CHAPTER 17

POPULATION GENETICS AND EVOLUTION

The mechanism of heredity and variation is basic to the study of evolution and was therefore a major concern of Darwin and many of his followers, including Galton and Weismann. It was an interest in evolution that led several of the earlier Mendelians, such as de Vries and Bateson, to the study of heredity. But, with the discovery of Mendel's work in 1900, the development of the new methods caused a temporary lack of interest in the evolutionary implications. As Bateson put it in 1909:

> It is as directly contributing to the advancement of pure physiological science that genetics can present the strongest claim. We have an eye always on the evolution-problem. We know that the facts we are collecting will help in its solution; but for a period we shall perhaps do well to direct our search more especially to the immediate problems of genetic physiology . . . willing to postpone the application of the results to wider problems as a task more suited to a maturer age.

Evolution is concerned with changes in populations, rather than in individuals, and what was needed was an analysis of the effects of the Mendelian scheme on populations of interbreeding individuals. The beginning of this analysis is a paper by Yule (1902) in which he pointed out that, if the members of an F_2 population, segregating for a single pair of genes (*A* and *a*), interbreed at random, the three types of individuals (*AA*, *Aa*, *aa*) will be represented in the same proportions in the following generations. He also raised the question: What will happen if all the *aa* individuals are removed? His analysis here was in error, but it was corrected by Castle in 1903. Castle pointed out also that, if such selection ceases in any generation, the newly established proportions will then be stable.

Here was the essence of the basic formula of population genetics, though it was derived by a longhand method and was not stated in simple algebraic form. The result, without selection, was also derived by Pearson (1904), that is, for the case where $p = q$, where p = the frequency of A genes, q = the frequency of a, and $p + q = 1$.

The generalization that the stable frequency of genotypes is $p^2AA : 2pqAa : q^2aa$ was made by Hardy and by Weinberg, independently, in 1908. Both knew of Pearson's result. To Hardy, who was a mathematician, the generalized result seemed so self-evident that he commented: ". . . I should have expected the very simple point which I wish to make to have been familiar to biologists." That it was not familiar is shown by the fact that it had been seriously suggested that dominant genes would automatically increase in frequency in mixed populations.

The Hardy-Weinberg formula is strictly valid only if several conditions are fulfilled:

1. The population must be large enough so that sampling errors can be disregarded. As Hardy pointed out, any chance deviations in the values of p and q will be as "stable" in succeeding generations as were those of the preceding one.
2. There must be no mutation, since change of A to a or of a to A will alter the values of p and q.
3. There must be no selective mating.
4. There must be no selection, that is, A and a must have no differential effect on the reproductive capacity of individuals bearing them.

(In the requirements 2 and 4, the wording given is not strictly correct, since it is possible that balanced mutation and selection rates may exist—in which case there will be no net changes in the frequencies of A and a.)

These are rather stringent requirements, and it may be doubted if they are ever fully met; nevertheless, they are often approximated closely enough to make the formula useful in analysis of populations.

The further developments in this field have depended on the algebraic analysis of the effects of deviations from the four requirements. This development dates from Haldane's (1924 and later) analysis of the effects of selection. He determined the number of generations required to alter gene frequencies, as related to the intensity of selection. This was worked out for dominant and for recessive genes, both in haploid and in diploid organisms, and for sex-linked as well as for autosomal genes.

This analysis has been followed by detailed studies on the algebraic

consequences of variations in each of the four requirements by Haldane, Fisher, Wright, and others. Perhaps the most useful general summary is that of Li (1955).

One of the early developments was a theory of the origin of dominance (Fisher, 1928). Fisher suggested that mutant genes are inherently neither strictly dominant nor strictly recessive but produce more or less intermediate heterozygotes. There are, however, numerous modifying genes that affect their dominance. Since most mutant genes have unfavorable effects on their bearers, an individual heterozygous for a mutant gene will leave more offspring if the modifiers it carries happen to make the gene more nearly recessive. This effect will be slight, since an unfavorable semidominant mutant will not persist long in the population. However, since mutants continually recur with a low frequency, the effects will be cumulative over very long periods. This hypothesis has played a large part in discussions of population genetics but has been criticized by Wright and others on several grounds. The postulated effect of modifiers is a second-order effect, and it seems likely that their frequencies in the population will usually be determined by more direct effects. Also, an alternative interpretation is that a favorable gene will undergo selection—probably largely among alleles of the gene itself—for a "factor of safety," so that it will be capable of producing an excess of its useful product in times of stress; and such an excess may be supposed to result directly in dominance.

The first attempt to assemble a coherent general account of the algebraic analysis of Mendelian population behavior was Fisher's book in 1930. Widely read and discussed, it certainly strongly influenced further developments. One of the elements it minimized was that of the effects of population size. This was analyzed by Wright in several papers, first summarized in 1932.

Wright pointed out that in small populations there is a possibility of chance shifts in gene frequencies which are not controlled by selection and that this may lead to the production of combinations of genes—sometimes favorable—that would have almost no chance of being produced in large populations. He suggested that the most favorable condition for rapid evolution is that of a large population that is split (geographically or otherwise) into a series of relatively small subpopulations, with gene flow between these possible, but strongly restricted. Under these conditions, what has come to be called "random drift" may produce more favorable combinations of genes in some subpopulations,

and these will gradually spread to neighboring populations. More often, probably, the result of random drift will be a less favorable combination of genes; when this happens, the subpopulation will diminish and probably be replaced by migrants from neighboring areas. Selection still remains the determining element, but emphasis is placed on selection among subpopulations rather than solely among individuals.

These analyses are largely theoretical, based on laboratory experiments—not on the actual properties of naturally occurring wild populations. It has turned out, not unexpectedly, that such populations are difficult to study. The situations encountered are so complex that it becomes difficult to evaluate separate variables. Further, when a quantitative analysis is made, it is specific for the population studied and cannot safely be applied to other populations, even of the same species. In other words, generalizations are difficult and dangerous. One result often obtained is that new kinds of problems have been suggested for further exact algebraic analyses.

In spite of these difficulties, much progress has been made. The first attempt to coordinate the existing quantitative data on natural populations and interpret them in terms of algebraic studies, was the book by Dobzhansky (1937).

It had, of course, long been evident that inherited diversities do occur in natural populations. Such examples as separate sexes, or heterostyly in plants, are obviously special cases. But the occurrence of "sports" and of inherited slight differences in such things as size, cold-resistance and color, among other things, was also familiar. It was not so clear that there is a store of recessive genes suitable for exact study by Mendelian methods. Examples of these had also been found by many observers, but the first attempts at a quantitative determination of their frequency in wild populations seem to have been made by a series of Russian investigators, using Drosophila. This work was initiated by Chetverikov (sometimes transliterated as Tschetwerikoff) in 1926 and culminated in the work of Dubinin and collaborators (1934, 1936), who studied a large series of wild populations of *Drosophila melanogaster* collected in the Caucasus. They found that up to 16 percent of the second chromosomes in these populations contained recessive lethals. It had not been supposed that wild populations contained such high frequencies of unfavorable recessives, but further work by Timoféeff-Ressovsky, Sturtevant, Dobzhansky and co-workers, and others has confirmed the result for several species and also has, in some cases, shown even higher frequencies of lethals and other unfavorable recessives.

Out of this work grew the realization that natural populations are full

of hidden genetic variability, most of it potentially unfavorable. The study of this "genetic load" is now being actively prosecuted, in part because of its importance in the practical breeding of animals and plants, and because of the increase in frequency of mutations in human populations that must be supposed to have followed increases in irradiation caused by medical and dental uses of X-rays and by exposure to radioactive fallout from atomic bombs.

Another approach to the study of the relation between genetics and evolution is through the use of species hybrids. This, as pointed out in Chapter 1, has a long history; but only with the development of Mendelian and cytological methods did it begin to yield really helpful results.

It was soon apparent that different species usually differ in many pairs of genes and accordingly give numerous recombination types in F_2; Mendel himself indicated this in his discussion of Gärtner's work. With the development of the multiple-gene interpretation of quantitative inheritance (Chapter 9), it became possible to give a more precise interpretation of the results.

One of the awkward circumstances about discussions in this field is the ambiguity in the use of the word *species*. There is no generally accepted definition, but most discussions have centered on the extent to which cross-sterility between populations is to be taken into account in deciding whether or not two groups are to be considered as separate species. The nature and origin of interspecific sterility has been recognized since Darwin as one of the major problems of evolution.

The genetic study of the nature of interspecific sterility is difficult. If sterility is complete, this very fact makes it impossible to study by conventional genetic methods; if it is partial, there usually arises a possibility (or certainty in some cases) that it involves so great a distortion of segregation and of the relative viabilities of the various recombination products as to render analysis difficult or impossible. When these complications are absent or unimportant in any given case it becomes questionable whether the data obtained are relevant to the general problem.

The outcome of crosses between distinct species varies widely from one case to the next. The eggs and sperm may never come together (through lack of mating or failure of pollen tube growth); they may fail to fuse even if they do come together; if fertilization is effected, some or all of the chromosomes derived from the sperm may be eliminated at cleavage in the foreign cytoplasm (Baltzer, 1909); cleavage and mitosis may be normal, with the development of the go-

nads of the hybrid blocked at almost any point—most often just before meiosis; meiosis may be abnormal or essentially normal; the gametes may be normal in structure and be functional but may give rise to some or many abnormal or sterile individuals in the next generation. In other words, there are many different kinds of mechanisms that prevent or hinder interspecific exchange of genes. The problem is, how do these arise?

There is probably no single general answer to this question, but there is an answer applicable to many cases, namely, polyploidy.

The study of polyploidy may be dated from the work of Boveri and others in the 1880's and 1890's on the two races of *Ascaris megalocephala*—univalens, with one pair of chromosomes in the germ line, and bivalens, with two pairs. There is some question about this being a simple case of polyploidy. The first unambiguous cases seem to be those reported by the Marchals (1906, 1907) in mosses (described in Chapter 13) and since studied in great detail by F. von Wettstein, and in Oenothera by Lutz, also in 1907 (see Chapter 10). Other more or less probable examples followed (Strasburger, Tischler, and others) together with cytological studies of meiosis in the triploids produced by crosses between diploids and tetraploids. The results were confusing and contradictory, until the analysis by Winge (1917), which began to clarify the situation.

Winge made a detailed study of the available data on chromosome numbers in plants and found that in many groups there was a basic number, with various multiples of this number represented in different species. He pointed out that, in a hybrid between two species, it sometimes happens that some (or all) chromosomes are sufficiently different so that they do not pair at meiosis, so that the resulting gametes do not all contain a single complete set of chromosomes, and partial, or essentially complete, sterility results. Now, if the chromosomes of such a hybrid are doubled (by chromosome division without an accompanying cell division), each chromosome will now have an exact mate, and meiosis can be expected to be normal—with a restoration of fertility, as had already been shown by Federley (1913). This process was demonstrated by Clausen and Goodspeed (1925) and Clausen (1928), in Nicotiana. *N. tabacum* (24 pairs of chromosomes) was crossed to *N. glutinosa* (12 pairs). The hybrid, with 36 chromosomes, was sterile and at meiosis showed only a few paired chromosomes. One hybrid individual was fertile, however, and was found to be a tetraploid, with 72 chromosomes that formed 36 bivalents at

meiosis; this plant was fully fertile and bred true to type.

The terminology which was suggested by Kihara and Ono (1926) has been generally accepted and has helped to clarify the relations. They suggested that:

> Under polyploidy we must distinguish two phenomena, namely, autopolyploidy and allopolyploidy. Autopolyploidy signifies the doubling of the same chromosome set; allopolyploidy the multiplication of different chromosome sets brought together by hybridization.

This distinction has proved to be convenient and useful—though it is not of an all-or-none type, since intermediate conditions occur. Allopolyploidy has been found to be widespread in the higher plants, but is rare in animals. Muller (1925) suggested that this is because it leads to difficulties in sex determination in species with separate sexes. It now seems more probable that the difficulty lies usually with the crossing of new tetraploids to diploids and the production of relatively sterile triploids; self-fertilizing hermaphrodites can avoid this difficulty.

The relative (or complete) sterility of the triploids is, however, of importance in that it leads to an immediate effective sexual isolation of an allopolyploid from both its parental forms. In fact, this sterility often operates to make difficult the production of any backcross offspring at all (Karpechenko and others), presumably because of an interaction between the tissues of the style and of the pollen tube, since it is known that, in some autopolyploid series, haploid pollen functions best in diploid styles, and diploid pollen in tetraploid styles.

Polyploidy, then, does form one method of bringing about interspecific sterility, but even in the higher plants where it is relatively frequent, it must be considered a rather unusual cause that has little bearing on the general question.

It should be added that the study of polyploidy has been of importance in other directions in genetics. Its use by Bridges and others in the analysis of sex determination has been discussed in Chapter 13, and it has also contributed largely to our understanding of chromosome mechanics, as developed by many authors. There have also been extensive applications in the breeding of cultivated plants. One of the important events in this field was the discovery (Blakeslee and A. G. Avery, 1937) that doubling of the chromosome set in plants may be induced by the use of colchicine.

Winge's interpretation of polyploidy grew out of a comparison of the

chromosomes of related species. Another type of conclusion was based on such comparative studies by Metz (1914 and later) with species of Drosophila and by Robertson (1916) with a series of genera and species of grasshoppers. In both of these groups there are rod-shaped chromosomes, and also (in other species) **V**-shaped ones. Both authors found that, if each **V** was counted as two rods, the haploid number of elements was constant within a given group: 5 in Drosophila, if the small dot (often difficult to see) is neglected; 7 in the grouse-locusts, and 12 in ordinary grasshoppers. The conclusion was drawn that, in general, the elements maintain their individuality (to a large extent at least) within such groups, and are separated, or united, in various ways in different species.

This conclusion was questioned by R. C. Lancefield and Metz (1921) as a result of studies on *Drosophila willistoni.* This species, like *D. melanogaster,* has a pair of rods and two pairs of **V**'s, but in melanogaster the rod is the X chromosome, whereas in willistoni, Lancefield and Metz showed that one of the **V**'s is the X.

This anomalous result was explained later as a result of studies on the mutant genes found in various species of Drosophila. Such studies, by Metz, Sturtevant, D. E. Lancefield, Weinstein, Chino, Moriwaki, and others, showed that mutants with phenotypes closely resembling those of melanogaster could often be found in other species, and in the case of simulans, which can be crossed to melanogaster (although the hybrids are all completely sterile), it was possible to show that many of these resemblances are in fact due to mutations of the same wild-type genes (Sturtevant, 1921). This conclusion rests on more indirect evidence for the other species, but, as more examples accumulated, one rule appeared: mutants in other species that closely resemble sex-linked mutants in melanogaster are also sex-linked—though the reverse relation does not hold so consistently. In retrospect, it is obvious that most of the clear exceptions to the reverse rule concerned mutants in *D. willistoni* and *D. pseudoobscura* that were sex-linked in those species but resembled autosomal mutants in melanogaster; in both of these species the X was known to be a **V**.

The obvious conclusion was implied by D. E. Lancefield (1922) but was not elaborated or made more specific until Crew and Lamy (1935) carried out a detailed comparison of the known mutants of pseudoobscura with the apparently similar types of melanogaster. This paper was not very critical and used terminology that made it difficult to understand, but the conclusions were fully confirmed and extended by the more substantial accounts of Donald (1936) and of Sturtevant and Tan

(1937). These results confirmed the conclusion of Metz and Robertson: the six elements of pseudoobscura do in fact correspond to those of melanogaster. The X of pseudoobscura is **V**-shaped, and one arm contains the material of the melanogaster X, the other that of the left limb of the melanogaster III. The remaining four elements of melanogaster are all intact but are separate. That is, both **V**'s of melanogaster have their two arms separated, and one of these (III L) is now the right arm of the X.

This type of comparison has been extended to several other species of Drosophila that have haploid chromosome numbers from 3 to 6 (summary and analysis by Sturtevant and Novitski, 1941). It appears that the 6 elements have retained their composition, with relatively few exceptions, but that within each element the sequences of loci are little more alike than would result from chance alone. In other words, inversions within an element have been frequent; translocations, or inversions including centromeres, have been rare, in the evolution of the genus, as they are within existing species. There are a few examples of persistent associations of closely linked genes, which may be due to chance or to the existence of favorable position effects, but such persistent sections are not common.

The studies just discussed lead to the conclusion that there is a long-time stability in the genetic basis of particular characters, but such a stability has often been questioned. Perhaps the most extreme formulation of this point of view is that by Harland (1936):

> The genes, as a manifestation of which the character develops, must be continually changing . . . we are able to see how organs such as the eye, which are common to all vertebrate animals, preserve their essential similarity in structure or function, though the genes responsible for the organ must have become wholly altered during the evolutionary process, since there is now no reason to suppose that homologous organs have anything genetically in common.

This conclusion was based on solid facts derived from extensive species crosses in the genus Gossypium (cotton). I have elsewhere (Sturtevant, 1948) given my reasons for an alternative interpretation of these facts, based on the polyploid nature of cotton. The more recent comparative biochemical data also favor the idea of the great stability of genetic systems, since they show essential identity of some of the gene-controlled basic biochemical pathways in bacteria, fungi, and vertebrates.

It is true, however, that in many, probably most, loci there exist series of isoalleles (Stern and Schaeffer, 1943), which carry on the function

characteristic of the locus, but with different efficiencies and different temperature characteristics and reactions to the presence of gene differences in other loci. These often lead to no phenotypic differences under normal conditions, and can only be studied by special methods. It seems probable that the efficiencies of those concerned with any given developmental system need to be properly adjusted among themselves to give a harmonious system, as postulated by Goldschmidt for "strong" and "weak" races of Lymantria (Chapter 13). But other equally effective systems are also possible, and may come to exist in related species. Disharmonies will then arise on species crossing, and there are examples that suggest that this is often the case (Sturtevant, 1948).