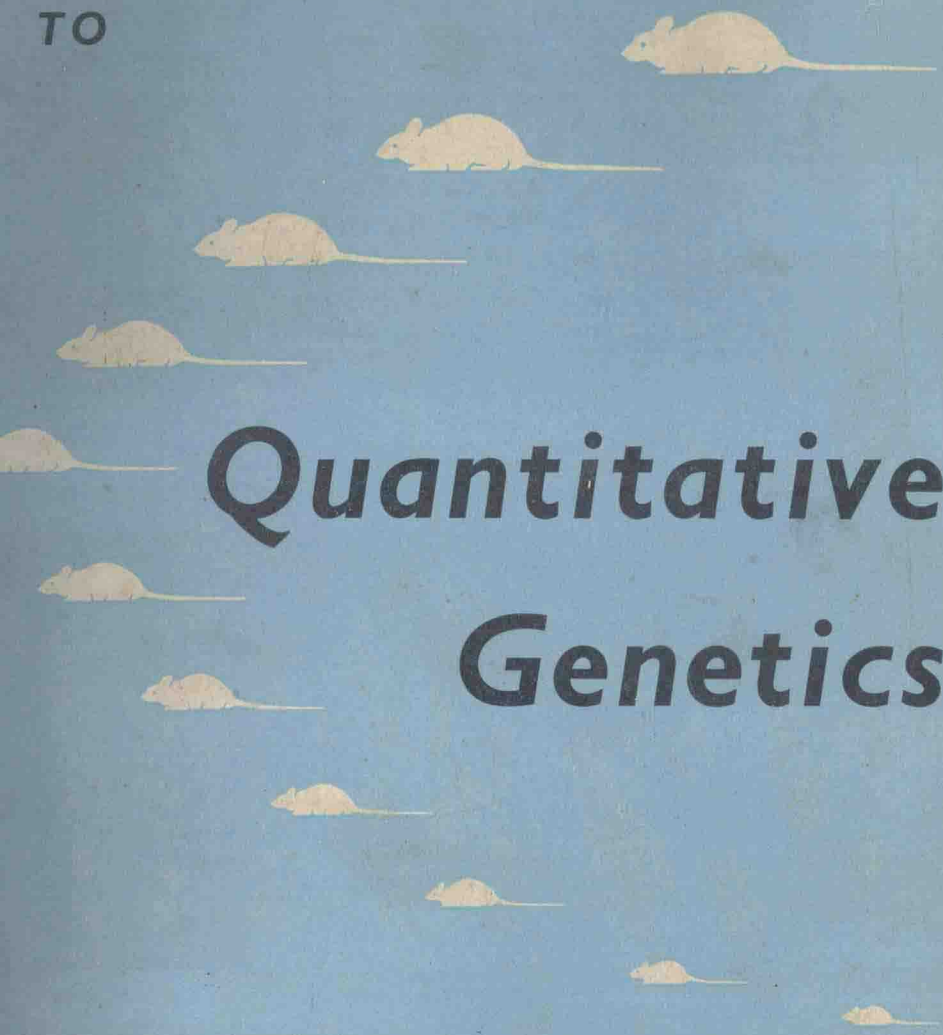
INTRODUCTION

TO



Quantitative Genetics

D. S. FALCONER

INTRODUCTION TO QUANTITATIVE GENETICS

D. S. FALCONER

Agricultural Research Council's Unit of Animal Genetics
University of Edinburgh

OLIVER AND BOYD
EDINBURGH AND LONDON
1960

OLIVER AND BOYD LTD

Tweeddale Court
Edinburgh 1

39a Welbeck Street
London W.1

FIRST PUBLISHED 1960

Printed in Great Britain for Oliver and Boyd Ltd
by Robert MacLehose and Company Limited, Glasgow

PREFACE

My aim in writing this book has been to provide an introductory textbook of quantitative genetics, with the emphasis on general principles rather than on practical application, and one moreover that can be understood by biologists of no more than ordinary mathematical ability. In pursuit of this latter aim I have set out the mathematics in the form that I, being little of a mathematician, find most comprehensible, hoping that the consequent lack of rigour and elegance will be compensated for by a wider accessibility. The reader is not, however, asked to accept conclusions without proof. Though only the simplest algebra is used, all the mathematical deductions essential to the exposition of the subject are demonstrated in full. Some knowledge of statistics, however, is assumed, particularly of the analysis of variance and of correlation and regression. Elementary knowledge of Mendelian genetics is also assumed.

I have had no particular class of reader exclusively in mind, but have tried to make the book useful to as wide a range of readers as possible. In consequence some will find less detail than they require and others more. Those who intend to become specialists in this branch of genetics or in its application to animal or plant breeding will find all they require of the general principles, but will find little guidance in the techniques of experimentation or of breeding practice. Those for whom the subject forms part of a course of general genetics will find a good deal more detail than they require. The section headings, however, should facilitate the selection of what is relevant, and any of the following chapters could be omitted without serious loss of continuity: Chapters 4, 5, 10 (after p. 168), 12, 13, and 15–20.

The choice of symbols presented some difficulties because there are several different systems in current use, and it proved impossible to build up a self-consistent system entirely from these. I have accordingly adopted what seemed to me the most appropriate of the

symbols in current use, but have not hesitated to introduce new symbols where consistency or clarity seemed to require them. I hope that my system will not be found unduly confusing to those accustomed to a different one. There is a list of symbols at the end, where some of the equivalents in other systems are given.

Acknowledgements

Many people have helped me in various ways, to all of whom I should like to express my thanks. I am greatly indebted to Professor C. H. Waddington for his encouragement and for the facilities that I have enjoyed in his laboratory. It is no exaggeration to say that without Dr Alan Robertson's help this book could not have been written. Not only has his reading of the manuscript led to the elimination of many errors, but I have been greatly assisted in my understanding of the subject, particularly its more mathematical aspects, by frequent discussions with him. Dr R. C. Roberts read the whole manuscript with great care and his valuable suggestions led to many improvements being made. Parts of the manuscript were read also by Dr N. Bateman, Dr J. C. Bowman, Dr D. G. Gilmour, Dr J. H. Sang, and my wife, to all of whom I am grateful for advice. I owe much also to the Honours and Diploma students of Animal Genetics in Edinburgh between 1951 and 1957, whose questions led to improvements of presentation at many points. Despite all the help I have received, many imperfections remain and there can hardly fail to be some errors that have escaped detection : the responsibility for all of these is entirely mine. To Mr E. D. Roberts I am indebted for drawing all the graphs and diagrams, and I greatly appreciate the care and skill with which he has drawn them. I am indebted also to the Director and Staff of the Commonwealth Bureau of Animal Breeding for assistance with the preparation of the list of references.

D. S. FALCONER

Institute of Animal Genetics, Edinburgh
December, 1958

CONTENTS

INTRODUCTION

Quantitative genetics is concerned with the inheritance of those differences between individuals that are of degree rather than of kind, quantitative rather than qualitative. These are the individual differences which, as Darwin wrote, "afford materials for natural selection to act on and accumulate, in the same manner as man accumulates in any given direction individual differences in his domestic productions." An understanding of the inheritance of these differences is thus of fundamental significance in the study of evolution and in the application of genetics to animal and plant breeding; and it is from these two fields of enquiry that the subject has received the chief impetus to its growth.

Virtually every organ and function of any species shows individual differences of this nature, the differences of size among ourselves or our domestic animals being an example familiar to all. Individuals form a continuously graded series from one extreme to the other and do not fall naturally into sharply demarcated types. Qualitative differences, in contrast, divide individuals into distinct types with little or no connexion by intermediates. Examples are the differences between blue-eyed and brown-eyed individuals, between the blood groups, or between normally coloured and albino individuals. The distinction between quantitative and qualitative differences marks, in respect of the phenomena studied, the distinction between quantitative genetics and the parent stem of "Mendelian" genetics. In respect of the mechanism of inheritance the distinction is between differences caused by many or by few genes. The familiar Mendelian ratios, which display the fundamental mechanism of inheritance, can be seen only when a gene difference at a single locus gives rise to a readily detectable difference in some property of the organism. Quantitative differences, in so far as they are inherited, depend on gene differences at many loci, the effects of which are not individually distinguishable. Consequently the Mendelian ratios are not exhibited by quantitative differences, and the methods of Mendelian analysis are inappropriate.

It is, nevertheless, a basic premiss of quantitative genetics that the inheritance of quantitative differences depends on genes subject to the same laws of transmission and having the same general properties as the genes whose transmission and properties are displayed by qualitative differences. Quantitative genetics is therefore an extension of Mendelian genetics, resting squarely on Mendelian principles as its foundation.

The methods of study in quantitative genetics differ from those employed in Mendelian genetics in two respects. In the first place, since ratios cannot be observed, single progenies are uninformative, and the unit of study must be extended to "populations," that is larger groups of individuals comprising many progenies. And, in the second place, the nature of the quantitative differences to be studied requires the measurement, and not just the classification, of the individuals. The extension of Mendelian genetics into quantitative genetics may thus be made in two stages, the first introducing new concepts connected with the genetic properties of "populations" and the second introducing concepts connected with the inheritance of measurements. This is how the subject is presented in this book. In the first part, which occupies Chapters 1 to 5, the genetic properties of populations are described by reference to genes causing easily identifiable, and therefore qualitative, differences. Quantitative differences are not discussed until the second part, which starts in Chapter 6. These two parts of the subject are often distinguished by different names, the first being referred to as "Population Genetics" and the second as "Biometrical Genetics" or "Quantitative Genetics." Some writers, however, use "Population Genetics" to refer to the whole. The terminology of this distinction is therefore ambiguous. The use of "Quantitative Genetics" to refer to the whole subject may be justified on the grounds that the genetics of populations is not just a preliminary to the genetics of quantitative differences, but an integral part of it.

The theoretical basis of quantitative genetics was established round about 1920 by the work of Fisher (1918), Haldane (1924–32, summarised 1932) and Wright (1921). The development of the subject over the succeeding years, by these and many other geneticists and statisticians, has been mainly by elaboration, clarification, and the filling in of details, so that today we have a substantial body of theory accepted by the majority as valid. As in any healthily growing science, there are differences of opinion, but these are chiefly

matters of emphasis, about the relative importance of this or that aspect.

The theory consists of the deduction of the consequences of Mendelian inheritance when extended to the properties of populations and to the simultaneous segregation of genes at many loci. The premiss from which the deductions are made is that the inheritance of quantitative differences is by means of genes, and that these genes are subject to the Mendelian laws of transmission and may have any of the properties known from Mendelian genetics. The property of "variable expression" assumes great importance and might be raised to the status of another premiss: that the expression of the genotype in the phenotype is modifiable by non-genetic causes. Other properties whose consequences are to be taken into account include dominance, epistasis, pleiotropy, linkage, and mutation.

These theoretical deductions enable us to state what will be the genetic properties of a population if the genes have the properties postulated, and to predict what will be the consequences of applying any specified plan of breeding. In principle we should then be able to make observations of the genetic properties of natural or experimental populations, and of the outcome of special breeding methods, and deduce from these observations what are the properties of the genes concerned. The experimental side of quantitative genetics, however, has lagged behind the theoretical in its development, and it is still some way from fulfilling this complementary function. The reason for this is the difficulty of devising diagnostic experiments which will unambiguously discriminate between the many possible situations envisaged by the theory. Consequently the experimental side has developed in a somewhat empirical manner, building general conclusions out of the experience of many particular cases. Nevertheless there is now a sufficient body of experimental data to substantiate the theory in its main outlines; to allow a number of generalisations to be made about the inheritance of quantitative differences; and to enable us to predict with some confidence the outcome of certain breeding methods. Discussion of all the difficulties would be inappropriate in an introductory treatment. The aim here is to describe all that is reasonably firmly established and, for the sake of clarity, to simplify as far as is possible without being misleading. Consequently the emphasis is on the theoretical side. Though conclusions will often be drawn directly from experimental data, the experimental side of the subject is presented chiefly in the form of

examples, chosen with the purpose of illustrating the theoretical conclusions. These examples, however, cannot always be taken as substantiating the postulates that underlie the conclusions they illustrate. Too often the results of experiments are open to more than one interpretation.

No attempt has been made to give exhaustive references to published work in any part of the subject; or to indicate the origins, or trace the history, of the ideas. To have done this would have required a much longer book, and a considerable sacrifice of clarity. The chief sources, from which most of the material of the book is derived, are listed below. These sources are not regularly cited in the text. References are given in the text when any conclusion is stated without full explanation of its derivation. These references are not always to the original papers, but rather to the more recent papers where the reader will find a convenient point of entry to the topic under discussion. References are also given to the sources of experimental data, but these, for reasons already explained, cover only a small part of the experimental side of the subject. In particular, a great deal more work has been done on plants and on farm animals than would appear from its representation among the experimental work cited.

Chief Sources

(For details see List of References)

Fisher, R. A. (1930), *The Genetical Theory of Natural Selection.*
Haldane, J. B. S. (1932), *The Causes of Evolution.*
Kempthorne, O. (1957), *An Introduction to Genetic Statistics.*
Lerner, I. M. (1950), *Population Genetics and Animal Improvement.*
Li, C. C. (1955), *Population Genetics.*
Lush, J. L. (1945), *Animal Breeding Plans.*
Malécot, G. (1948), *Les Mathématiques de l'Hérédité.*
Mather, K. (1949), *Biometrical Genetics.*
Wright, S. (1921), Systems of Mating. *Genetics* **6**: 111–178.
— — (1931), Evolution in Mendelian Populations. *Genetics* **16**: 97–159.

CHAPTER 1

GENETIC CONSTITUTION OF A POPULATION

Frequencies of Genes and Genotypes

To describe the genetic constitution of a group of individuals we should have to specify their genotypes and say how many of each genotype there were. This would be a complete description, provided the nature of the phenotypic differences between the genotypes did not concern us. Suppose for simplicity that we were concerned with a certain autosomal locus, A, and that two different alleles at this locus, A_1 and A_2, were present among the individuals. Then there would be three possible genotypes, A_1A_1, A_1A_2, and A_2A_2. (We are concerned here, as throughout the book, exclusively with diploid organisms.) The genetic constitution of the group would be fully described by the proportion, or percentage, of individuals that belonged to each genotype, or in other words by the frequencies of the three genotypes among the individuals. These proportions or frequencies are called *genotype frequencies*, the frequency of a particular genotype being its proportion or percentage among the individuals. If, for example, we found one quarter of the individuals in the group to be A_1A_1, the frequency of this genotype would be 0·25, or 25 per cent. Naturally the frequencies of all the genotypes together must add up to unity, or 100 per cent.

Example 1.1. The M-N blood groups in man are determined by two alleles at a locus, and the three genotypes correspond with the three blood groups, M, MN, and N. The following figures, taken from the tabulation of Mourant (1954), show the blood group frequencies among Eskimoes of East Greenland and among Icelanders as follows:

		Blood group			*Number of individuals*
		M	MN	N	
Frequency, %.	Greenland	83·5	15·6	0·9	569
	Iceland	31·2	51·5	17·3	747

Clearly the two populations differ in these genotype frequencies, the N blood group being rare in Greenland and relatively common in Iceland. Not only is this locus a source of variation within each of the two populations, but it is also a source of genetic difference between the populations.

A population, in the genetic sense, is not just a group of individuals, but a breeding group; and the genetics of a population is concerned not only with the genetic constitution of the individuals but also with the transmission of the genes from one generation to the next. In the transmission the genotypes of the parents are broken down and a new set of genotypes is constituted in the progeny, from the genes transmitted in the gametes. The genes carried by the population thus have continuity from generation to generation, but the genotypes in which they appear do not. The genetic constitution of a population, referring to the genes it carries, is described by the array of *gene frequencies*; that is by specification of the alleles present at every locus and the numbers or proportions of the different alleles at each locus. If, for example, A_1 is an allele at the A locus, then the frequency of A_1 genes, or the gene frequency of A_1, is the proportion or percentage of all genes at this locus that are the A_1 allele. The frequencies of all the alleles at any one locus must add up to unity, or 100 per cent.

The gene frequencies at a particular locus among a group of individuals can be determined from a knowledge of the genotype frequencies. To take a hypothetical example, suppose there are two alleles, A_1 and A_2, and we classify 100 individuals and count the numbers in each genotype as follows:

	A_1A_1	A_1A_2	A_2A_2	*Total*	
Number of individuals	30	60	10	100	
Number of genes A_1	60	60	0	120	200
Number of genes A_2	0	60	20	80	

Each individual contains two genes, so we have counted 200 representatives of the genes at this locus. Each A_1A_1 individual contains two A_1 genes and each A_1A_2 contains one A_1 gene. So there are 120 A_1 genes in the sample, and 80 A_2 genes. The frequency of A_1 is therefore 60 per cent or 0·6, and the frequency of A_2 is 40 per cent or 0·4. To express the relationship in a more general form, let the frequencies of genes and of genotypes be as follows:

	Genes		*Genotypes*		
	A_1	A_2	A_1A_1	A_1A_2	A_2A_2
Frequencies	p	q	P	H	Q

so that $p+q=1$, and $P+H+Q=1$. Since each individual contains two genes, the frequency of A_1 genes is $\frac{1}{2}(2P+H)$, and the relationship between gene frequency and genotype frequency among the individuals counted is as follows:

$$\left.\begin{aligned} p &= P+\tfrac{1}{2}H \\ q &= Q+\tfrac{1}{2}H \end{aligned}\right\} \ldots\ldots (1.1)$$

EXAMPLE 1.2. To illustrate the calculation of gene frequencies from genotype frequencies we may take the M-N blood group frequencies given in Example 1.1. The M and N blood groups represent the two homozygous genotypes and the MN group the heterozygote. The frequency of the M gene in Greenland is, from equation 1.1, $0{\cdot}835+\frac{1}{2}(0{\cdot}156)=0{\cdot}913$, and the frequency of the N gene is $0{\cdot}009+\frac{1}{2}(0{\cdot}156)=0{\cdot}087$, the sum of the frequencies being 1·000 as it should be. Doing the same for the Iceland sample we find the following gene frequencies in the two populations, expressed now as percentages:

	Gene	
	M	N
Greenland	91·3	8·7
Iceland	57·0	43·0

Thus the two populations differ in gene frequency as well as in genotype frequencies.

The genetic properties of a population are influenced in the process of transmission of genes from one generation to the next by a number of agencies. These form the chief subject-matter of the next four chapters, but we may briefly review them here in order to have some idea of what factors are being left out of consideration in this chapter. The agencies through which the genetic properties of a population may be changed are these:

Population size. The genes passed from one generation to the next are a sample of the genes in the parent generation. Therefore the gene frequencies are subject to sampling variation between successive generations, and the smaller the number of parents the greater is the sampling variation. The effects of sampling variation will be considered in Chapters 3–5, and meantime we shall exclude it from

the discussion by supposing always that we are dealing with a "large population," which means simply one in which sampling variation is so small as to be negligible. For practical purposes a "large population" is one in which the number of adult individuals is in the hundreds rather than in the tens.

Differences of fertility and viability. Though we are not at present concerned with the phenotypic effects of the genes under discussion, we cannot ignore their effects on fertility and viability, because these influence the genetic constitution of the succeeding generation. The different genotypes among the parents may have different fertilities, and if they do they will contribute unequally to the gametes out of which the next generation is formed. In this way the gene frequency may be changed in the transmission. Further, the genotypes among the newly formed zygotes may have different survival rates, and so the gene frequencies in the new generation may be changed by the time the individuals are adult and themselves become parents. These processes are called selection, and will be described in Chapter 2. Meanwhile we shall suppose they are not operating. It is difficult to find examples of genes not subject to selection. For the purpose of illustration, however, we may take the human blood-group genes since the selective forces acting on these are probably not very strong. Genes that produce a mutant phenotype which is abnormal in comparison with the wild-type are, in contrast, usually subject to much more severe selection.

Migration and mutation. The gene frequencies in the population may also be changed by immigration of individuals from another population, and by gene mutation. These processes will be described in Chapter 2, and at this stage will also be supposed not to operate.

Mating system. The genotypes in the progeny are determined by the union of the gametes in pairs to form zygotes, and the union of gametes is influenced by the mating of the parents. So the genotype frequencies in the offspring generation are influenced by the genotypes of the pairs that mate in the parent generation. We shall at first suppose that mating is at random with respect to the genotypes under discussion. *Random mating*, or *panmixia*, means that any individual has an equal chance of mating with any other individual in the population. The important points are that there should be no special tendency for mated individuals to be alike in genotype, or to be related to each other by ancestry. If a population covers a large geographic area individuals inhabiting the same locality are more

likely to mate than individuals inhabiting different localities, and so the mated pairs tend to be related by ancestry. A widely spread population is therefore likely to be subdivided into local groups and mating is random only within the groups. The properties of subdivided populations depend on the size of the local groups, and will be described under the effects of population size in Chapters 3–5.

Hardy-Weinberg Equilibrium

In a large random-mating population both gene frequencies and genotype frequencies are constant from generation to generation, in the absence of migration, mutation and selection; and the genotype frequencies are determined by the gene frequencies. These properties of a population were first demonstrated by Hardy and by Weinberg independently in 1908, and are generally known as the *Hardy-Weinberg Law*. (See Stern, 1943, where a translation of the relevant part of Weinberg's paper will be found.) Such a population is said to be in Hardy-Weinberg equilibrium. Deduction of the Hardy-Weinberg Law involves three steps: (1) from the parents to the gametes they produce; (2) from the union of the gametes to the genotypes in the zygotes produced; and (3) from the genotypes of the zygotes to the gene frequency in the progeny generation. These steps, in detail, are as follows:

1. Let the parent generation have gene and genotype frequencies as follows:

A_1	A_2		A_1A_1	A_1A_2	A_2A_2
p	q		P	H	Q

Two sorts of gametes are produced, those bearing A_1 and those bearing A_2. The frequencies of these gametic types are the same as the gene frequencies, p and q, in the generation producing them, for this reason: A_1A_1 individuals produce only A_1 gametes, and A_1A_2 individuals produce equal numbers of A_1 and A_2 gametes (provided, of course, there is no anomaly of segregation). So the frequency of A_1 gametes produced by the whole population is $P+\frac{1}{2}H$, which by equation 1.1 is the gene frequency of A_1.

2. Random mating between individuals is equivalent to random union among their gametes. We can think of a pool of gametes to which all the individuals contribute equally; zygotes are formed by