed No. 1 (about 73 per cent of all individuals being of that phenotype with about 24 per cent more deviating in only one of the three characteristics), while such an individual would occur only once in a thousand times in breed No. 2. By extending the number of characteristics observed, we can reach any degree of assurance we wish.

This is the process widely used for classification in such sciences as anthropology, where it is often difficult to find any differences which are 100 per cent true of all individuals in the races being compared although average racial differences certainly exist. This process is also used to some extent in taxonomy, although in dealing with differences between species there is more chance of finding one or two criteria which are so nearly true of all individuals in the species that they can be used alone for identification. The taxonomist may rarely need more than a half dozen criteria in order to be sure of his classification, while the anthropologist, because of the more extensive overlapping of his groups, may need a score or more to reach the same degree of assurance. Table 9 shows in some detail the average number of deviations from "type" which may be expected per individual in a population where  $n$  independent characteristics are being examined and in each of those characteristics  $t$  is the fraction of individuals which deviate from type. The average number of deviations expected in an individual is  $nt$ . The

figures which follow the  $\pm$  signs are standard deviations computed by the formula,  $\sigma = \sqrt{nt(1-t)}$ , which must be interpreted with reservation on account of the distinct skewness of the distribution where  $t$  is far from .5. The standard deviations show that, when the average num-

TABLE 9

AVERAGE NUMBER OF CHARACTERISTICS PER INDIVIDUAL EXPECTED TO DEVIATE FROM THE "TYPE" OF THE GROUP WHERE  $n$  CHARACTERISTICS ARE OBSERVED AND  $t$  IS THE FRACTION OF INDIVIDUALS WHICH DEVIATE FROM TYPE IN EACH RESPECT

$t$	$n$		
	5	10	20
.01.....	.05 $\pm$ .2	.1 $\pm$ .3	.2 $\pm$ .4
.05.....	.25 $\pm$ .5	.5 $\pm$ .7	1.0 $\pm$ 1.0
.20.....	1.0 $\pm$ .9	2.0 $\pm$ 1.3	4.0 $\pm$ 1.8
.40.....	2.0 $\pm$ 1.1	4.0 $\pm$ 1.5	8.0 $\pm$ 2.2

ber of deviations from type is large, there may be no individuals which are exactly like the "type." For example, when  $n = 20$  and  $t = .40$ , the average individual will deviate from "type" in 8 respects, only about one-eighth of them will deviate from type in as few as 5 respects, only 1 in 20 in as few as 4 respects. Less than 1 in 27,000 will conform exactly to type in all 20 respects. In this sense it may be true that in finite populations there is *no such thing as an average individual*, but that does not impair the usefulness of the average for describing the group. The description of the group will be more complete if something is also stated about the variation to be expected in each characteristic.

The same principles used in classifying individuals may be extended to the classification of groups wherever it appears likely that the groups are random samples or systematic samples selected fairly. One will often see a Jersey cow which is broken-colored and in that respect is more like the Guernsey breed than the Jersey; but one will almost never see a herd of Jerseys which are all broken-colored, although that is the usual description of a herd of Guernseys. A broken-colored Jersey without any black pigment, although rare, sometimes occurs and might be mistaken for a Guernsey if color alone were considered. One would never see a herd of Jerseys which were all broken-colored and free from black pigment.

#### DIFFERENCES BETWEEN SPECIES

A species is a more or less continuous and interbreeding group as actually found in nature. To practical taxonomists *discontinuity* between two groups is the criterion of whether or not they are really different species. This leaves room in some cases for disagreement as to

whether the continuity and freeness of interbreeding are sufficient to justify calling the population one species. Two groups of similar individuals which are not connected by any intermediates are universally recognized as valid, or "good," species. On the other hand, two groups of individuals may be widely different in the averages of many of their traits and yet, if they are connected by a complete series of intermediate individuals, most practical taxonomists would regard them as really one species with a tendency to exist in several varied forms or subspecies. This situation is much the same as might be encountered in deciding in a mountainous region whether one were describing two distinct mountain ranges or one group of mountains. If there were a broad and deep valley separating the mountains, everyone would agree that they were two distinct ranges. If the intervening space were occupied by a group of large hills or low mountains, irregular in outline, some might still wish to call them two mountain ranges, but others would think the whole group should be called by one name. The fact that there is sometimes room for argument about whether a group is one species with two subspecies or is two separate species does not contradict the fact that there are such things as species groupings in nature, any more than the reality of mountains is contradicted by a dispute as to whether some particular elevation is one mountain with two peaks or is two separate mountains.

Besides being a matter of practical convenience, the taxonomic system attempts to describe the general fact that the existing plants and animals are not scattered uniformly all through the infinite field of possible combinations of form and function and appearance and behavior, but are definitely and irregularly clumped around certain types, most of which are separated from the nearest similar types by a considerable void of conceivable intergrades which do not occur. This is like the distribution of matter in astronomical space, which is highly discontinuous and irregular with definite clumps, like planets, stars, etc., which are themselves clustered in irregular groups like solar systems, constellations, nebulae, etc. The interesting cases about which there is dispute as to species classification are those where a few of those conceivable intergrades actually do occur and form a nearly continuous connection between two nearby clumps, or where the clump itself seems to have two or more separate centers of density around which the existing forms are clustered more closely than elsewhere in the clump. Quite conceivably these may be centers of incipient formation of new species, but in particular cases they may be still so close together and connected by so many intergrades that there can be no reasonable argument for elevating them to specific rank.

The early taxonomists often had an exaggerated idea of the supposed uniformity of wild species. Traces of that idea still linger in biological literature. Much of this doubtless comes from a more or less unconscious deference to the opinions of Linnaeus (Carl von Linné, 1707-1778), the Swedish naturalist who devised the present binomial system of naming plant and animal species.<sup>1</sup> Today it is recognized with increasing clarity that the more one studies a wild species, the more differences he finds among the individuals composing it. If he makes several different collections of individuals of the same species, each collection having been trapped in a different locality, he will usually find each collection different enough from the others that if he were to see another *collection* made from one of those same localities he would be able to identify the locality from which it came, although he would not be able surely to identify the locality of each animal presented to him singly. In short, even the "good" species differ from locality to locality, particularly among animals which do not travel far.<sup>2</sup>

The definition of a species has a peculiar relation to the idea of special creation. Linnaeus believed firmly in the special creation of each individual species, as did most other people of his time. The belief in special creation necessitated a belief that there was a real nature-made difference between species. Hence, there would be some fundamental valid distinction between species which, if one could only discover it, would serve as an accurate touchstone for deciding in accordance with natural laws what was and what was not a true species. Linnaeus, as well as all contemporary and succeeding naturalists, recognized that orders, genera, etc., were divisions made by man, as a matter of convenience, for classifying together organisms which had certain general resemblances. They regarded subspecies, varieties, breeds and such divisions of a species as matters of convenience or as the result of man's handiwork and therefore outside the scheme of nature. This insistence that the difference between species is a fundamental one on an entirely different basis from other differences in classification explains the intensity of many arguments about what constituted a real species in particular cases. The acceptance of the idea of organic evolution carries with it

<sup>1</sup> For example, on page 128 of *Biology and Its Makers*, by Locy, we read: "Ray had spoken of the variability of species but Linnaeus in his earlier publications declared that they were constant and invariable"; and on page 129: "While Linnaeus first pronounced upon the fixity of species, it is interesting to note that his extended observations upon nature led him to see that variation among animals and plants is common and extensive, and accordingly in later editions of his *Systema Naturae* we find him receding from the position that species are fixed and constant. Nevertheless, it was owing to his influence more than to that of any other writer of the period, that the dogma of fixity of species was established."

<sup>2</sup> Sumner, Francis B. 1934. "Taxonomic Distinctions Viewed in the Light of Genetics." Amer. Nat., 68:137-149.

the consequence that species differences are no less and no more man-made than differences between genera or orders. The difference between species is taken out of a special category and becomes only a matter of degree in the general system of classification.

The modern genetic idea of species differences is that they are similar in kind to the breed differences explained above but are much more extreme, so that discontinuity is an essential part of the species definition. Also, in most cases, the two sets of genes are so different that they will not work together harmoniously in producing their physiological effects. Hence, crosses between species are usually (but not always) impossible or, if possible, are usually sterile. Since inheritance is in duplicate, it will sometimes happen that the first cross hybrids which have a complete but single set of the genes from each parent can function all right in their own physiology (as is the case with the mule) but cannot reproduce because the genes from the different kinds of parents are so unlike that they will not pair properly enough for the reduction division to take place. Even if reduction does occur, the resulting sample of genes will rarely have all that are necessary for harmonious functioning. It is somewhat as if one were to try to make a workable automobile from two different kinds by fastening together cylinders from one, pistons from the other, a fuel pump from the first, a carburetor from the second, cam shaft from one, distributor from the other, etc. If the two makes of automobiles were very different, it would be the rarest kind of a coincidence if such an automobile would run at all.

Discontinuity of ancestry between breeds is a fairly recent thing and does not go back much farther than the beginnings of the herdbooks in some cases. Separateness of ancestry between species is a vastly older thing, going far back into geologic time in most cases. Whenever two subgroups of a single group cease to interbreed, their genetic averages tend to become different, either through the random drift of gene frequencies in the sampling which takes place each generation during the reduction division, or because the direct effects of selection may be different if the subgroups live in different regions or otherwise occupy different ecological niches. Both processes can be supplemented by mutations which may not be the same in both groups. Once discontinuity of interbreeding is established, it is easy to see how two subgroups of a species not only might but must in geologic time drift apart so far that they could not cross.

The most puzzling problem in the origin of species is how the discontinuity of interbreeding first arose in each case of a dividing species. Natural selection differently directed in different regions might have made portions of a species become unlike; but, unless they were so dis-

tinctly separated that there was almost no interbreeding between the two groups, it seems likely that crosses between them would usually prevent that from going far enough to form two separate species. Discontinuity of interbreeding might have been brought about by geographic isolation; but, as nearly as we can interpret the geologic record, changes of sufficient magnitude to bring that about seem to have been too slow to provide enough isolation to account for the facts of evolution. Irregularities in the chromosome mechanism, such as polyploidy or frequent and extensive inversions, might have brought about discontinuity of interbreeding even without geographic isolation. These probably did play a considerable part in the evolution of plants, but it is generally thought that their effect on animals must have been much less important, both because the chromosome numbers among so many of the modern mammals are so similar and because self-fertilization and asexual reproduction are not possible in higher animals. One with a chromosome abnormality probably could not find a mate like itself, and it or its offspring from normal mates would usually be sterile. Assortive mating may have played a part, although most students of the subject now think the importance of that was overestimated by Darwin in his writings on sexual selection. Inbreeding is a powerful force for differentiation of groups, but there is uncertainty about how much of that actually occurs in nature.<sup>3</sup> Perhaps all these processes played a part, varying in importance in different cases. The subject is full of interest to the philosophy of biology but contributes little to the routine practice of the animal breeder since he can establish discontinuity in his own breeding operations in whatever way he wishes.

#### SUMMARY

1. Because of the lessened variability of averages it is often possible to distinguish breeds (or other groups) whose averages are not very far apart even though the individuals within each breed vary widely and the breeds overlap.
2. Breeds may differ in one's having a gene which is entirely absent from the other, but more commonly they differ in the frequencies of genes which are present in both.
3. By considering a sufficient number of characteristics in which the averages of two breeds are different, it is possible to identify with any desired degree of certainty the members of each, even though both breeds overlap each other's range in all those characteristics.

<sup>3</sup> For a general survey of the possibilities in that, see: Wright, Sewall. 1940. "Breeding Structure of Populations in Relation to Speciation." *Amer. Nat.* 74:232-48. See also the chapter on isolating mechanisms in T. Dobzhansky's *Genetics and the Origin of Species*. 1941. Columbia University Press.

4. Differences between species are thought to have the same genetic basis as the differences between breeds but are so much more extreme that crosses between species are usually sterile, if possible at all.

5. A species is a more or less continuous and interbreeding group as actually found in nature. To the practical taxonomist the discontinuity rather than the magnitude of the differences between species is the most satisfactory criterion of whether they are really two species or a single variable one.

6. The most puzzling problem in the origin of species is how the discontinuity of interbreeding first arose in each case. Several explanations are possible, but it is not certain which of them has been the main explanation in most cases.

## CHAPTER 10

### The Means Available for Controlling Animal Inheritance

Man can do only a few kinds of things to change the heredity of his animals. First of all he has some power to decide which of them shall have many offspring, which shall have few and which shall have none. That is selection. Selection has always been practiced by animal breeders, and among many of them it is almost the only breeding method used. There can be various degrees of it. It can be based on individuality, on ancestry, on progeny, or on combinations of those in different degrees.

In the second place, since those chosen to be parents will not be exactly alike, either in pedigree or in their own somatic appearance and performance, there are many different ways in which the breeder may decide which of the chosen males are to be mated to which of the chosen females. But in their genetic nature and practical consequences all these systems of mating, if they deviate at all from random mating within the group of chosen parents, may be classified as the mating of like to like or as the mating of unlikes. Likeness or unlikeness may be based either on blood relationship or on individual appearance and performance.

Mating systems, wherein the mates have a closer blood relationship to each other than if mating were at random, are *inbreeding* in the broad sense of that word, although most animal breeders reserve that term for the closest degrees of inbreeding. Mating systems wherein the mates are less closely related to each other by blood than they would be under random mating are *outbreeding*. The consequences and uses of inbreeding and outbreeding were but vaguely known in pre-Mendelian days. Inbreeding and outbreeding are still used only a little by most breeders of purebreds; but now that the reasons for their results are understood and measures of their intensity have been found, considerable increase in their use seems likely in the future. The results of mating like to like or of mating unlikes on the basis of their own individual characteristics, regardless of pedigree, are very different from the results

of inbreeding and outbreeding. Since selection and these four general systems of mating are the only tools with which man can change the inheritance of his animals, it is important for the practical breeder to know what kind of change each is apt to produce, what things each will do well and what each will do poorly or not at all, what are the chief dangers or difficulties in each, and what are the most useful means of overcoming those dangers and difficulties.

Any of these four kinds of mating systems can be practiced in combination with or in alternation with any of the others. They are almost always accompanied by some degree of selection. That makes possible an almost infinite number of specific breeding plans. The probable consequences of each of those may be predicted in a general way, but the chance involved in Mendelian segregation and recombination will leave room for surprising results in individual cases. Moreover, the combination of one mating system with another will sometimes give results which are not simply the sum of what each would accomplish if practiced separately—an epistatic effect among the breeding systems themselves, so to speak! There is no immediate prospect that reliable predictions of the outcome of breeding plans can become so detailed and accurate that they will remove from the business of livestock breeding the sporting element of hope and uncertainty which has been one of its great attractions and has led many wealthy men to take it up as a hobby.

Man's knowledge of how the mechanism of inheritance operates is fairly complete, but that knowledge has not yet given him any ability to interfere with some of the processes so as to change their outcome in the direction he wishes. Thus, no way has yet been found to control the segregation of genes so as to produce from heterozygous parents gametes which contain more than a random proportion of the desired genes. Nor is there any prospect that such control over assortment at segregation will ever be achieved. All that man can yet do in this respect is to select from among those animals available for parents the ones which suit him best and then accept whatever gametes they produce. But even after the gametes are produced, he cannot select those which most nearly have the genes<sup>1</sup> he prefers or promote the union of those which are most like each other or least like each other. All he can do is to let the array of gametes from the chosen sire unite at random with whatever

<sup>1</sup> P. C. Mangelsdorf has shown (1931, *Proc. Nat. Acad. Sci.*, 17:698-700) that it is physically possible by the use of certain mechanical sieving methods to separate corn pollen grains which carry the gene for "sugary" from those pollen grains which carry the allelic gene for "starchy," but the method has not found practical application. Several attempts to separate male-producing from female-producing spermatozoa by physical or chemical methods have been tried without success.

ova the chosen dam has produced. There is extensive evidence from plants that selective fertilization exists in nature,<sup>2</sup> but the general importance of this is in some doubt. In the mildest forms of selective fertilization in plants, pollen tubes containing genes like those in the tissue of the plant being fertilized grow down the style toward the ovule a little faster (or a little more slowly) than pollen tubes which carry unlike genes. This gives some kinds of genes an advantage over others in reaching the ovule and fertilizing it, although the handicapped genes are perfectly capable of doing so if they have no competition from the favored genes. Whether there is anything to correspond to this in the higher animals is uncertain, although the processes of animal courtship may possibly indicate that there is considerable assortive mating in nature.<sup>3</sup> The most extreme forms of selective fertilization in plants are cross-sterility or self-sterility, which are phenomena well-known among certain horticultural crops such as some varieties of apples. Definite genes for self-sterility, often long series of multiple alleles, are well known in some cases (*Genetics* 27:333-38. 1942). Occasionally it happens with animals that a female is bred several times to one male without conceiving but conceives promptly when bred to another male, although the first male was fertile in matings with other females. Yet such cases rarely furnish any very plausible indication of selective fertility since one cannot know whether that female would have conceived if she had been bred again that last time to the first sire. There are almost never enough such cases at any one time for a statistical investigation to be decisive.

Neither can the breeder change the laws of Mendelism nor the number of genes nor their linkage relations. He cannot change their mutual physiological interactions, such as dominance, except as he can find and increase the frequency of other genes which modify in the desired way the physiological effects of the first genes. To a very limited extent genes can be changed into other kinds by such violent treatments as exposure to X-rays, radium, etc.; but such treatments usually result in a high degree of sterility in the treated animals, and the mutations produced are so nearly all undesirable<sup>4</sup> that the production of mutations offers no help to the practical breeder.

This leaves as the breeder's only practical means of controlling the heredity of his animals his partial freedom to decide how many offspring each animal shall have and his freedom to choose, within the

<sup>2</sup> Jones, D. F. 1928. *Selective Fertilization*. 163 pp. University of Chicago Press.

<sup>3</sup> For a summary of knowledge concerning that in *Drosophila*, see *Biological Symposia* 6:277-79. 1942.

<sup>4</sup> Gowen and Gay found in their material (*Genetics* 18:1-31) that 92.2 per cent of all mutations were actually lethal and many of the rest obviously caused low viability.

group selected for breeding, which shall be mated with which. These opportunities the breeder possessed and used before Mendel's work was known. But he used them with many mistakes, and he neglected many opportunities to use mating systems which could have forwarded his work. Full use of the genetic knowledge available today should make the mistakes in selection fewer, although it cannot prevent them all, and should enable freer use of inbreeding, outbreeding, and the crossing of types than breeders would have dared before the principles underlying those practices were understood. Perhaps the analogy is not too fanciful if we compare pre-Mendelian animal breeding with the animal breeding now possible in the same way we would compare the

Systems of mating which may be combined with various kinds of selection	Mating like to like	<i>By pedigree:</i> Inbreeding, including linebreeding, staying within one family, etc.
	Random mating	<i>As individuals:</i> Mating big with big, little with little, medium with medium, compact with compact, rangy with rangy, active with active, sluggish with sluggish, etc.
	Mating unlikes together	<i>By pedigree:</i> Outbreeding, ranging from species crossing through crossbreeding, to crossing strains within the breed.  <i>As individuals:</i> Compensating for defects, crossing extremes to produce intermediates, mating large with small, coarse with refined, active with sluggish, etc.

The above classification may make clearer the kinds and definitions of breeding systems.

common practice of making soap on farms less than a century ago with the processes now used in modern soap factories. The fundamental chemistry of soapmaking has not changed; but the product has changed tremendously in variety, usefulness, adaptability and dependability, as a result of accumulated refinements in the purity of the ingredients, closer control of temperatures and concentrations and the use of small amounts of certain ingredients whose effects were formerly but dimly understood.

*Read  
Carefully*

## CHAPTER 11

### How Selection Changes the Genetic Composition of a Population

Causing or permitting some kinds of individuals to produce more offspring than other kinds do is selection. It is the number raised and added to the breeding herd rather than the number born which matters, since those which are born but get no chance to reproduce cannot affect the composition of the future population. Under some circumstances selection may quickly cause large and permanent changes in the population. Under other circumstances it may cause marked changes, but the moment selection is relaxed the population returns to its original condition. Under still other circumstances selection may be almost powerless to produce any change unless it is combined with some mating system like inbreeding.

The changes which selection produces in the underlying genetic composition of a population can rarely if ever be seen or measured directly, since the observer will not know what genes are present, nor the frequencies of each, nor the frequencies of their various combinations, nor the amount of change which selection makes in those.

Selection creates no new genes. It merely causes the possessors of some genes or of some combinations of genes to have more offspring than those which lack those genes or combinations. Its primary genetic effect is to change gene frequency and the frequency of gametes carrying certain gene combinations. All its other effects are consequences of that, and their magnitude depends on how much the gene and gamete frequencies were changed. Changes in gene frequency are permanent even if selection ceases, unless counter-selection in the opposite direction begins and is effective. Changes in gamete frequency, other than those which result from changes in gene frequency, are temporary because the genes recombine when segregation takes place in forming the gametes for the next generation. Because of this segregation and recombination, the gains from selecting for epistatic differences are temporary, and selection must be continued merely to hold those, although the gains from selecting for additive differences are permanent and remain even when selection is relaxed and abandoned.

Selection can be creative only in the sense that new types can be

produced when selection has moved the average of the population far from the original position, as is shown in Figure 15. For example, suppose there are five desirable genes, each having a frequency of .1. If mating is random, only one individual in ten billion will be homozygous for all five of the desired genes. For practical purposes this is nonexistent. But if selection in favor of the individuals with the larger number of these genes were practiced long enough to raise the frequency of each gene to .7, then about 28 individuals in each thousand would be homozygous for all five genes. This is frequent enough that some of them would be found. In that sense selection may be said to have created something new, somewhat as an architect can create a building of an original design, although all the materials were already in existence before he began.

#### ONE PAIR OF GENES

If from a population containing the three genotypes,  $AA$ ,  $Aa$ , and  $aa$ , only the  $AA$  individuals are allowed to reproduce, the next generation will be homozygous for  $A$  which will then have a frequency of 1.0 in that population. Selection will in one generation have done all it could if it were to be continued for many generations. Similarly, if only the  $aa$  individuals had offspring, the whole population in the next generation would be homozygous for  $a$ , the frequency of  $A$  would have fallen to zero and change by selection would come to an end, as far as that pair of genes is concerned. In actual practice, selection can practically never be that accurate and extreme. Instead, some of the undesired genotypes are kept and some of the desired ones are culled, either because there are not enough of the desired ones to permit culling all the others, because the breeder is careless, or because dominance and environmental variation mislead him. The net result is that selection increases the frequency of the favored gene by at least a little each generation and thus leads to some change in the genetic composition of the population.

There may not be enough individuals to permit discarding at once all of the undesired ones. If all Shorthorn breeders were to decide suddenly that they wished their breed to be white, there are probably not enough white Shorthorn bulls alive that every breeder could secure a white bull to head his herd, even if no attention were paid to anything but color. Some would have to use roan bulls for at least a generation until the number of white bulls had increased. As for cows, they would not only have to use all the whites but probably all of the roans and even some of the reds. In the next generation they might have enough whites and roans that they could cull all of the reds, but it would prob-

ably be several generations before they could discard all of the roan cows.

The animal is the smallest unit which the breeder can select or reject, but in the animal the genes come in pairs rather than singly. This makes progress by selection slower than if selection could be gene by gene. For example, suppose 25 per cent of the animals in a Short-horn population were white, 50 per cent were roan, and 25 per cent were red, and the breeders should suddenly decide that they wanted the breed red, but could only afford to cull half of each generation. The best they could do in the first generation would be to keep all of the red animals and half of the roans, discarding the whites and the other half of the roans. There were enough genes for red in the population that the breeders could have discarded all the genes for white and have changed the population completely in one generation if they could have selected gene by gene, but instead they must select or reject the genes a pair at a time. Selection between zygotes thus changes gene frequency more slowly than selection between gametes would. It seems unlikely that the breeder will ever be able to select between gametes, although some natural selection at that stage does take place in plants and perhaps some also in animals.

The effects of environment may duplicate or hide the effects of genes, thus causing the breeder to discard some animals which he would keep and to keep others which he would discard if he knew what their genotypes really were. Dominance may do the same thing by preventing him from knowing which individuals are *AA* and which are *Aa*. Naturally every mistake of this kind means that the undesired genes are transmitted by more individuals and the desired genes by fewer than would have been the case if these mistakes had been avoided. Such mistakes lower the rate at which selection increases the frequency of the desired gene. They do not stop the process but merely cause more time to be required to produce the same amount of change.

#### RATE OF CHANGE IN GENE FREQUENCY

The amount which gene frequency is changed by one generation of selection could be computed if we knew the rates of reproduction for each of the three genotypes and the frequency of each genotype. If the numbers of offspring produced by the same number of *AA*, *Aa*, and *aa* individuals are in the ratio: 1 : 1 — *hs* : 1 — *s*, we can consider *s* as a measure of the intensity of selection against the *aa* individuals and *hs* as measuring the intensity of selection against the heterozygote. For example, if for every 100 offspring which *AA* individuals produced, the same number of *Aa* individuals would on the average produce 95 offspring

*Important*

and an equal number of  $aa$  individuals would produce only 80 offspring, then  $s$  would be .2,  $hs$  would be .05, and hence  $h$  would be .25. In a random breeding population the change ( $\Delta q$ ) produced in the frequency ( $q$ ) of gene  $A$  by one generation of selection is a tiny bit larger than  $sq(1-q)[1-q+h(2q-1)]$ . The height of the three curved lines in Figure 13 shows how large  $\Delta q$  would be at each value of  $q$  and for each of three conditions of dominance. The values of  $h$  are, respectively 0, .5, and 1. Selection is most effective, ( $\Delta q$  is largest) when gene frequency is somewhere near the middle of its possible range, and is least when  $q$  is near zero or 1.0.

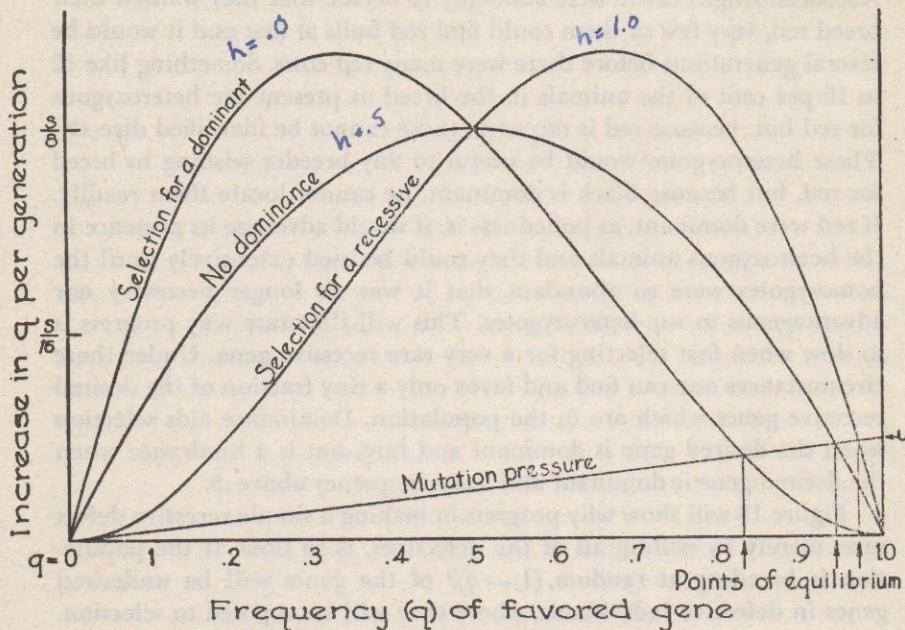


FIG. 13. Rate of change in gene frequency under constant selection,  $s$ , which is opposed by a constant mutation rate,  $u$ . Drawn with  $u$  equal to  $.03 s$ , which is rather weak selection. The height of the curved lines indicates the rates at which selection would change the frequency,  $q$ , of the desired gene under the three conditions specified for dominance. The height of the straight line, "mutation pressure," indicates the rate at which mutation would change gene frequency in the absence of selection. The difference between the height of the curved lines and the height of the "mutation pressure" line indicates the net rate at which gene frequency is changed by selection and opposing mutation. Arrows indicate gene frequencies at which selection pressure and opposing mutation pressure are equal. (After Wright in *Genetics*, 16:104.)

Dominance of the favored gene is a help to selection when that gene is rare, but a hindrance when the favored gene is more abundant than the undesired recessive. Thus, those who are breeding Hereford cattle for polledness are now fortunate that the gene they want is dominant,

because that enables them to distinguish between the heterozygotes and the homozygous recessives. The gene for polled is still rare enough in the Hereford breed that its frequency can be increased by using heterozygous animals for breeding purposes and there are not yet enough homozygous polled animals to permit discarding all of the heterozygotes. If the trend toward polledness continues long enough, the time will come when the favored gene will be more abundant than its recessive allele. Then the breeders will want to discriminate more strongly against the heterozygous polled animals. When that time comes, they will wish that the gene for polledness were recessive. If breeders of Aberdeen-Angus cattle were suddenly to decide that they wanted their breed red, very few of them could find red bulls at first and it would be several generations before there were many red cows. Something like 12 to 16 per cent of the animals in the breed at present are heterozygous for red but, because red is recessive, those cannot be identified directly. These heterozygotes would be useful to any breeder wishing to breed for red, but because black is dominant, he cannot locate them readily. If red were dominant, as polledness is, it would advertise its presence in the heterozygous animals and they could be used extensively until the homozygotes were so abundant that it was no longer necessary nor advantageous to use heterozygotes. This will illustrate why progress is so slow when first selecting for a very rare recessive gene. Under those circumstances one can find and favor only a tiny fraction of the desired recessive genes which are in the population. Dominance aids selection when the desired gene is dominant and rare, but is a hindrance when the desired gene is dominant and has a frequency above .5.

Figure 13 will show why progress in making a simple recessive defect rare, merely by culling all of the defectives, is so slow. If the population is breeding at random,  $(1 - q)^2$  of the genes will be undesired genes in defective individuals, where they will be exposed to selection. But  $q(1 - q)$  of the genes will be undesired genes hidden in the heterozygous individuals. Thus  $q$  of the undesired genes will be in heterozygous individuals where dominance shields them against selection. This becomes a larger and larger fraction as the undesired recessive becomes rarer. Consequently, although discarding the defectives is always to be recommended if the defect is serious, and will decrease the proportion of defectives rapidly when they are abundant, it produces only slight changes when the defectives are already rare. Selection is abundantly able to make an undesired gene rare but is almost powerless to eliminate it entirely from the population.<sup>1</sup>

<sup>1</sup> Among the nonrandom mating systems, only inbreeding alters this situation much. Under it the heterozygotes shield from selection only  $q(1 - F)$  of the undesired genes,  $F$  being the inbreeding coefficient (Chap. 21) and ranging from 0 to +1.0.

It is sometimes said that selection makes more rapid progress at first and that further progress per generation becomes slower and slower. Inspection of Figure 13 will show that this need not be so. If the desired gene is very rare, the increase in gene frequency made by selection would be small at first, simply because there is not enough genetic variability in the population. As the gene frequency rises toward the values near the middle of its range, progress would become faster and faster until it reached a maximum. After that it would decrease.

In actual practice  $s$  cannot be large against many genes unless each is rare, as lethals are. If 10 per cent of the population is  $aa$ , 10 per cent is  $bb$ , 10 per cent is  $cc$ , etc. each of these being undesirable; and if many such traits are to be considered, it will be impossible to find animals which are free from all of these defects. In any population which is constant in numbers, at least enough offspring must reach breeding age to replace their parents. Some animals which have a few defects must be saved for breeding because they are better than average in other respects. Any mistakes caused by dominance or by the confusing effects of environment will also decrease  $s$ . Since many genes affect the net desirability of each animal it is reasonable to use a general value something like .01 for  $s$  in these formulas, although of course  $s$  will vary widely from one gene to another. It will be as high as 1.0 for lethals and doubtless will approach zero for many genes. In actual practice  $s$  is likely to change as selection changes the population. Then more intense selection for some genes may become possible, and less intense selection than formerly may be needed for others.

We can compute how many generations will be required to change gene frequency from one value to another if  $s$  is known and remains constant and if the frequencies of the different genotypes are known. The figures necessary for such computations<sup>2</sup> are shown in Table 10 for a random mating population. These figures will show how much time selection may need for increasing the frequency of a gene by a large amount. Other than for demonstrating this principle Table 10 is not of much practical use since one will rarely know the frequency of any of the genes in his herd. Still more rarely will he know how intense his selection for each gene actually is. The following example will show how Table 10 may be used. The time required to change  $q$  from .01 to .05 when selecting for a complete dominant is  $\frac{1.69}{s}$  generations, which equals 169 generations if  $s = .01$ , but only 1.69 if  $s = 1.0$ . In either case

<sup>2</sup> They are derived from integrating the equations for  $\Delta q$  under those three special conditions of dominance. The correction factors in the last column and in the bottom row allow for a denominator which is not quite 1.0.

there is a correction, 1.61, to be subtracted. The final answer is a little more than 167 generations for the mild selection, and only .08 generations for the intense selection. For the same problem, except that selection is for a complete recessive, the final answers are 8,083 generations for the mild selection and .04 generations for the intense selection. The

TABLE 10  
APPROXIMATE TIME REQUIRED FOR SELECTION TO INCREASE THE FREQUENCY ( $q$ ) OF A FAVORED GENE BY VARIOUS AMOUNTS

$q$ to be Changed From $q_1$ to $q_2$		Time, Expressed in $1/s$ Generations			Correction Factor $x$
$q_1$	$q_2$	Selection for a Complete Dominant ( $h = 0$ )	Selection When There Is No Dominance ( $h = .5$ )	Selection for a Complete Recessive ( $h = 1.0$ )	
.01	.05	1.69	3.30	81.65	1.61
.05	.10	.81	1.49	10.75	.69
.10	.20	.95	1.62	5.81	.69
.20	.30	.72	1.08	2.21	.41
.30	.40	.68	.88	1.28	.29
.40	.50	.74	.81	.91	.22
.50	.60	.91	.81	.74	.18
.60	.70	1.28	.88	.68	.15
.70	.80	2.21	1.08	.72	.13
.80	.90	5.81	1.62	.95	.12
.90	.95	10.75	1.49	.81	.05
.95	.98	30.95	1.89	.98	.03
.98	.99	50.70	1.41	.71	.01
.99	.995	100.70	1.40	.70	.00
From answer in generations subtract:		$x$	$2x$	$x + 1/q_1 - 1/q_2$	

corrections in the right-hand column, to be used as indicated at the bottom of the table, are unimportant when  $s$  is small but are considerable when  $s$  is large.

#### EQUILIBRIUM BETWEEN SELECTION AND OPPOSING MUTATION

Mutations are very rare, and nearly all of them are harmful. Therefore they are to be considered as opposing selection, although perhaps at extremely rare intervals a favorable mutation does occur. The more abundant the desired genes are, the more of them are exposed to the risk of mutating to something less desirable. Hence the higher the frequency of the desired gene, the more strongly mutation tends to lower that frequency. That is shown in Figure 13 by the height of the straight line which shows "mutation pressure."

Selection will be a far more powerful force than mutation except when the undesired gene is very rare. This gives rise to an equilibrium

value for gene frequency at a point where the undesired gene is already so rare that the few undesired genes eliminated each generation by selection are balanced by an equal number newly produced by mutation. A numerical example may make this clearer. If there is perfect selection against recessive gene  $a$  ( $s = 1.0$ ) in a random breeding population of a million individuals, and if the mutation rate from  $A$  to  $a$  is such that in each generation one out of every million  $A$  genes mutates to  $a$ , then the equilibrium point for  $q_A$  will be about .999. At that point the proportion of  $aa$  individuals born will be 1 in 1,000,000, while about 1 in 500 will be  $Aa$ . Culling the  $aa$  individuals will in each generation eliminate two  $a$  genes from every 2,000,000 genes in that allelic series in that population: At the same time there will be 1,998,000  $A$  genes exposed to mutation. A mutation rate from  $A$  to  $a$  of about 1 in 1,000,000 will provide two new  $a$  genes each generation to replace the two culled by selection, and the frequency of  $A$  will not change, even though selection for it continues.

A general formula for the value of  $q$  at the equilibrium point may be had by letting  $s$  equal the selection coefficient as before and  $u$  equal the net rate of opposing mutation. Then an undesired complete recessive is at equilibrium when its frequency is approximately  $\sqrt{u/s}$ . The corresponding equilibrium point for an undesired dominant is  $u/s$  and for an undesired gene where the heterozygote is exactly intermediate in undesirability is  $2u/s$ . All these frequencies at equilibrium will be low if  $s$  is large. Complete dominance shields the undesirable recessive from selection to such an extent that at equilibrium its frequency is  $\sqrt{s/u}$  times as large as that of an equally undesired dominant. Since  $s$  to be detectable would usually need to be larger than .01, and  $u$  seems usually to be something of the order of .000,01 to .000,000,1, it is not at all surprising that undesired recessive genes should be anywhere from thirty to a thousand times as frequent as undesired dominant genes in a population which has long been under selection. This may be the major explanation for the widely observed fact that recessive genes, uncovered by inbreeding or otherwise, are nearly always less desirable than their dominant alleles.

#### ABUNDANCE OF RECESSIVE UNDESIRABLE GENES

While  $\sqrt{u/s}$  is a very low equilibrium frequency for any one gene which is seriously undesirable, yet if the number of loci which can mutate to undesired alleles is several hundreds or a few thousands, as the evidence indicates, then the total number of lethal genes which can exist in the population is large. It might happen that nearly every individual would carry at least one lethal, although only rarely would the

male and female which mate together both carry the same lethal gene. If so, the proportion of defective individuals born would be small, as long as the population is large and mating practically at random, but would increase sharply whenever inbreeding is begun. As a numerical example, suppose  $u$  is .000,001,  $s$  is 1.0 as it would be for a lethal gene, and  $h$  is zero as it would be for a completely recessive lethal which had no desirable effects at all. Then the frequency of a lethal gene would be maintained at about .001 in a very large and freely interbreeding population. Only about one in a million among those born would be homozygous and die. Yet about one individual in every 500 would carry the gene and would be capable of transmitting it. If there are 1,000 such loci capable of supporting lethals at equilibrium frequencies of .001, only about 14 per cent ( $= .999^{2000}$ ) of all individuals will be entirely free from all lethals. The rest will carry one or more which they could transmit. Only about one individual in a thousand among those actually born will show any one of these 1,000 individually rare defects.

These figures will illustrate why lethal or otherwise undesirable genes may be so abundant in a population that inbreeding will be almost sure to uncover them in large numbers and yet, if the population is large and breeding at random, any one of those defects may be seen only rarely. They will also explain why genes against which selection has always been directed since time began still recur in appreciable numbers. Lethals are examples of such genes, although there is always the possibility that a gene now lethal in combination with the present genes of the species once may have been neutral or even advantageous at an earlier stage when the species had other genes.

The actual evidence on the abundance of lethals in farm animals is still scanty. It consists mostly of the considerable number of lethals for which the Mendelian basis has been discovered and reported already and of general observations concerning the effects of inbreeding. There is more evidence concerning natural populations of such organisms as *Drosophila*, although the question of whether the situation is the same among farm animals remains an open one. For example in one study (*Genetics* 26:25) of *Drosophila pseudoobscura* from the Death Valley region in California, over 15 per cent of all third chromosomes in wild flies carried lethals. In another population from Mexico and Guatemala the corresponding figure was 30 per cent. In another California population (*Genetics* 27:373) of 1,292 chromosomes the figure was 14 per cent. Another study (*Biological Symposia* VI:18) of New England, Ohio, and Florida wild populations of *Drosophila melanogaster* showed that 41 to 67 per cent of the second chromosomes carried lethals or semilethals. While such evidence is still meager, yet it seems to indicate that few

individuals are absolutely free from all undesired genes. A small amount of the breeder's freedom to select will be used in combating the generally destructive tendencies of mutation.

#### SELECTION IN FAVOR OF THE HETEROZYGOTE

Selection can never fix the heterozygote. Examples of such heterozygotes are the Blue Andalusian fowl, the Erminette fowl, the cream color of guinea pigs, the yellow color of mice and the roan color of such breeds of cattle as the Shorthorn, the Blue Albion, the "race bleue" (in France), and the blue color inside the ears of Wensleydale sheep, and the Palomino color in horses. Selection of nothing but roans in Shorthorns would lead toward a ratio of 1 red: 2 roans: 1 white. Aside from a few exceptions possibly caused by other modifying genes, it would be possible to produce calves which all were roans by mating whites to reds, but that would be temporary if it were practiced all over the breed, because only roans would then be available for parents of the next generation.

Preference for the heterozygote over both homozygotes means that  $h$  in the general formula for  $\Delta q$  is negative. One of the two homozygotes may be preferred over the other, but if the heterozygote produces more offspring than either homozygote, such selection carries the frequency

of the gene toward a stable equilibrium value,  $\frac{1-h}{1-2h}$ . As a numerical

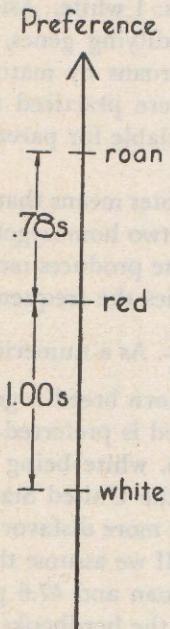
example we may take roan color, which in the Shorthorn breed is generally preferred over both white and red, although red is preferred to white. These preferences vary from region to region, white being in more disfavor in the southern and western parts of the United States than it is in the eastern Cornbelt or in Canada, and in more disfavor in the Argentine and South Africa than it is in Britain. If we assume that Wright's count of 8.6 per cent white, 43.8 per cent roan and 47.6 per cent red among 6,000 animals, equally distributed in the herdbooks of the United States, Canada and Great Britain, represents the equilibrium condition of the breed with respect to color, and if we further accept the monofactorial explanation of the inheritance of these colors, then, by setting  $\Delta q$  equal to zero a numerical expression can be had for the degree to which roan is preferred over red and to which red is preferred over white. That is shown graphically in Figure 14. The example illustrates how a population can cease to change while yet selection for a heterozygote continues and  $q$  has an intermediate value. The example is particularly instructive in showing how it can happen that the most highly preferred color (the roan) is not necessarily the most abundant when equilibrium is reached. This happens because red is preferred to white even more than roan is preferred to red.

Whether the heterozygote is often preferred over both the homozygotes is not yet clear. The cases for which the definite Mendelian basis is known are few but the phenomenon of heterosis, which is widespread and important, is believed by some to rest almost wholly on this. This is likely to be true if (1) each gene has several effects and if (2) the more favorable effect tends to be dominant over the less favorable effect, regardless of the other effects of the same gene. If this is generally true, ideal breeding systems for producing maximum vigor, health, and growth rates, as in animals destined directly for the market, should be

based even more than at present on maintaining purebred but unrelated seedstocks and crossing these for the production of market and work stock.

A preference for the heterozygote may be partly responsible for some lethal genes being as abundant as they are. The yellow mouse, the "creeper" fowl, extreme short leggedness in Dexter cattle, and probably the abnormally thick muscles of "doppelender" cattle, are examples of genes which are lethal when homozygous but have highly prized dominant effects when heterozygous. If a lethal gene has even one dominant effect, favorable enough to give the heterozygote a 1 per cent advantage over the more desirable homozygote, then its equilibrium frequency will be more nearly .01 than .001; nearly 2 per cent of all individuals would carry it, and one such lethal individual would appear among each 10,000 born. If the advantage of the heterozygote over the normal is 5 per cent, the lethal gene will have an equilibrium frequency of  $1/21$ , nearly 9 per cent of all individuals would carry it, and about one individual in 484 would be lethal.

FIG. 14. Scale showing the average degree of preference for roan over red and for red over white among Shorthorn breeders if this preference for the heterozygote is the only thing holding these colors in constant proportions in the breed.



an economic need and reward (an *ecological niche*, the biologist would say) for a few each of tailors, bakers, lawyers, doctors, etc., but if any

#### INTENSITY AND DIRECTION OF SELECTION MAY VARY WITH GENE FREQUENCY

The conditions under which and the purposes for which the breed is kept may be complex enough that there is need for at least a few of each genotype, just as in human societies there is

one of these professions becomes overcrowded, its members are at a competitive disadvantage. If there is no pedigree barrier to the free interbreeding of types in similarly complex animal populations, the result is the same as if the heterozygote were preferred; namely, gene frequency is rather quickly carried to near the value which will furnish the optimum ratio between each of the two homozygotes and the heterozygotes under those conditions. In terms of the general formula this means that the size and even the sign of  $s$  and of  $h$  depend in part on  $q$ . Little is known definitely about whether this situation is rare or widespread and important, either in animal breeding or in nature. Presumably it will be frequent wherever ecological or economic conditions are highly varied.

#### SELECTION AND HOMOZYGOSES

Selection changes homozygosity but little in any one generation. Such change as it does produce may be either to increase or to decrease homozygosity. If mating is random among those selected to be parents,  $2q(1-q)$  of the whole generation out of which the parents are selected will be heterozygous and  $2(q + \Delta q)(1 - q - \Delta q)$  of the next generation will be heterozygous. The change in heterozygosity will be  $2\Delta q(1 - 2q - \Delta q)$  which will depend on both  $\Delta q$  and  $q$  for its size and will be negative when  $q$  is larger than .5, provided that selection does have some effect and hence that  $\Delta q$  is positive. As numerical examples, consider first a case where  $q = .2$  and selection is so effective that  $\Delta q = .03$ , and then a case where  $q = .7$  and selection again is effective enough that  $\Delta q = .03$ . In the first case heterozygosity was .32 in the parental generation and rose to .3542 in their offspring. Here the successful selection decreased homozygosity by .0342. In the second case heterozygosity was .42 in the parental generation and fell to .3942 in the offspring. Here the successful selection increased homozygosity by .0258.

Referring back to Figure 3 it will be noted that  $2q(1 - q)$  changes only a little with changes in  $q$  while  $q$  has values near the middle of its range. It does change rather rapidly with changes in  $q$  when  $q$  is near zero or 1.0, but those are the values of  $q$  at which selection cannot change  $q$  rapidly. Hence, under any but laboratory conditions, where selection might perhaps be extremely intense and directed entirely at the effects of one gene, selection has only tiny effects in any one generation on the homozygosity of the population. This is in marked contrast to the rather powerful effects it can have on the mean of the population when  $q$  is near .5.

Among the nonrandom mating systems, only inbreeding will increase homozygosity much. The amount of help or hindrance which

selection will be to inbreeding in that respect will be so small in any one generation that for practical purposes it can be disregarded, although the accumulated effects may become important if selection is continued over many generations and if the inbreeding is mild.

#### SELECTION AND SEX-LINKAGE

Selection between sex-linked genes is more effective in the heterogametic sex than in the homogametic sex, both because the deceiving effects of dominance are absent and because the heterogametic sex shows the gametic ratio of sex-linked genes directly instead of the square of that ratio as the homogametic sex does. If  $s$  against  $a$ - individuals is the same as  $s$  against  $aa$  individuals,  $\Delta q$  pertaining to sex-linked genes in the heterogametic sex is exactly twice as large as it is for an autosomal gene which shows no dominance at all. In the homogametic sex, selection for sex-linked genes proceeds at exactly the same rate as selection for equally desirable autosomal genes.

#### MANY PAIRS OF GENES—SIMPLEST CONDITIONS

If there are  $n$  pairs of genes equal in effects, without dominance or epistasis, and if the characteristic is not affected by the variations in environment within that population, then the range between the most extreme individuals possible is  $2n$  and the standard deviation is  $\sqrt{2nq(1-q)}$  times the effect of one gene. Thus the range is

$$\sqrt{\frac{2n}{q(1-q)}}$$
 times the standard deviation. Individuals varying from the

mean of a normal population by as much as two times the standard deviation in either direction are unusual (1 in 22), while those varying as much as three times are quite rare (1 in 370), and those varying more than four times scarcely occur at all except among truly enormous populations. Consequently, if the ideal is an extreme type and if more than three or four pairs of genes are involved, individuals homozygous for the desired genes may be so rare that they do not exist at all in the population from which the initial selections must be made. Perfect animals cannot be selected for parents simply because they have not yet been born! Instead the best of those available will be selected and, as this increases the frequency of the genes with the desired effects, each generation will average nearer to the desired goal than the preceding one did, but several generations of selection may be necessary before any ideal individuals appear. The rate of improvement from one generation to the next rises or falls with the standard deviation and therefore is generally maximum when gene frequencies are near .5. Improvement continues until the goal is eventually reached, or selection comes

into equilibrium with opposing mutation rates. The distribution of the population becomes increasingly asymmetrical as the goal is approached. This is the situation usually pictured in generalized discussions of the results of selection. It is represented in Figure 15, where selection begins when the frequencies of all gene pairs are .5, as, for example, in an  $F_2$  generation. Environmental effects and dominance change this situation chiefly in making progress per generation slower, and the changes in variability and symmetry less.

As  $n$  increases,  $\Delta q$  for each gene decreases, other things being equal. But, since the effects of more genes are involved, the rate of change in the population mean, which is proportional to  $2n\Delta q$ , is unaffected. Changes in things like the variability and homozygosity of the population, which depend on the rate of change in  $q$ , are made slower as  $n$  increases. But for the practical breeder the main difference resulting from whether a fixed amount of variability is caused by many genes each with small effects or by a few genes each with large effects is that in the former case the ultimate limit to which the population mean can be carried is much farther off, and he can expect progress per generation to be steadier and not to increase or decrease so much or so soon as in the latter case.

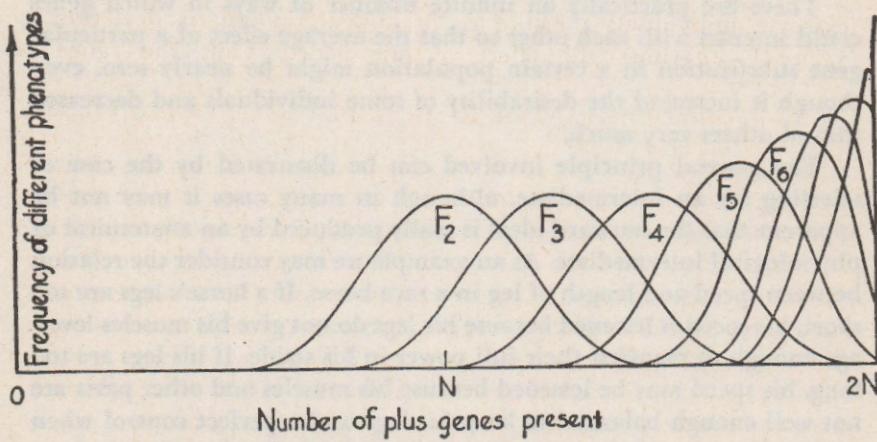


FIG. 15. Distribution of successive generations under intense selection toward an extreme, with few mistakes from dominance or from environmental causes and with no epistasis.

Doubtless some genes have large and some have small effects. Because of that, the variability of the population behaves as if  $n$  were smaller than the actual number of genes but larger than the number of those which have major effects.  $\Delta q$  will be larger for the genes with the larger effects than for genes with minor effects. When selection succeeds

in making  $q$  for the more important genes approach such high values that they no longer contribute much to the genetic variability of the population, the situation becomes more nearly as if many genes each have minor effects.

#### SELECTION FOR EPISTATIC DIFFERENCES

Some genes have one effect in some combinations and another effect in other combinations. In combination with  $Aa$  or  $AA$ ,  $bb$  may have an undesirable effect, and  $aa$  may have an undesirable effect when in combination with  $BB$  or  $Bb$ ; but  $a$  and  $b$  may supplement each other's effects so well that  $aabb$  is as desirable as  $AABB$ . In such a case it would be meaningless to speak of  $A$  or  $B$  as desirable genes. They are desirable when together but undesirable when separate. Selection for net merit is against  $A$  when  $B$  is absent but for  $A$  when  $B$  is present. A partial analogy may be had by considering whether shoes help or hinder the speed of a man running a foot race. If he takes off one shoe, his speed will almost certainly be lowered; but, if he takes off the other one also, his speed will be raised again, perhaps even to a higher level than when he wore both shoes. Whether taking off a shoe makes him faster or slower depends in part on whether the other shoe is on or off!

There are practically an infinite number of ways in which genes could interact with each other so that the average effect of a particular gene substitution in a certain population might be nearly zero, even though it increased the desirability of some individuals and decreased that of others very much.

The general principle involved can be illustrated by the case of selecting for an intermediate, although in many cases it may not be apparent that the outward ideal is really produced by an anatomical or physiological intermediate. As an example we may consider the relation between speed and length of leg in a race horse. If a horse's legs are too short, his speed is lessened because his legs do not give his muscles leverage enough to manifest their full power in his stride. If his legs are too long, his speed may be lessened because his muscles and other parts are not well enough balanced to keep his legs under perfect control when racing. The maximum of speed may come neither with extremely long nor extremely short legs but with legs perfectly balanced with other parts so that the animal is a harmonious whole with all parts co-operating perfectly with each other. The genes which affect the length of leg might be entirely additive in their effects on length of leg, but they will not be additive in their effects on speed.

This simple situation is illustrated in Figure 16. The change from  $a$  to  $A$  may tend to make a horse's legs longer, almost regardless of how

long they already are or of what other genes are present, but it will not consistently make the horse speedier or slower. We may speak of *A* as a gene which lengthens legs. We cannot consistently speak of it as a gene which makes a horse speedier. If substituted for *a* in a short-legged horse, it makes him speedier; but, if the same gene substitution is made in a long-legged horse, it makes him slower. Selection for speed is selection for *A* in short-legged horses and selection against *A* in long-legged horses.

The simplest scheme which will describe in Mendelian terms the consequences of selecting for an epistatic effect is to suppose that a characteristic is affected by two pairs of genes lacking dominance, equal in effect, and combining their effects by addition, but that the intermediate phenotype—the one with two plus genes—is considered the most desirable. If in each pair the gene with the plus effect has a frequency of .5, the nine possible genotypes will be grouped into five phenotypes as follows:

Genotypes:	$1aabb$	$2aaBb$	$1AAbb$	$2AaBB$	$1AABB$
No. of plus genes in each animal	0	1	2	3	4
Ratio of numbers in each phenotype	1	4	6	4	1

Saving for breeding purposes only individuals from the middle phenotype would increase the proportion of that phenotype in the next gen-

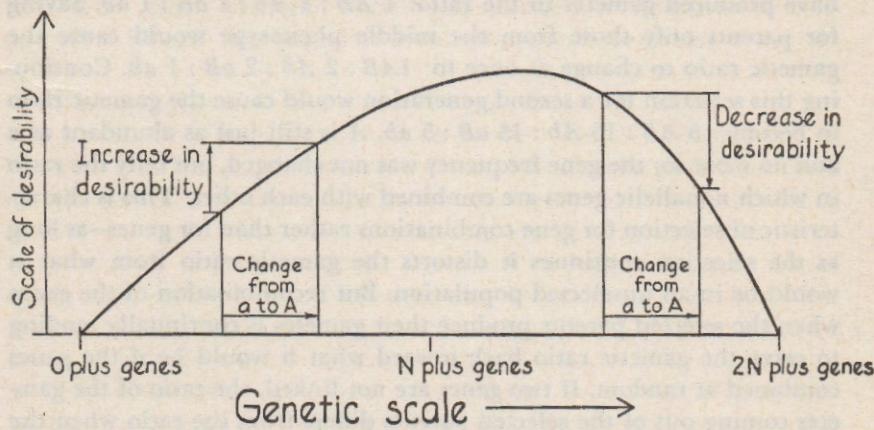


FIG. 16. Illustration of a simple case where the most desirable individuals are intermediates on the genetic scale. Whether the substitution of *A* for *a* increases or decreases the merit of the individual depends on the other genes which are present.

eration from  $37\frac{1}{2}$  to 50 per cent and would reduce the variability of the population. The breeder would appear to be making rapid progress. In the second generation of selection the percentage in the most desir-

able phenotype would increase from 50 to 56 per cent, and the variance which was reduced to 67 per cent of its original value by one generation of selection would be further reduced to 56 per cent of the original. Progress in the second generation is less than in the first. In the third generation of such perfect selection for the intermediate phenotype, the percentage of individuals in that phenotype would rise only from 56 to 57 per cent, and the additional decrease in variance would be only 2 per cent of the original amount. Progress would come nearly to an end with the second or third generation of such selection.

Not only is selection helpless to make much change beyond this point, but continued selection is necessary to hold the gains already made. If selection ceased at the end of the third generation, the percentage of individuals in the desired middle phenotype would fall in the next generation from 57 to 45 per cent and in another generation or two would be practically where it was ( $37\frac{1}{2}$  per cent) before any selecting began.

In this example selection was neither for nor against *A* or *B*; it was for animals which had *any combination* of two of those genes. The result was a change in the gametic ratio because selection eliminated more of the individuals which would produce extreme gametes (*AB* and *ab*) than of those which would produce gametes (*Ab* or *aB*) containing only one of the plus genes. The unselected original population would have produced gametes in the ratio: 1 *AB* : 1 *Ab* : 1 *aB* : 1 *ab*. Saving for parents only those from the middle phenotype would cause the gametic ratio to change at once to: 1*AB* : 2*Ab* : 2*aB* : 1*ab*. Continuing this selection for a second generation would cause the gametic ratio to become: 5*AB* : 13*Ab* : 13*aB* : 5*ab*. *A* is still just as abundant as *a* and no more so; the gene frequency was not changed, but only the ratio in which nonallelic genes are combined with each other. This is characteristic of selection for gene combinations rather than for genes—as long as the selection continues it distorts the gametic ratio from what it would be in an unselected population. But recombination of the genes when the selected parents produce their gametes is continually tending to carry the gametic ratio back toward what it would be if the genes combined at random. If two genes are not linked, the ratio of the gametes coming out of the selected parents differs from the ratio when the genes are combined at random only about half as far as did the ratio of the gametes which united to form those selected parents. With each additional generation, after selection ceases, the remaining difference between the actual gametic ratio and what that ratio would be under random combination tends to be halved. If the two genes are linked the rate of approach toward the random distribution is *c* (instead of one-

half) of the remaining difference each generation,  $c$  being the percentage of recombination.

Selection for epistatic effects is somewhat like building a sand pile on the seashore exposed to each incoming wave. It is easy to build a little pile between waves, but each wave which rolls over it tends to flatten out the pile. When building is stopped, some traces remain after the first wave and perhaps even a few after the second and third, but soon practically all traces of the pile are leveled away. If building continues between waves, the pile can be built a little higher before the second and third waves than it was built before the first wave but soon a size is approached which can just be maintained, the building between waves being just enough to repair the leveling action of the preceding wave.

It should be emphasized that selection for an intermediate is not necessarily selection for hererozygosis. In the example just given, selection was for the  $AAbb$  genotype which is entirely homozygous, just as much as it was for the  $AaBb$  genotype which is entirely heterozygous. Intermediacy and heterozygosis are almost unrelated to each other, provided the characteristic is affected by more than two or three pairs of genes.

The existence of environmental effects and dominance to confuse in the selections, and the usual necessity of saving more than one phenotypic class in order to have parents enough, weaken the intensity of selection for epistatic gene combinations. Probably the Mendelian example just used showed more extreme effects than would often be met in practice, although the multiplicity of kinds of epistatic effects possible throws some doubt on the validity of that conclusion.

Also this example was somewhat artificial in its supposition that the frequencies of  $A$  and of  $B$  were both exactly .5 and would remain at that level. If one had been larger and the other smaller, selection would ultimately have made the whole population homozygous for the gene with the larger initial frequency and for the allele of the gene with the smaller initial frequency. The frequencies are in equilibrium when both are .5, but this equilibrium is essentially unstable. When disturbed by sampling variations, it would tend to depart from equilibrium at an increasing rate. Hence, the population would ultimately become either  $AAbb$  or  $aaBB$ . In more complicated epistatic situations the equilibrium might well be a stable one toward which each gene would tend to return when disturbed.

The general principle which the example illustrates is that, where a genetic intermediate is the goal, selection will carry a population rather quickly to the point where the number of plus genes will *average* nearly what is desired, but some individuals will have more of them and some

will have less. Further selections can do little more than hold down the variation, unless the epistatic equilibrium is an unstable or moving one. If selection ceases, the average number of plus genes will not change but variability will at once increase, and the average merit of the population will decline sharply, most of that decline occurring in the first generation.

For all of the differences caused by the  $Aa$  pair of genes to be epistatic requires that the average effect of changing  $A$  to  $a$  shall be zero; i.e., that the cases in which the possessors of  $A$  have higher reproductive rates than the possessors of  $a$  shall be exactly balanced by the cases in which the possessors of  $A$  have the lower rates. Then the net selection pressure for or against  $A$  (the average  $s$ ) would be zero, and selection would not tend to change the frequency of  $A$  in either direction. However, if selection changes the frequency of other genes which alter the difference between  $AA$  and  $aa$ , the proportion of genotypes in which  $A$  is at an advantage or disadvantage may change. This would then give  $A$  some average effect, and selection for or against it would begin. In short,  $s$  for  $A$  would partly depend on the frequency of genes in other allelic series, as well as on the physiological differences between  $A$  and of  $a$  themselves and on the kinds and frequencies of the different environments in which the population lives.

The general effect of epistatic interactions is to decrease the rate at which selection changes the frequency of a gene, but they may help gene frequency to drift about irregularly to an extent which may sometimes be important, especially in small populations being inbred.

#### AUTOSOMAL LINKAGE AND SELECTION

Autosomal linkage makes new combinations rarer. It is therefore a factor for stability, making it harder to get desired new combinations but easier to hold existing combinations of desirable genes while trying to add other genes to them. Although linkage is a drag on progress, it does not actively tear down any of the breeder's past accomplishments. It can be compared with friction in a machine which requires effort to overcome but helps keep the machine in position wherever it stops and can be useful in such parts of the machine as brakes and governors.

How linkage works can be seen by supposing an extreme case in which there is no crossing over. Then each chromosome would behave as one large gene with many effects, some favorable and some unfavorable. These effects would be distributed more or less randomly along the chromosome, but it would be a remarkable coincidence if two chromosomes of a pair were exactly equal in selective value. In the whole population the chromosomes of each pair would constitute an

indefinitely extended series of multiple alleles. With any general tendency for dominance of the favorable effects, selection would favor the heterozygotes and tend toward an equilibrium at which the population would continue to keep all chromosomes which had any dominant favorable effects at all. But those chromosomes which had only a few favorable effects would be kept at a low frequency. As mutations occurred, the selective values of the chromosomes containing them would alter, and the equilibrium frequencies would shift.

Now if some crossing over takes place, the selective value of each chromosome will alter at a still more rapid rate. Any chromosome which loses more of the desirable genes than it gains by crossing over tends to be reduced to a lower frequency. One which gains more than it loses tends to be increased to a higher frequency.

Crossing over is continually tending to bring each pair of genes into random distribution with every other pair, so that the coupling phases of the double heterozygotes tend to become just as numerous as the repulsion phases. But whenever the genes are not in this equilibrium, the approach toward equilibrium is slower if there is linkage than if the genes were independent. Selection disturbs this randomness of the combinations of the genes with each other, as will be discussed in more detail in the section on selection and variability. The gametes coming from the selected individuals tend to include more of the intermediate combinations and fewer of the extreme combinations than if the same genes were combined entirely at random. If linkage exists this will persist longer and will build up to a wider discrepancy from the random distribution than if the genes are all independent.

If the population is large enough that the inbreeding effect can be ignored, the net result is that linkage will make the offspring of selected parents less variable, and this in turn will prevent the selected and the culled individuals in the next generation from averaging as far apart as they might otherwise. Linkage may constitute some reason for allowing two or three generations of interbreeding following a cross before selecting intensely to combine the desirable characteristics of two different races into one new one. That would give more time for the various genes to cross over so that their coupling and repulsion phases would have more chance to become equally abundant.

In selection for such epistatic effects as when the intermediate is more desirable than either extreme, linkage may play a still more active part in keeping the percentage of desired offspring higher than it would be otherwise.<sup>3</sup>

<sup>3</sup> For details about this see Mather's article in *Jour. of Genetics* 41: 159-93. 1941.

**SELECTION AND THE VARIABILITY OF A POPULATION**

Mass selection of the parents has little effect on the variation among the offspring, although of course the variation remaining among the selected parents themselves will be distinctly less than was in the population from which they were chosen.

Eliminating 10 per cent of the very poorest from a normal distribution will decrease the standard deviation of the remainder by 16 per cent, eliminating 20 per cent will decrease it 24 per cent, and eliminating 50 per cent will decrease it 40 per cent. Thus even a small amount of culling can make rather striking effects on the uniformity of the group of survivors. Probably this is the main cause of the rather widely held opinion that selection is an effective way of increasing uniformity. In most herds some selection is practiced, and the visitors see only the selected survivors of at least a little culling. When they do see a herd where, through the owner's carelessness or financial difficulties or other reasons, no selection is being practiced they see for the first time that rather rare sight—an entirely unselected population. They are likely to compare such a herd with the other herds they know, which are selected populations, rather than with the unselected offspring of selected parents. The latter is what should be done to find how selection of the parents really affects the uniformity of their offspring.

Sometimes in thinking about selection and variability we are contrasting two herds which are the products of selection in different directions. Since selection can shift the mean of a population considerably, even when it does not change the variability within that population, two herds started from the same foundation stock but selected toward different goals for two or three generations may differ rather sharply in their means. If so, it may appear by contrast that selection has made each herd uniform, whereas really it only increased the differences between herds without making much change in the variation within herds.

Sometimes when we think of selection and uniformity we are comparing the offspring of one selected sire with a whole breed or other large population. The offspring of one sire have some extra uniformity because they are half brothers and hence get half of their inheritance as samples from the very same genotype. By contrast individuals whose parents were not the same but merely were selected because they were much alike phenotypically may get widely different kinds of inheritance, since those parents will generally be less alike in breeding value than they are phenotypically.

Often when we think of selection and variability we are comparing the variation within one herd with variation within the whole breed.

Usually each herd has some environmental conditions which are different from those of other herds but tend to affect alike all the members of the same herd. These effects of common environment may often be enough to make each herd distinctly less variable phenotypically than a population composed of a fair sample from all herds of the breed.

All these things may be mistaken for the effects of selection. One or more of them often are. It is not surprising that many persons, without having seen experiments or herds where the actual contrast is only that the parents were a selected group in the one case and a random sample of their population in the other, should have inferred that selection would increase uniformity distinctly. This opinion is still widespread, notwithstanding the fact that actual experiments have contradicted it and that these have been published. As long ago as 1907 E. D. Davenport wrote<sup>4</sup> "We often speak of 'fixing' the type by selection, meaning by that the reduction of variability. All recent studies, however, go to show that we do not greatly reduce variability by selection, however much we alter the type." and "The principal function of selection, therefore, is to *alter the type, not to reduce variability, . . .*"

Selection of the parents alters the variability of the next generation in two ways from what it would have been if the parents had been unselected. First, it changes gene frequency, and that automatically has some effect on variation. Second, the gametes from selected parents contain a somewhat larger proportion with intermediate combinations of desirable and undesirable genes than would be the case if the very same genes with the same actual frequencies were combined entirely at random.

The first effect is almost always very slight and may either increase or decrease variability. It has already been discussed in connection with selection and homozygosis (page 131). Variance goes up and down in proportion to a term which always has  $2q(1-q)$  as a factor, the other factors depending on whether there is dominance and upon the nature of the mating system, if that is not random. Successful selection will increase variability if  $q$  is small at the start but will decrease variability after  $q$  is much larger than .5. But the change is very small when  $q$  has values in its middle range, and  $\Delta q$  will be small when  $q$  is near zero or 1.0. Therefore, this effect of selection in changing variability may be plus instead of minus but is exceedingly small in any one generation.

The other effect of selection in producing an excess of intermediate gametes, as compared with what there would be if the same genes were combined with each other entirely at random, is the same process already discussed (pages 136 and 137) in connection with selection for a

<sup>4</sup> Pages 534 and 537 in *Prin. of Breeding*. Ginn & Company.

genetic intermediate except that here it is usually less extreme, since individuals are discarded from only one end of the curve and not from both. As an extreme Mendelian example, consider again the case on page 135 and suppose that only the two phenotypes on the extreme right were saved for breeding. The frequency of *A* and of *B* among the gametes they produce would both be .8. The actual array of gametes from those parents as contrasted with what would occur if the same genes were combined strictly at random would be as follows:

Gametes	<i>ab</i>	<i>Ab</i>	<i>aB</i>	<i>AB</i>
Actual frequency .....	.00	.20	.20	.60
Frequency if random .....	.04	.16	.16	.64

Although the example is extreme and selection is assumed to be without mistake, the departures from the random distribution are small.

This excess of intermediate gametes tends to correct itself in subsequent segregations and recombinations of the genes. Whatever difference there is between the actual gametic array and what it would be if the genes were combined at random tends to be halved with each succeeding generation, so far as concerns any two unlinked genes, and to lose *c* of its amount in each generation if the genes are linked, *c* being the percentage of recombination.

This effect of selection through narrowing the gametic array is generally slight. It depends for its size on the intensity of selection as well as on the square of the heritability of the characteristic being selected. As a numerical example, if only the best half of a hitherto unselected and random breeding population is saved for breeding, the standard deviation of the offspring will be 17 per cent less than the standard deviation of the population from which their parents were taken if the characteristic being selected is not affected at all by environment, dominance, or epistasis. But if heritability is 50 per cent, the reduction in standard deviation will be only 4 per cent. If heritability is as low as 30 per cent the reduction in standard deviation will be less than 2 per cent. If the culling could be so extraordinarily intense that only the best 10 per cent were saved for parents, the corresponding reductions in standard deviations for those three levels of heritability would be 24 per cent, 5 per cent, and 2 per cent, respectively. For characteristics with heritabilities much under 50 per cent, the reduction in the variability of the offspring caused by selection of the parents is thus only a tiny amount.

The reduction in variability proceeds only a little farther in following generations if selection is continued. As in selection for epistatic effects, most of what can be done is done in the first generation or two of selection. Further selection only does enough to cancel the tendency

for the genes to recombine at each segregation to produce a gametic array which would be more nearly random. When selection ceases, this slight reduction in variability caused by selection having made the gametic array nonrandom disappears quickly, most of it going in the first generation produced from unselected parents.

#### SUMMARY

The primary effect of selection is to change gene frequency. Its outward results are consequences of that. Conditions which modify the rate at which selection changes populations, but do not change the ultimate goal, are:

1. The proportion needed for replacements may be so large that not all of the undesired homozygotes can be discarded at first.
2. Selection must be between individuals, which are pairs of genes, rather than gene by gene.
3. Environment may duplicate or hide the effects of genes, and dominance may cause two or more genotypes to be indistinguishable, so that the breeder makes mistakes which cause his selection to be less intense than it would be if he knew the genotypes perfectly.
4. The amount of selection that can be practiced depends in part on the amount of genetic variability which is present, that is, upon the gene frequency, and on the mating system if that was not random.

Conditions which modify the ultimate goal which selection can attain, as well as the rate at which that is approached are:

5. Selection becomes progressively feebler at eliminating the undesired genes as those become rare. Hence, as an undesirable gene becomes nearly extinct from a population, the power of selection to make it still rarer comes into equilibrium with opposing mutation rates, even when the latter are very low. Generally this equilibrium frequency of the undesired gene is so low that it is not of much importance in practical breeding, but a tiny fraction of the breeder's efforts is required for combating mutation.

6. When the heterozygote is more desirable than either homozygote, selection ceases to change gene frequency while yet the gene frequency has an intermediate value which may be rather far from either 1.0 or zero.

7. If economic or ecological conditions provide a useful place in the population for at least a few individuals of each of the homozygous types, and if the population is freely interbreeding, this has the same effect as if the heterozygote were preferred. Progress in changing the population by selection comes to a halt when gene frequency reaches whatever value will most nearly provide that proportion.

8. Epistatic effects tend to lower the rate at which selection changes gene frequency because selection for a gene in some combinations tends to be balanced by selection against the same gene in other combinations. If all the variation caused by a certain gene is epistatic, the net selection pressure for or against that gene is zero. Selection then merely tends to keep the frequency of the gene at this equilibrium point.

In any one generation selection has very little effect on homozygosity or variability of the population.

The number of genes responsible for a given amount of genetic variability does not affect the amount of progress which selection can make in the next generation, but if the gene number is large the rate of progress will not change so much or so soon, and the ultimate limits to which selection can change the population will not be so near as if the genes which cause this same amount of genetic variation are few.

Autosomal linkage lessens the effectiveness of selection slightly by making the array of gametes from selected parents less widely diverse than it would be if the genes were independent.

Selection for sex-linked genes is roughly twice as effective in the heterogametic sex as is selection for autosomal genes. In the homogametic sex, selection is equally effective for sex-linked and for autosomal genes.

#### REFERENCES

The classic work on selection is Darwin's *Origin of Species*, which, however, was written before the mechanism of inheritance was discovered. It seems to have been wrong chiefly in assuming that inheritance was blending in nature and, therefore, that hereditary variations must be seized by selection almost at once after they occur, else they would be "swamped" in the subsequent matings and lost. Hence, also, it assigned too much importance to mutation. Also, the qualitative distinction between hereditary and nonhereditary variations was not entirely clear. Sexual selection probably was overemphasized. R. A. Fisher's *The Genetical Theory of Natural Selection* shows how Darwin's conclusions are modified or extended by modern genetics. Sewall Wright's "Evolution in Mendelian Populations" (*Genetics*, 16:97-159, March 1931) and also his "The Roles of Mutation, Inbreeding, Crossbreeding, and Selection in Evolution" (*Proc. Sixth International Congress of Genetics*, 1:356-366) treat extensively of the interplay of selection, mutation rates, and inbreeding systems. It is his notation which is mostly followed here. Wright's articles in the *Journal of Genetics*, 30:243-266, are at this writing the most comprehensive study yet published of the genetic consequences of selection for epistatic effects.