252

Croonian Lecture

Population genetics

By Sir Ronald Fisher, F.R.S.

(*Delivered* 11 *June* 1953—*Received* 12 *June* 1953)

1. The classification of genetics

Genetics, the study of the hereditary mechanism, and of the rules by which heritable qualities are transmitted from one generation to the next, is a science sufficiently new for its subdivisions and their mutual relationships to be ill-defined, or, at least, in process of finding their definitions. Consequently, of the many adjectives which one hears and sees applied to the word 'genetics' some are meaningful, while many others imply some distinction scarcely worth making. The wide applicability of Mendel's principles invites a classification based on the subject-matter, such as human genetics, *Drosophila* genetics, mouse genetics, plant genetics, and so on, but no distinction of importance can be made in this way. In recent years indeed, organisms much more different than the higher animals and green plants have been shown to exhibit closely analogous genetic phenomena; the list now includes Protozoa, fungi, bacteria and viruses, and to these obscure and difficult forms some of the finest of modern researches have been devoted (Sonnerborn 1947; Lederberg 1951). The very existence of sexual reproduction in many of these groups was only discovered, and can still only be demonstrated, by genetic methods.

With such an enormous range of diverse forms, though all conforming to the same principles of inheritance first glimpsed by Gregor Mendel, the technical methods of culture differ greatly according to the breeding system current—self-fertilizing hermaphrodites, self-sterile hermaphrodites, animals and plants with separate sexes, ephemeral or long-living, capable perhaps of vegetative or of apomictic reproduction. In the microfungi heterocaryosis, exhibiting fusion of the cytoplasm without using the recombination mechanism of the nucleus, forms a branch of genetics which has been greatly forwarded by Pontecorvo at Glasgow, and by Mather at Birmingham.

2. Biometry and Mendelism

It is not, however, on such distinctions that the great branches into which modern genetics is developing can be defined. The first major distinction that I should make is between the genetics of quantitative characters, dependent on the cumulative

action of numerous factors, and the genetics of distinct and recognizable *loci* in the germplasm. These differ in the technical equipment required of the geneticist, and in their fields of application. In the evolution of species by natural selection, certainly, it is the quantitative characters that are of principal importance; this is equally true in the selective improvement of domesticated animals, and cultivated plants; or, again, in any serious effort to control the mental and physical deterioration of the human race. The means of study are of course through experimental breeding, but *faute de mieux* what can be observed are biometrical quantities, means, variances and covariances, and third degree statistics, rather than the frequencies of manifestation of individual items characteristic of Mendelian work proper, and by means of which alone we can hope to build up cumulative knowledge of the structure and content of the germplasm, and of the fundamental laws of its behaviour. In making this distinction the last thing I should wish to imply is that the study of singly recognizable gene substitutions should be neglected at centres for animal and plant improvement; on the contrary, far too little attention has been given to this aspect of genetic knowledge in our larger farm animals, although only this can supply the basis of any improvement in our knowledge of the genetic situations presented by them. By contrast, the study of human genetics has been completely revolutionized in recent years by the genetical study, not of quantitative characters, but of the blood groups, yet it is not in respect of blood groups that human deterioration is to be feared. The fact is that we must rid ourselves of the last traces of that antagonism which so sorely impeded the progress both of biometry and of Mendelism fifty years ago, and recognize that both types of technical training are needed if serious progress is to be made in any real genetical problem, and that university education should be reformed so as to allow both types of study to be pursued by the same students, without allowing or, still less, inculcating, the presumption that proficiency in either one is any substitute for proficiency in the other.

3. Levels of organization

A second principle of classification has been proposed by Darlington & Mather (1949), namely, that corresponding with the three levels of organization represented by the cell, the organism and the population, there have developed cytogenetics, individual genetics and population genetics. The principle of classification is an important one, but the material to be classified will not easily divide itself into these tidy packages. At the cytogenetic level we do indeed have, largely owing to Darlington himself, an orderly scheme for the interpretation of the normal cytological phases of the reproductive process, and of the commonly occurring cytological abnormalities; at the two higher levels of organization, however, no such simple contrast as that between individual and population is adequate to characterize the newer genetical disciplines which have been developing.

4. Polysomic inheritance

We have first of all the inheritance of those polyploid forms in which not only is the germinal material doubled or tripled in relation to the cell and the organism,

but in which the increase has yielded a multiplicity of *like* elements capable of pairing and interchange in the normal reproductive cycle. The laws of inheritance in such cases, though derivable from Mendel's principles, are immensely more complex than in diploid forms, or in those polyploids which behave as diploids, amphidiploids, etc., all of which may be comprehended by the single word *disomic*, in contra-distinction to those more characteristic polyploid forms which we may distinguish as *polysomic*. The increase of complexity is such that the development of adequate ideas, and of an adequate vocabulary, will necessarily take some time (Fisher 1947, 1949*a*, 1950). The ideas, particularly, of combinatorial mathematics, partitions, multipartitions, permutative operations, groups, etc., have been needed, and have been freely employed in the theoretical exploration, which has accompanied the experimental invasion of this territory. It must be emphasized that not only wild plants like the purple loosestrife, but also some of the most important of the world's crop plants, such as potatoes and the black medick, lucerne or alfalfa, illustrate inheritance of this scarcely investigated kind.

5. Inbreeding

Secondly, I should mention the theory of inbreeding (Fisher 1949*b*), and the study of such lineages as are produced by a predetermined mating system. Although such a lineage may be spoken of as a population, it is only a population in a very peculiar and exceptional sense of the word. I do not think it would be helpful to consider the study of inbreeding as a part of population genetics, though undoubtedly the lineage is not an individual, but a concatenation of individuals. The genetic properties of an inbreeding system are very simply expressed in terms of matrix algebra; the breeding system defines the generation matrix, a mathematical operator transforming each generation into the next. Most of the ideas of population genetics specifying the vital statistics, and the selective intensities arising from them are either quite inapplicable to an inbred line, or are applicable only rather artificially and fruitlessly. I doubt if any modern worker on inbreeding is tempted to overlook these differences.

6. Selective improvement

In my third class of modern genetical developments, namely, the study of programmes of selective improvement in quantitative characters, a study which has been greatly forwarded in this country by the work of Mather (1949) at Birmingham, and in the United States by that of Lerner (1950) at Berkeley, California, there *has* been at least some *verbal* confusion. Selection and inbreeding are the two means by which genetical material may be acted on and modified in a calculable way without our being able to recognize the factors by which the modification is taking place. A selective improvement programme, like a system of inbreeding, produces an artificial population with properties which it is the intention of the procedure constantly to modify. Although many of the ideas of population genetics, such as the genotypic variance and the divisions into which this quantity can be analyzed, are immediately helpful to the selectionist's work, the corresponding quantities, being to some extent human artifacts, cannot have

to the naturalistic observer the same scientific significance as quantities bearing the same name have in the study of real natural populations. Mather was, in my opinion, more judicious in entitling his book on this subject *Biometrical genetics* than was Lerner in using the very wide term *Population genetics* for a topic which some might think scarcely to lie in the field of population genetics at all, and which, however this may be, covers only an exceedingly specialized, though practically very important, section of this field. In respects other than the use of this word in the title, Lerner's book is undoubtedly a contribution to the subject of sterling value.

These three modern developments: polysomic inheritance, inbreeding, and biometrical genetics in the sense of Mather, which I have mentioned primarily to distinguish and make clear the place and scope of population genetics proper, will certainly increasingly occupy the attention of all geneticists who are aware even remotely of the needs of animal and plant improvement. They are, however, but natural outgrowths or branches of a central body of pure genetic analysis, and do not at all supersede it. This central preoccupation, to which I refer, is common to all the laboratory sciences, namely, the need to improve our capacity to foresee, and to predict, the consequences of all laboratory operations within our power. In terms of genetics, we need for every sort of organism suitable for laboratory or garden study, a more detailed knowledge of the architecture of its entire germplasm, and a more accurate formulation of the rules by which the gametes which the organism can produce may be predicted from prior knowledge of the gametes which entered into its composition (Owen 1950). Vogue and fashion notwithstanding, the science of genetics cannot afford to neglect the continued improvement of the tools of our craft, and particularly the continued investigation of the all-too-little understood rules of genetic recombination. The only adequate evidence of progress in this central field is the publication of data from experimental breeding, and though good data can never be easy to secure, the paucity of good three- or four-point linkage data, or equally the paucity of data on tetrads, complete sets of the four products of the same reduction division, would seem to indicate that promise has been taken as a substitute for performance in many departments. This is the more striking since observations of tetrads, such as can be obtained in the micro-fungi, have been for so long and so confidently advocated as the key to fuller quantitative knowledge of the process of gamete formation. How much there is still to learn of the fundamentals of this process is illustrated by the report only this year, by Michie (1953) from Oxford, and Wallace (1953) from Cambridge, of evidence of a new phenomenon, designated 'affinity' between non-homologous chromosomes, acting apparently through their centromeres.

What I have to say about population genetics *senso strictu*, that is, of the genetics of populations in which the selection of individuals for breeding, or for elimination, is not in the hands of the geneticist, so that the genetic structure of the population is not a human artifact, but exists, as it were, in its own right, as an object of naturalistic study—on this subject, what I have to say falls into four parts. First, I should like to sketch the background of theoretical notions in terms of which the genetical structure of a population, as an entity continuing and

changing in time, may be specified; next and finally, I should like to give three illustrations of the fields in which these ideas are required. In these fields several activities characteristic of genetical research are practically precluded. We cannot set about, for example, a course of matings needed to prepare a given genotype, or a line supplying a set of genotypes by simultaneous segregation, nor multiple recessive analyzer stocks, nor inbred lines, nor multiple heterozygotes for linkage tests; nevertheless, the genetic situations presented by a natural population may to a useful extent be fathomed, as I shall hope to show.

7. The theory of populations

The foundations of population genetics are laid in population statistics (Fisher 1930). Starting with the life table of actuarial science with its derived vectors the age-specific death-rate and the expectation of life, we associate with these the complementary aspect of vital statistics supplied by the age-specific rate of reproduction. Without ignoring, but without taking the time to elucidate the rather intricate theoretical minutiae which arise from bisexual reproduction—that the parents have two ages, so that a full specification of reproduction must be in terms of a bivariate function, taking account of the differences in the vital statistics of the two sexes—we may, certainly in mankind, and with a strong presumption in all other species, effectively average out the sex differences, and derive the fundamental equation for rate of population increase in the form that has become familiar:

$$\int_0^\infty \mathrm{e}^{-mx} l_x b_x \,\mathrm{d}x = 1,$$

where m is the parameter of population increase, named in honour of Malthus, l_x is the proportion of the population which survive to age x, and b_x is the rate of reproduction at that age. The integral is taken over all ages at which reproduction can take place. There is one and only one real root of this equation for m, though an infinite number of complex roots may exist. If the real root is positive, the vital statistics of the population are such as to favour an increase in numbers, if it is negative they will lead to a decrease. The real root necessarily exceeds the real parts of all the complex roots, so that such a population will approach a definite stable age distribution calculable from the real, or dominant, root. There is a complex age distribution associated in like manner with each complex root. A second, or conjugate, associated vector is provided for each root by the evaluation of the present value of the future offspring to be expected from any individual. For the dominant root, this measures how important each age class is for its contribution to future generations, and therefore as material for natural selection. It may be shown that if each age class is weighted with the appropriate valuation, the total value of the population will increase at the same rate, measured by m, whatever its composition by age may be.

These ideas in the field of vital statistics must be combined with our concepts of the genetic structure of individuals to build up a specification of the genetic structure of a population. This transition is of the utmost importance in the theoretical development of population genetics, for it is, I believe, at the root of

the paradoxes and differences of opinion associated with Sewall Wright's doctrine of the evolutionary importance of random variation in small isolated populations. At any locus, of which there must be thousands, at which more than one allelomorph exists in a population, all individuals may be classified in a number, perhaps only three, of distinct, or discontinuous, genotypes. Two whole brothers, offspring of the same two parents, may without restriction be of any two of these genotypes arbitrarily chosen. However many loci are considered simultaneously this discontinuity remains, and nearly related individuals may be unlike in many factors. Now, the frequencies with which the different genotypes occur define the gene ratios characteristic of the population, so that it is often convenient to consider a natural population not so much as an aggregate of living individuals as an aggregate of gene ratios. Such a change of viewpoint is similar to that familiar in the theory of gases, where the specification of the population of velocities is often more useful than that of a population of particles.

Unlike the genotypes of individuals the gene ratios can vary effectively continuously, so that the evolution of a species can be specified by a path in the appropriate multi-dimensional continuum. Even if such a path were to reach a stationary point—a very hypothetical proposal, for in such a case evolution would have ceased—the population would still contain among its myriads of genotypes, homozygotes, that is, extreme forms, in all the loci represented. If any of these, or combinations of them, present, in some particular locality, a selective advantage, there is nothing to prevent a shift in the gene ratios yielding higher adaptation, for the desired genotypes will occur already in the population and only require that their frequencies shall be increased, whereas Wright seems to ascribe selective value to the gene ratios, and not only to the genotypes, and pictures the population as tied by selective forces so closely to its point of representation that it cannot move to another 'peak', as he so describes it, offering greater advantages, except through the agency of enhanced random or chance variation. Actually, the effects of chance are the most accurately calculable, and therefore the least doubtful, of all the factors of an evolutionary situation, and only a rather deep-seated misunderstanding of the nature of the genetic specification of a population can explain the existence of differences of opinion about its evolutionary importance.

If in a population multiply-classified according to the genotype at each locus, a measurement of any kind is taken on each individual, the variation among these measurements is seen to be easily resolved according to the general principles of the analysis of variance, into a few intelligible fractions. If by the fixed environment of the species we mean not an environment determinately fixed for each individual, but a system of chances equally applicable to all such individuals, then the variance within groups having the same genotype, measured at the same stage of development, may be called the *environmental* component of variance; the remainder is *genotypic*. Of the genotypic variance, a fraction, usually a considerable fraction, is the portion calculable as a linear function of the gene frequencies, and is distinguished as the *genetic* variance; the remainder may be divided into two portions, a portion ascribable to *dominance*, or to interaction solely between substitutions at the same locus, and a second portion ascribable to *epistasy*, or to

interactions between substitutions at different loci. In controlled experimental work, as in a selective improvement programme, or in the analysis of the effects of crossing inbred lines of maize, such as is in progress at Raleigh, North Carolina, by Comstock & Robinson (1948), these four elements can be distinguished quantitatively; and their relative magnitudes are used to decide on one procedure of improvement rather than another. In natural populations, however, they are features of the genetic picture, comparatively permanent, or, at least, changing but slowly from one generation to the next.

These quantities are well defined for any quantitative measurement. If we consider the Malthusian parameter itself as a quantitative characteristic of the genotype, it is quite easily shown that the genetic variance of this quantity may be equated to the rate of increase per generation of its average value in the population. In this way the principle of natural selection is put upon an exact quantitative basis, and the course of spontaneous population change is related to its state at any instant. The improvement in relation to a fixed external environment is always positive, so that this relationship, like the second law of thermodynamics, is polarized relative to the time sequence. Of course, for any particular species, it does not follow that it is gaining on its competitors and enemies, or even that its increasing efficiency is enough to meet the needs of a deteriorating climate. Natural selection does not imply that every species is progressively better off, but that all species tend by all possible changes of gene ratio to improve their average condition.

There is, therefore, a fairly extensive theoretical background available to aid the task of understanding the populations of the real world, in so far as they are open to observation. The uses to which such a framework of ideas can be put vary greatly from case to case.

8. Virus genetics

It has been known for some time when strains of the same bacterial virus differing in two or more distinguishable characters are induced to infect the same bacterial cell that the virus particles which emerge, when finally the infected cell is disrupted, are not only of the two parental strains, but that other strains may be present exhibiting various combinations of the parental characters. This has been confirmed later by Burnet with the influenza virus. The series of forms obtainable at first strongly suggested an ordinary genetic cross, between organisms with at least three chromosomes, but showed some exceptional features which led to the idea that the process might be altogether different. For example, a mixture of three strains infecting the same cell could yield some offspring showing characters derived from all three parents.

An exceedingly interesting paper published earlier this year by Visconti & Delbrück (1953) puts a new interpretation on the whole of the phenomena. The biological foundation that these authors postulate is that the infecting virus particle, on entering the host cell, passes into what they (rather curiously) call a 'vegetative' phase, in which state they are capable of multiplication, and also of sexual reproduction by pairing and reduction; they are not apparently capable of living outside the host, or of infecting other cells. As the process of infection

advances, however, these particles proceed more and more to pass into the second or 'mature' phase, in which they are capable of leaving the disrupted host as infective particles. The relation between what goes *in* to the infected cell, and what comes *out*, is therefore that between the ancestors and the descendants in a freely interbreeding population, and Delbrück & Visconti claim that quantitative agreement is observed with this simple theory. It certainly explains three of the features which were previously puzzling. (*a*) Matings in successive generations are available to bring together three or more genes from different ancestors. There is no longer need to postulate a complete dissolution of the virus to explain this feature. (*b*) The apparent linkage observed between factors which on comparative evidence must be assigned to different linkage groups, is the natural consequence of a randomly variable number of generations behind the mature particles, so that some have experienced no sexual reproduction at all. If y_0 is the apparent recombination fraction between two such factors, the average number of generations experienced may be plausibly estimated as

$$m = -2 \ln (1 - 2y_0),$$

so that an observed value for the recombination fraction, $y_0 = 0{\cdot}45$, would correspond with m about 4·6 generations on the average. If the cell is prematurely lysed, an appropriately lower value of m is found. Also if y, a number less than y_0, is the apparent recombination fraction between two linked factors in the same test, the recombination fraction between them may be calculated as

$$p = \frac{1}{2} \frac{\ln(1-2y)}{\ln(1-2y_0)}.$$

(*c*) Thirdly, since we now have apparently unbiased estimates of recombination fractions, the previously puzzling feature of apparently negative interference, when three factors are used in the same chromosome, has disappeared. At this stage, if Delbrück & Visconti are right, sexual reproduction in phage particles has all the important features with which we are familiar in the higher animals and plants.

Apart from the brilliant scientific penetration which we should expect from these authors, the remarkable thing about this instance is that the uniformity of behaviour of the germinal material, which has always been so striking a feature in comparative genetics, should have been so far extended by the explanation of the apparent exceptions offered by viruses. It is very much to the point that in this same year Watson & Crick (1953) at Cambridge, with their associates of the University of London, should have been able to put forward a theory of the chemical structure of nucleic acid which seems admirably fitted to supply the physical basis of this uniformity.

9. Capture sampling

The study, in the field or in the sea, of wild species of motile animals has been greatly facilitated by that process of successive sampling known as the method of capture-recapture. Primarily, as in the classical pioneer researches of Jackson

(1940) with tsetse-fly populations in Tanganyika territory, the aim has been to estimate the population's density and its vital statistics. That on this foundation can be built, in favourable cases, an evolutionary study of gene ratios, of their clines in space and their changes with time, has been demonstrated in this country by Ford (Dowdeswell, Fisher & Ford 1940–9; Sheppard 1951) and his associates at Oxford. In the United States Dobshansky has made rather similar sampling studies, without much emphasis on population density or vital statistics, but with clear evidence of natural selection.

An efficient technique of capture-recapture is hedged about with a variety of precautions and checks, many of which were pointed out by Jackson. It is important that the means of marking shall be permanent, that they shall not be injurious, that they shall not expose the animal to danger, including any enhanced danger of recapture. Much experience is needed to arrive at satisfactory methods, and especially to demonstrate that they are satisfactory. Only entirely uninjured specimens should be released, and they should be distributed, or distribute themselves, uniformly over the unmarked population. In isolated communities receiving no immigrants, emigration is included in the death-rate as a single estimate; in continuous populations special measures may be needed to obtain separate estimates of the death-rate and the emigration rate, or equally of the accretions by birth or emergence, and by immigration. The vital statistics are often somewhat different in the two sexes. We must always expect that they may be different for different genotypes.

It will readily be understood that first attempts at such work encounter difficulties which it takes time and pertinacity to overcome; moreover, the collecting season of any one year may be short. Consequently, more than in most kinds of scientific work, ecological genetics needs the accumulation of experience, and a continuity of tradition, such as Ford and his associates have established at Oxford. Methods for the quantitative study of wild populations should, moreover, be more widely taught to foresters, and to all others who have responsible care of wild life, to which ignorance is at least as great a danger as indifference.

10. Polymorphism

A somewhat intricate genetical situation, namely, *selectively balanced polymorphism*, has been found in very diverse animals, insects, molluscs, fishes and man, with remarkably similar features, the causes of which are at present scarcely understood. Some years ago Dr R. K. Nabours of Manhattan, Kansas, studied the genetics in the genera *Paratettix*, *Apotettix* and *Acrydium* of the grouse locusts of Mexico and the southern United States, of the varied markings to be seen on the pronotum, and the conspicuous thigh-joints of the leaping legs. He uncovered the paradoxical situation, since paralleled elsewhere, of numerous variant genes completely dominant to their allelomorphs in the standard form of the species, linked in inheritance, in some species so closely that at first they were taken for a series of twenty or so allelomorphs at the same locus. Extensive breeding experiments showed indeed that the linkage was not absolute, and that many of these dominants could be combined in the same chromosome, so as to be inherited

together in coupling. Insects taken in the wild and showing two dominants are said, however, always to carry them in opposite chromosomes, or in repulsion.

From the incidence of the dominance I had inferred that the heterozygotes must in nature be at a selective advantage compared with the standard type, and that for equilibrium of the gene ratios, these advantages must be counterbalanced by other selective effects, presumably by a more severe disadvantage of the homozygous dominants, which are indistinguishable in appearance from the heterozygotes. It was partly to test this view that Nabours and Sabrosky carried out a collecting expedition in Mexico and Texas in 1933, some of the results of which are very clear and striking (Fisher 1939). In *Paratettix texanus* for example, of which the six samples taken yielded over 3000 insects, the double dominants must have suffered an elimination relative to other genotypes of 40 to 50 % between the fertilization of the egg and the time of capture. Considerations of stability require that the elimination of homozygotes should be just as intense. The elimination of double heterozygotes sufficiently explains the absence of chromosomes with two dominants in coupling, for these would be subject to this intense elimination generation after generation, whereas the doubles in repulsion will be made up afresh by singly dominant parents. The violent selective interaction of such pairs of factors is probably a potent factor in producing the close linkage observed between them. The single dominants as a whole must be at a selective advantage, the magnitude of which is estimated at 7, 10 and 14 % respectively in the three best represented localities.

Passing over other cases illustrating population genetics in various aspects, I should like to stress the similarities of the situation so far uncovered in the grouse locusts with that apparent in the blood groups in man. In 1900 Landsteiner discovered the *A* and *B* factors, and the first genetic interpretation was that there were two different factors with variant forms completely dominant to the common type *O*. Later, Bernstein pointed out on statistical grounds that they must be allelic, since double dominants *AB* always segregated *A* versus *B* and never both (*AB*) versus neither (*O*). In other words, in the families observed, the two dominants were always in repulsion, as in the grouse locusts, although the opposite interpretation, allelism rather than closely linked factors, was being placed on the observations. In 1927, however, Kirihara & Haku (1928) did propose the interpretation of close linkage.

11. The Rhesus blood groups

In the case of the Rhesus system, the first inclination was to speak of the numerous types that could be distinguished as *alleles*. In 1944 the position was sufficiently clear for Race and myself (Race 1944) to point out that the seven combinations then known could be interpreted as seven of the eight combinations possible to three pairs of genes *C* versus *c*, *D* versus *d*, and *E* versus *e*, and that the four antibodies known at that time were four out of six specific for these six antigens. Before the year was over Mourant (1945) had discovered the fifth antibody, anti-*e*, and in the two following years Diamond at Harvard and Hill & Haberman in Louisiana had reported the sixth, anti-*d* (Race & Sanger 1950; Hill & Haberman,

1948). The eighth antigenic combination R_y was found by Miss Van den Bosch in Belgium (1948) some years after its existence as a rare allelomorph, or combination, had been anticipated. We now have examples in England by reason of further discoveries in this system made by Race & Sanger towards the end of last year.

Now the immediately observable consequences of the two views (*a*) that we have a complicated system of multiple allelomorphs, and (*b*) that we have a series of three closely linked factors, are so slight as to be negligible. In the first case there will be no recombination in each generation, in the second case there may be none, or perhaps a very little; and these two possibilities are not directly distinguishable if other causes, errors of record, false ascriptions of paternity, and mutations, produce anomalous results with greater frequency than the hypothetical crossovers. If there is a controversy between the two views it is one that the practical pathologist can treat with complete indifference. On the other hand, to an evolutionist interested in the different ways in which peculiar genetic situations can establish themselves in natural populations, it matters a great deal, and I should now like to put forward, as an example of the sorts of evidence and reasoning available in population genetics, my reasons for thinking the loci of the Rhesus factor to be separable, and that in European conditions (though not in Africa) at least four of the combinations are being steadily eliminated by selection, and are only maintained by rare recombinations between the separable loci.

As the material is a little intricate I should like to make use of a few slides, which will appear in print as tables. In table 1 are shown the eight combinations of the letters *d*, *c*, *e*, with the frequencies of these combinations in the English population; naturally the lowest frequencies are the least accurately observed, and that given for R_y is derived by calculation and not by enumeration. In the arrangement shown the three common alleles known briefly as R_1, r and R_2 occupy the second line, under R_0, from which each can be derived by a single substitution at one of the three loci. These constitute 96 % of all alleles present. The third line shows the three rare combinations which like R_0 can be derived from the common ones by a single change. Finally, there is R_y opposite to R_0 in all three steps, and two steps away from each of the common combinations. Its low frequency puts it in a class by itself.

TABLE 1

	$R_0\,Dce$ 2·5 %	
$R_1\,DCe$ 43 %	$rdce$ 39 %	$R_2\,DcE$ 14 %
$R''\,dcE$ 1·2 %	$R_z\,DCE$ 0·25 %	$R'\,dCe$ 0·5 %
	$R_y\,dCE$ 0·01 %	

Table 2 shows the explanation I should give of these frequency classes by very rare crossing-over accompanied by a very gentle continuous elimination of its products. It is seen that R_0 is produced by the same processes as R', R'' and R_z, and this supplies a reason for expecting that its frequency should be about as great

as theirs put together. That R_y is not produced by any of these processes is a clear reason for its exceptional rarity. The frequencies of R', R'' and R_z relative to the heterozygotes that produce them suggest that crossing-over between D and E is more frequent than between D and C, or between C and E, so it is probable that C is situated between D and E. If this is so the only two triple heterozygotes capable of producing R_y must be rR_z and R_2R', for R_1R'' would need a double cross-over. The expected frequency of R_y implies that it should be very rare, though common enough to be found occasionally. It is neither too high nor too low for our rather vague positive knowledge of its frequency.

TABLE 2

heterozygote	segment	products
R_1r 33 %	C–D	R_0, R'
R_1R_2 11 %	C–E	R_0, R_z
rR_2 10 %	D–E	R_0, R''

the highest ratio:

$$\frac{R''}{rR_2} > \frac{R_z}{R_1R_2} > \frac{R'}{R_1r}$$

order

$D \quad C \quad E$

sources of R_y

$$\frac{r}{R_2}\frac{R_z}{R'}\ 0{\cdot}20\,\%\ D\text{–}C \qquad 0{\cdot}13\,\%\ C\text{–}E \qquad R_y = \frac{2R_zR'}{R_1} = 0{\cdot}006\,\%$$

This view receives useful confirmation from comparative data from other peoples. For example, among the Mexicans r is rare, consequently the only common double heterozygote is R_1R_2 capable of producing R_0 and R_z, but not R' and R'', and this is just what has been found among the Mexicans. Similarly, among the Basques R_2 is rare; the only common double heterozygote should yield R_0 and R', and not R'' and R_z. This also is confirmed factually.

A piece of direct evidence of the production of an R_0 gene by an R_1r father has been reported by Bentley Glass from Baltimore; in view of the possibilities of error already mentioned it would need a series of such cases to be wholly convincing. More remarkable is the deletion reported by Race & Sanger, observed in a lady homozygous for the deleted chromosome D– –. This definitely establishes the separability of the genes, though not necessarily separability by crossing-over; it also confirms that D is not the locus in the middle.

The most remarkable confirmation has come from a new antibody (Rosenfield *et al.* 1953) termed anti-f, found by Race and his associates in the serum of a haemophilic patient. With this antibody positive reactions are almost invariable with r, and common to the extent of about 90 % with R_0; they are unknown or very exceptional with the other Rhesus combinations. Table 3 shows how this unexpected situation could arise for an antigen f originating with r (*cde*), and located in the chromosome close to and probably outside e. The confirmation of the origin by crossing-over of R_0, R'', R', R_z is strong, for it explains the positive reactions of most R_0, and the negative reactions of the others. The use of the new reagent to discover the rare combination R_y is incidental. Before anti-f was discovered rR_y could not be distinguished from the commoner combination $R'R''$

for each contains *dCcEe* but not *D*. The first genotype, however, since it contains *r*, also reacts with *f*, and can now be picked out; family evidence is no longer necessary to recognize these cases.

TABLE 3

heterozygote	segment	products
R_1F/rf	$D\ C$	$R_0f,\ R'F$
	$C\ F$	$R_1f,\ rF$
R_1F/R_2F	$C\ E$	$R_0F,\ R_2F$
rf/R_2F	$D\ E$	$R_0f,\ R''F$
	$E\ F$	$R_2f,\ rF$

The moral I should draw from these examples is a trite one in this company, though often overlooked elsewhere; that our best way towards understanding our own species is often through the study of what St Francis might have called our little brother, the grouse locust, and even our little sister, the bacteriophage.

REFERENCES

van den Bosch, C. 1948 The very rare *Rh* genotype R_yr (*CdE/cde*) in a case of erythroblastosis foetalis. *Nature, Lond.*, **167**, 781.

Comstock, R. E. & Robinson, H. R. 1948 The components of genetic variance in populations of bi-parental progenies. *Biometrics*, **4**, 254–266.

Darlington, C. D. & Mather, K. 1949 *The elements of genetics.* London: George Allen & Unwin, Ltd.

Dowdeswell, W. H., Fisher, R. A. & Ford, E. B. 1940–1949 The quantitative study of populations in the lepidoptera. *Ann. Eugen., Lond.*, **10**, 123–136; *Heredity*, **3**, 67–84.

Fisher, R. A. 1930 *Genetical theory of natural selection.* Oxford: Clarendon Press.

Fisher, R. A. 1939 Selective forces in wild populations of *Paratettix texanus*. *Ann. Eugen., Lond.*, **9**, 109–122.

Fisher, R. A. 1947 The theory of linkage in polysomic inheritance. *Phil. Trans.* B, **233**, 55–87.

Fisher, R. A. 1949*a* The linkage problem in a tetrasomic wild plant, *Lythrum salicaria*. *Proc. VIII Int. Conf. Genet. Stockholm*, Suppl. Vol., pp. 225–233.

Fisher, R. A. 1949*b* *The theory of inbreeding.* Edinburgh: Oliver and Boyd.

Fisher, R. A. 1950 A class of enumerations of importance to genetics. *Proc. Roy. Soc.* B, **136**, 509–520.

Hill, J. M. & Haberman, S. 1948 Two examples of sera containing the anti-*d* agglutinin predicted by Fisher and Race. *Nature, Lond.*, **161**, 688.

Jackson, C. H. N. 1940 The analysis of a tsetse-fly population. *Ann. Eugen., Lond.*, **10**, 332–369.

Kirihara, S. & Haku, R. 1928 The hereditary law of the human blood groups. *Nagoya J. Med. Sci.*, **2**, 75–102.

Lederberg, J. 1951 *Papers in microbial genetics.* Madison: University of Wisconsin Press.

Lerner, I. M. 1950 *Population genetics and animal improvement.* Cambridge University Press.

Mather, K. 1949 *Biometrical genetics.* London: Methuen and Co. Ltd.

Michie, D. 1953 Affinity: a new genetic phenomenon in the house-mouse. Evidence from distant crosses. *Nature, Lond.*, **171**, 26–27.

Mourant, A. E. 1945 A new *Rhesus* antibody. *Nature, Lond.*, **155**, 542.

Owen, A. R. G. 1950 Theory of genetical recombination. *Advanc. Genet.*, **3**, 117–157.

Race, R. R. 1944 An 'incomplete' antibody in human serum. *Nature, Lond.*, **153**, 771.

Race, R. R. & Sanger, R. 1950 *Blood groups in man.* Oxford: Blackwell Scientific Publications.

Rosenfield, R. E., Vogel, P., Gibbel, N., Sanger, R. & Race, R. R. 1953 A 'new' *Rh* antibody, anti-*f*. *Brit. Med. J.* i, 975.

Sheppard, P. M. 1951 Fluctuations in the selective value of certain phenotypes in the polymorphic land-snail, *Cepaea nemoralis* L. *Heredity*, **5**, 125–134.

Sonnerborn, T. M. 1947 Recent advances in the genetics of *paramecium* and *euplotes*. *Advanc. Genet.*, **1**, 264–358.

Visconti, N. & Delbrück, M. 1953 The mechanism of genetic recombination in phage. *Genetics*, **38**, 5–33.

Wallace, M. E. 1953 Affinity: a new genetic phenomenon in the house-mouse. Evidence from within laboratory stocks. *Nature, Lond.*, **171**, 27–28.

Watson, J. D. & Crick, F. H. C. 1953 Molecular structure of nucleic acids. *Nature, Lond.*, **171**, 737–741.