185

AVERAGE EXCESS AND AVERAGE EFFECT OF A GENE SUBSTITUTION

By R. A. FISHER

1. INTRODUCTION

S O M E years ago (Fisher, 1930), in the course of a study of the relations between the quantitative variation of natural populations, and the elementary factors which may contribute to this variation, I found it necessary to draw certain distinctions between ideas which had previously been ill developed or confused. In particular, as a first step towards clarifying the effects of selection, it was necessary to distinguish between what I defined as the *average excess* of a quantitative characteristic in relation to any given gene substitution, and a second quantity which I defined as the *average effect* of the gene substitution on the quantitative character in question. That these two quantities are distinct in definition, and in their quantitative values, may be made clear by quite simple examples.

The *average excess* of any measurement in respect of any gene substitution, such as that of **G** for **g**, is defined directly in terms of the numbers and mean measurements of the three genotypes **GG**, **Gg** and **gg**, into which the population may be divided. If the numbers of these are P, $2Q$ and R, and the mean measurements i, j and k, then

$$\frac{Pi + Qj}{P + Q} - \frac{Qj + Rk}{Q + R} = a, \tag{1}$$

is defined as the *average excess*. Thus, if gametes bearing **G** or **g** are chosen at random, and the zygotes to which they ultimately give rise are measured, the average for those bearing **G** will exceed the average for those bearing **g** by the difference a. This difference is not, however, to be ascribed to the gene substitution, as though it were necessarily an effect of it, and of it only; for by reason of such common phenomena as homogamy, or the mating of like with like, or a variety of similar causes, it may well be, and is probably the case in all real populations, that the two moities, of which the averages are compared, differ not only in the one gene substitution selected, but in the frequencies of a number of other genes affecting the measurement in question.

We require, therefore, to form an equally definite conception, explicitly defined, of what is meant by the *average effect* of a given gene substitution. It is natural to conceive this as the actual increase in the total of the measurements of a population, when without change in the environment, or in the mating system, the gene substitution is experimentally brought about, as it might be by mutation. If, however, we think thus in experimental terms it is necessary, for the idea to be applicable to all cases, to draw a further distinction. A change in the proportion of any pair of genes itself constitutes a change in the environment in which individuals of the species find themselves. In mammals, for example, a gene

substitution may affect milk yield, and thus directly influence infantile nutrition. To find the effect of the gene substitution in a constant environment, any effect of the gene substitution itself on the environment must be discounted. In other words, the effect that is wanted is only that due to the change in the frequencies of the different possible genotypes, not including any change in the average measurement of a given genotype, which the change in gene frequency may in fact bring about.

The logic of this distinction must be pressed even further, for, since the number of genotypes greatly exceeds the number of gene ratios on which their frequencies depend, we are concerned only with those changes of genotype frequency directly consequent on the proposed change of gene ratio, *in the actual condition of the population*, and not on any further changes in genotype frequency of such kinds as might be imposed without change of gene frequency, but which may in fact be brought about experimentally by a change in gene frequency. This might happen for example, by a gene increasing the frequency of self-fertilization, or by the existing frequency of self-fertilization exerting a more or less powerful effect when the primary gene ratio is changed.

The direct mathematical measure of the average effect of a proposed gene substitution is the partial regression, in the population as actually constituted, of the genotypic measurement on the numbers 0, 1 or 2 of the allelomorphic genes in each genotype. This is the natural measure of the average effect of the gene substitution, and avoids the difficulties which surround a definition based on the conceptual experiment discussed above. The experiment would determine the average effect combined with other effects acting through the gene environment brought about by change in gene ratios. These other effects which must often be infinitesimal, but may conceivably be important, are irrelevant to the effect of the gene substitution considered, in the actual circumstances in which we wish to evaluate it.

A simple example which brings out the main quantitative features, is provided by the case of a single factor in a population not mating at random; this may be due to homogamy, or to the partial isolation of sections of the population. It is well known that if mating were at random the frequencies P, $2Q$ and R of the three possible genotypes would be related so that $Q^2 = PR$. We shall consider the case in which this is not true, but, on the contrary, $Q^2 = \lambda PR$, where λ is a constant differing from unity. No complexity beyond this is needed to bring out the distinction we are illustrating. All individuals of the genotype **GG** may be characterized by a measurement i, all **Gg** by j, and all **gg** by k. Then, if a is the average excess we have

$$a = \frac{Pi + Qj}{P + Q} - \frac{Qj + Rk}{Q + R}.$$

We may now calculate the average effect, of substituting **G** for **g**, from the regression in the actual population of the genotypic measurement on the number of **G** genes; the 'expected' or genetic values of the genotypes will be $\mu + \alpha$, μ and $\mu - \alpha$, so that α and μ may be determined by minimizing the sum of squares

$$P(i - \mu - \alpha)^2 + 2Q(j - \mu)^2 + R(k - \mu + \alpha)^2$$

for variations of μ and α. This yields two equations for the unknowns, μ and α specifying the regression lines, namely,

$$P(\mu + \alpha - i) + 2Q(\mu - j) + R(\mu - \alpha - k) = 0,$$

$$P(\mu + \alpha - i) \qquad\qquad - R(\mu - \alpha - k) = 0,$$

or
$$(P + 2Q + R)\mu + (P - R)\alpha = Pi + 2Qj + Rk,$$

$$(P - R)\mu + (P + R)\alpha = Pi - Rk.$$

Solving these simultaneous equations for μ and α we find

$$(PQ + QR + 2PR)\alpha = P(Q + R)i - Q(P - R)j - R(P + Q)k. \tag{2}$$

It is therefore clear that even in the simplest case of non-random mating the values of a and α may be unequal. To equate one with the other is always inaccurate unless the full consequences of random mating are assured.

For a single factor, we may express a and α in a form in which they are more readily compared, namely,

$$\left.\begin{aligned} a &= \frac{P(Q+R)(i-j) + R(P+Q)(j-k)}{(P+Q)(Q+R)}, \\ \alpha &= \frac{P(Q+R)(i-j) + R(P+Q)(j-k)}{PQ+QR+2PR}, \end{aligned}\right\} \tag{3}$$

which are clearly equivalent when

$$(P + Q)(Q + R) = PQ + QR + 2PR,$$

or, otherwise stated, when $\qquad\qquad Q^2 = PR.$

When, on the contrary, Q^2 is less than PR, as is frequently the case, α will be less than a. It should be noticed that α, unlike a, is a true average of the measurable differences $i - j$, and $j - k$ produced by substituting **G** for **g** in heterozygotes and in homozygotes respectively, for in the case of α the denominator

$$PQ + QR + 2PR$$

is the sum of the coefficients of $i - j$ and $j - k$ in the numerator.

The nature of the quantity α measuring the average effect of a gene substitution may be illustrated by another fact. If the increase of genes **G** is distributed so as to give an increase of **GG** homozygotes

$$dP = \frac{2P(Q+R)}{P(Q+R) + R(P+Q)} dp,$$

the decrease of **gg** homozygotes will be

$$-dR = \frac{2R(P+Q)}{P(Q+R) + R(P+Q)} dp,$$

the same change requiring an increase of heterozygotes

$$2dQ = \frac{2Q(R-P)}{P(Q+R) + R(P+Q)} dp.$$

Then the increase in the total measurement of the population will be just α times the number of **g** gametes replaced by **G** gametes. In this case we observe that

$$\frac{dP}{P} + \frac{dR}{R} = \frac{2dQ}{Q},\qquad(4)$$

so that the change in genotypic frequencies ascribable only to change in gene ratio is that occurring when Q^2/PR maintains a constant ratio λ. If λ remains constant the actual change in the mean or total measurement in constant environment will be that due to change of gene ratio only. Constant genic environment is assured by a constant value of λ.

If, in fact, the value of λ changes, the change in the population mean will differ from that ascribable merely to the change in gene ratio, and this whether the change in λ is due to the change in gene ratio or to other causes. Whether λ remains constant or changes is, in fact, one of the things which cannot be predicted merely from the initial genotypic constitution of the population; for it requires knowledge of the ulterior causes governing differential viability, for example, or the choice of mates, by which the genotypic constitution is moulded.

This exact specification of the meaning of the *average effect* of a gene substitution is necessary for clarity of thought, and obviously also necessary for the development of its properties. It has, for example, been proved in general that the genetic variance of the population of diploid individuals in respect of any quantitative character is equal to

$$2\Sigma pq\alpha\alpha,$$

where the summation is taken over all pairs of allelomorphic genes. In the case of a single factor, in the absence of random mating, discussed above, the genetic variance is the variance of the genetic values $\mu + \alpha$, μ and $\mu - \alpha$ occurring with frequencies P, $2Q$ and R, namely,

$$\left\{(P+R) - \frac{(P-R)^2}{P+2Q+R}\right\}\alpha^2 = 2(PQ+QR+2PR)\frac{(P+Q)(Q+R)}{PQ+QR+2PR}\,a\alpha$$
$$= 2pqa\alpha,$$

the expression to which the general formula reduces in the case of a single factor. The multiplier 2 represents the number of loci in each zygote, and would be four for autotetraploids. In the case of a single factor it follows that a and α must be of the same sign, but this is not necessarily true if more than one factor is effective.

2. Application to the Measure of Biological Fitness

The most important application of this analysis (which can be applied to any measurable character) is to give a rational account of the action of natural selection. For this purpose I pointed out that we might choose as the metrical variable, characteristic of the different genotypes, the Malthusian parameter of population increase (m) measuring the survival value of each. If we do this, then a, the difference between the logarithmic rates of increase of two allelomorphic genes, is equal to

$$a = \frac{d}{dt}\log\frac{p}{q} = \frac{1}{pq}\frac{dp}{dt}.$$

The increase of average fitness of the population ascribable to a change in gene frequency dp will be
$$2\alpha dp.$$

Hence the rate of increase in the average value of the Malthusian parameter ascribable to natural selection acting on a single factor is
$$2pqa\alpha,$$

and the rate of increase due to all factors will be
$$2\Sigma pqa\alpha, \tag{5}$$

equal to the genetic variance of fitness due to all factors.

Since 1936, Sewall Wright has repeatedly put forward a formula of his own, namely,
$$\Delta p = \frac{pq}{2\overline{W}}\frac{d\overline{W}}{dp}. \tag{6}$$

In this we may recognize the portion $\quad \dfrac{1}{pq}\Delta p,$

as the average excess a of my formula expressed, not as a differential coefficient with respect to time, but as a finite difference for the lapse of one generation. As Wright does not regard his formula as limited to organisms having separated generations, and as he equates his expression to a differential coefficient on the right, this seems to be merely an accident in his notation. The remainder of Wright's formula, to which a is equated by him, is
$$\frac{1}{2\overline{W}}\frac{d\overline{W}}{dp},$$

and is evidently intended to represent the genetic effect of a change in the gene ratio p on $\frac{1}{2}\log\overline{W}$, where $\overline{W}$ is the mean fitness (e^m in my notation) of the different diploid zygotes. If all values of m are small, this is equivalent to α in my notation. Apart, then, from some inaccuracy of expression, which is probably unintentional, Wright's fundamental formula is merely
$$a = \alpha, \tag{7}$$

a relationship which is certainly not true in general, or approximately true in such a variable as the survival value of different genotypes.

It is, I think, clear from Sewall Wright's allusions to the subject that he has never clearly grasped the difficulties of interpretation of such expressions as
$$\frac{d\overline{W}}{dp},$$

in which the numerator involves the average of W for a number of different genotypes greatly exceeding the number of gene frequencies p on which their frequencies are taken to depend. It is likely, therefore, that he does not share my reasons for putting a particular and well defined meaning upon the phrase '*average effect* of a gene substitution'. The procedure of obtaining the rate of gene change by equating a to α is not, however, to be redeemed by finding a new definition for α. This may be seen even in the case of a single factor, where the effect of a given substitution must be in some cases $i - j$, and in others

$j - k$. For, as we have seen, the quantity a cannot be equated to any sort of average of these two quantities. It may, indeed, exceed both of them.

We have only to put, for example, $P = 0.3$, $Q = 0.2$, $R = 0.3$ for the phenotypic frequencies, then
$$a = \frac{0.15(i - j) + 0.15(j - k)}{0.25},$$
which, if $i - j$ and $j - k$ are equal, exceeds both of them in the ratio 6 : 5.

The attempt to equate a, measuring the selective intensity in favour of a given gene substitution, to the average effect of that substitution on the mean fitness of the population, however this average effect may be defined, is, therefore, foredoomed to failure just so soon as the simplifying, but unrealistic, assumption of random mating is abandoned.

3. Selection in favour of self-fertilization

Wright's conception embodied in equation (6) that selective intensities are derivable, like forces in a conservative system, from a simple potential function dependent on the gene ratios of the species as a whole, has led him into extensive but untenable speculations. For example, in the *New Systematics*, p. 170, we find: 'As already noted, $\overline{W}$ is a function of all gene frequencies. In the practically infinite field of gene combinations, possible from differences in only a few thousands, or even hundreds of loci, there are likely to be an enormous number of different harmonious combinations of characters. These would appear as peak values of $\overline{W}$, separated by valleys or saddles in a multidimensional surface.'

Prof. Wright here confuses the number of genotypes, e.g. 3^{1000}, which may be distinguished among individuals, with the continuous field of variation of gene frequencies. Even if a potential function, such as $\overline{W}$ is supposed to be, really existed, the large number of genotypes supplies no reason for thinking that even one peak, maximal for variations of *all* gene ratios should occur in this field of variation.

In regard to selection theory, objection should be taken to Wright's equation principally because it represents natural selection, which in reality acts upon individuals, as though it were governed by the average condition of the species or inter-breeding group. Early selectionists, following in this respect the language of earlier theological writers on organic adaptation, often speak of selection as directed 'for the good of the species'. In reality it is always directed to the good, as measured by descendants, of the individual. Unless individual advantage can be shown, natural selection affords no explanation of structures or instincts which appear to be beneficial to the species. Yet in Wright's equation the whole evolutionary sequence would appear to be governed by the principle of increasing the 'general good'. It may therefore be worth while to examine in detail a model involving powerful selection, in which the fitness of the species as a whole, judged by external criteria is entirely inoperative.

The model chosen is that of a gene, which, without affecting the fertility or pollen production of a plant, insures the self-fertilization of the ovules, as may be done, for example,

by cleistogamic flowers. In particular we shall suppose that homozygotes **gg** are fertilized at random by open pollen; in heterozygotes **Gg** half the ovules are fertilized by open pollen, and half are self-fertilized; while in homozygotes **GG** all ovules are self-fertilized, although the plant contributes its share to the openly disseminated pollen.

If P, $2Q$ and R stand for the frequencies of these three types of plant in any one generation, the frequencies in the next generation are calculated as in the following table, in which all progeny are counted twice:

Frequency	Genotype of parent		Genotype of offspring		
			gg	**Gg**	**GG**
P	**gg**	Open pollinated seed	$P(P+Q)$	$P(Q+R)$	—
		Pollen	$P(P+\frac{1}{2}Q)$	$\frac{1}{2}PQ$	—
$2Q$	**Gg**	Open pollinated seed	$\frac{1}{2}Q(P+Q)$	$\frac{1}{2}Q(P+2Q+R)$	$\frac{1}{2}Q(Q+R)$
		Selfed progeny	$\frac{1}{4}Q(P+2Q+R)$	$Q(P+2Q+R)$	$\frac{1}{2}Q(P+2Q+R)$
		Pollen	$Q(P+\frac{1}{2}Q)$	$Q(P+Q)$	$\frac{1}{2}Q^2$
R	**GG**	Selfed progeny	—	—	$2R(P+2Q+R)$
		Pollen	—	$R(P+\frac{1}{2}Q)$	$\frac{1}{2}QR$

In all, if P', $2Q'$ and R' are the frequencies in the second generation, it appears that

$$P' = P^2 + \tfrac{7}{4}PQ + Q^2 + \tfrac{1}{4}QR,$$
$$Q' = (P+Q)(Q+\tfrac{1}{2}R), \qquad (8)$$
$$R' = \tfrac{1}{4}PQ + Q^2 + PR + \tfrac{11}{4}QR + R^2.$$

These constitute the recurrence equations by which the constitution of each generation may be calculated from that of the last; in respect of gene frequency it follows that

$$P' + Q' = P + Q - \tfrac{1}{4}(PQ + 2PR + QR), \qquad Q' + R' = Q + R + \tfrac{1}{4}(PQ + 2PR + QR),$$

showing that selection always favours the gene **G**, which must constantly increase in frequency. Natural selection, indeed a selection of great intensity, is at work, although by external criteria the average fitness of the species as a whole is entirely independent of the gene ratio. Individuals differ greatly in their fitness to the circumstances in which they are placed. The mean fitness of the species as a whole is, however, no guide to the selective intensities in action.

The analytic solution of a set of equations such as (8), which constantly occur in genetic problems, offers apparent difficulties. These may frequently, perhaps always, be overcome by seeking an analytic expression valid for extreme values, as in the solution (Fisher, 1930, p. 85) of the exponential recurrence equations encountered in the theory of rare mutations. In the present instance, if the gene **G** is rare, both Q and R must be small, and P is brought as near as may be wished to unity, then the equations may be written

$$Q' = Q + \tfrac{1}{2}R, \qquad R' = \tfrac{1}{4}Q + R. \qquad (9)$$

If now we seek a solution such that

$$Q' = \lambda Q, \qquad R' = \lambda R,$$

t appears that

$$(\lambda - 1)^2 = \tfrac{1}{8};$$

and since λ must exceed unity $\qquad \lambda = 1 + \dfrac{1}{2\sqrt{2}},$

and $\qquad\qquad\qquad\qquad Q = 4(\lambda - 1)\,R = R\sqrt{2}.$

If, then, $\qquad\qquad\qquad\qquad x = \lambda^{\nu},$

where ν stands for the generation number, when x is small (ν large and negative) Q and R will change proportionately to x, and we may anticipate the convergence in the appropriate region of expansions such as

$$
\begin{aligned}
Q &= x\sqrt{2} + q_2 x^2 + q_3 x^3 + \ldots = \sum_{t=1}^{\infty} q_t x^t, \\
R &= x + r_2 x^2 + r_3 x^3 + \ldots = \sum_{t=1}^{\infty} r_t x^t.
\end{aligned}
\tag{10}
$$

These expressions may now be substituted in the exact recurrence equations (8)

$$
\begin{aligned}
\Sigma q_t \lambda^t x^t &= Q' = Q + \tfrac{1}{2}R - (Q+R)(Q+\tfrac{1}{2}R), \\
\Sigma r_t \lambda^t x^t &= R' = \tfrac{1}{4}Q + R + \tfrac{1}{2}Q(Q+R),
\end{aligned}
$$

and equations for the coefficients found by equating those of equal powers of x. In general,

$$
\begin{aligned}
\lambda^t q_t &= q_t + \tfrac{1}{2}r_t - \tfrac{1}{2}\sum_{s=1}^{t-1}(2q_s + r_s)(q_{t-s} + r_{t-s}), \\
\lambda^t r_t &= \tfrac{1}{4}q_t + r_t + \tfrac{1}{2}\sum_{s=1}^{t-1} q_s(q_{t-s} + r_{t-s}),
\end{aligned}
\tag{11}
$$

of which the summations on the right are known when q and r have been evaluated from 1 to $t-1$.

We thus obtain pairs of equations such as

$$
\begin{aligned}
0 \cdot 8321 q_2 - 0 \cdot 5 r_2 &= -4 \cdot 6213, \\
-0 \cdot 25 q_2 + 0 \cdot 8321 r_2 &= 1 \cdot 7070,
\end{aligned}
$$

and
$$
\begin{aligned}
1 \cdot 4798 q_3 - 0 \cdot 5 r_3 &= 21 \cdot 3659, \\
-0 \cdot 25 q_3 + 1 \cdot 4798 r_3 &= -9 \cdot 7634,
\end{aligned}
$$

which yield the series of solutions

t	q	r	t	q	r
1	$1 \cdot 4142,1356$	$1 \cdot 0000,0000$	4	$-25 \cdot 7043,377$	$11 \cdot 8437,469$
2	$-5 \cdot 2730,5835$	$0 \cdot 4671,4547$	5	$44 \cdot 9719,708$	$-23 \cdot 4828,802$
3	$12 \cdot 9477,773$	$-4 \cdot 4102,2710$	6	$-69 \cdot 5245,614$	$39 \cdot 9406,205$

Convergence seems sufficiently rapid if x is as small as $0 \cdot 05$. At this point $\nu = -9 \cdot 895616$, and the genotypic frequencies are

$$P \quad 0 \cdot 8313,1837,$$
$$Q \quad 0 \cdot 0589,9889,$$
$$R \quad 0 \cdot 0506,8385.$$

From these the frequencies in subsequent generations may be obtained by the direct application of equations (8).

A similar analysis is possible when the gene **g** is rare. In this case, putting $R = 1$, we have

$$P' = \tfrac{1}{4}Q, \quad Q' = \tfrac{1}{2}(P + Q); \tag{12}$$

and, if

$$P = \mu P', \quad Q = \mu Q',$$

then

$$1 - \tfrac{1}{2}\mu - \tfrac{1}{8}\mu^2 = 0,$$

or

$$\mu = 2(\sqrt{3} - 1),$$

and

$$Q = \frac{4}{\mu}P = (\sqrt{3} + 1)\,P.$$

Then, if

$$x = \mu^{-\nu'},$$

and in the exact equations (8) we substitute

$$P = x + p_2 x^2 + \ldots = \sum_{t=1}^{\infty} p_t x^t, \quad P' = \sum_{t=1}^{\infty} p_t \mu^{-t} x^t,$$

and

$$Q = (\sqrt{3} + 1)\,x + q_2 x^2 + \ldots = \sum_{t=1}^{\infty} q_t x^t, \quad Q' = \sum_{t=1}^{\infty} q_t \mu^{-t} x^t,$$

we obtain the equations for the coefficients

$$\left. \begin{aligned} p_t \mu^{-t} - \tfrac{1}{4}q_t &= \tfrac{1}{2}\sum_{s=1}^{t-1}(2p_s + q_s)(p_{t-s} + q_{t-s}), \\ -\tfrac{1}{2}p_t + (\mu^{-t} - \tfrac{1}{2})\,q_t &= -\tfrac{1}{2}\sum_{s=1}^{t-1}p_s(p_{t-s} + q_{t-s}), \end{aligned} \right\} \tag{13}$$

which lead to the solutions

	p	q		p	q
1	1·0000,0000	2·7320,5081	4	669·1361,56	− 1627·0013,3
2	5·4205,1219	− 25·2056,946	5	− 5907·93074	13641·1755
3	− 72·7411,373	201·7580,69	6	52629·6872	− 118904·651

The coefficients increase more rapidly than in the first case, but convergence seems sufficiently rapid when x is so small as 0·01. To find the relationship between ν and ν' the successive frequencies were calculated for sixteen generations as follows:

ν	P	Q	R
− 9·895616	0·8313,1837	0·0589,9889	0·0506,8385
− 8·895616	0·7811,5069	0·0750,9008	0·0686,6915
− 7·895616	0·7197,7318	0·0936,9385	0·0928,3912
− 6·895616	0·6470,4364	0·1139,7764	0·1250,0108
− 5·895616	0·5642,7809	0·1343,0365	0·1671,1461
− 4·895616	0·4746,8131	0·1521,9369	0·2209,3131
− 3·895616	0·3833,1747	0·1646,5458	0·2873,7337
− 2·895616	0·2963,5435	0·1689,6240	0·3657,5125
− 1·895616	0·2194,2404	0·1637,0543	0·4531,4510
− 0·895616	0·1563,5435	0·1495,3083	0·5445,8399
0·104384	0·1080,7882	0·1290,2935	0·6338,6248
1·104384	0·0731,8066	0·1057,4090	0·7153,3754
2·104384	0·0489,8849	0·0829,1398	0·7851,8355
3·104384	0·0326,5848	0·0627,2038	0·8419,0076
4·104384	0·0217,8611	0·0461,3197	0·8859,4995
5·104384	0·0145,7927	0·0332,1920	0·9189,8233
6·104384	0·0097,9558	0·0235,5080	0·9431,0282

The frequencies of the last line tally well with those obtained using $x = 0\cdot009374385$, these being

	By direct calculation	By expansion	Difference
P	0·0097,9558	0·0097,9558,52	‚52
Q	0·0235,5080	0·0235,5079,53	−‚47

This value of x corresponds with $\nu' = 12\cdot248851$, so that

$$\nu' - \nu = 6\cdot144467 \text{ generations.}$$
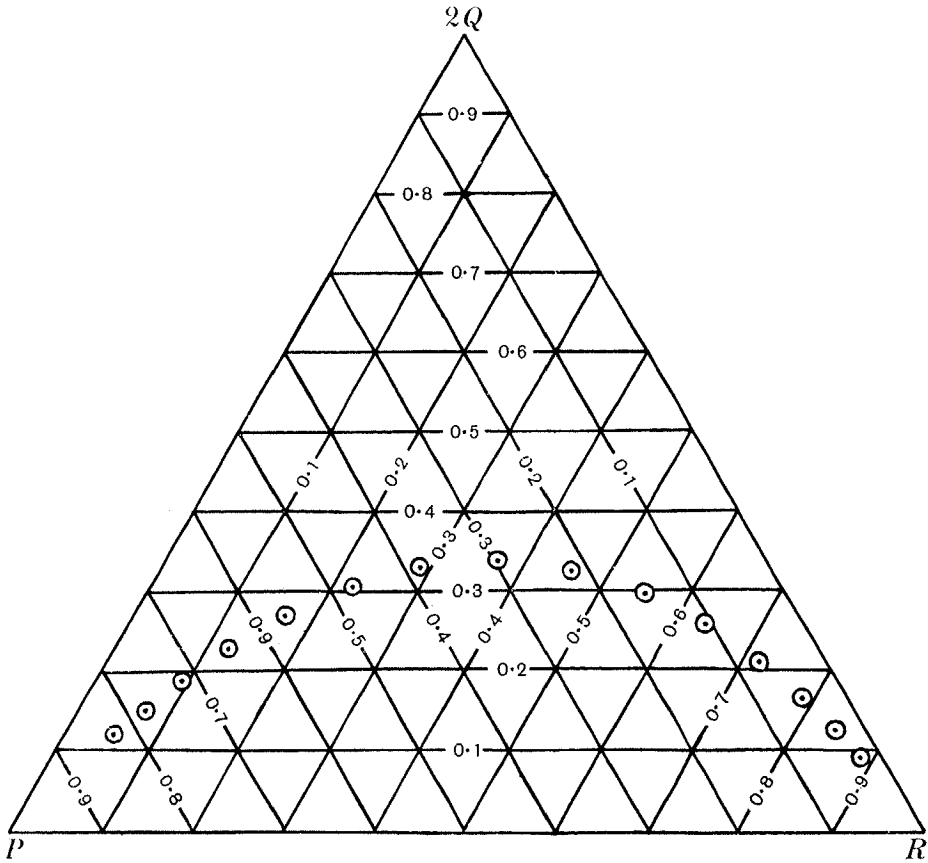
Trilinear diagram representing the field of variation of the coordinates P, $2Q$, R and the constitution of the populations in successive generations.

This constant, which suffices to complete the solution, arises from the convention that p_1 and r_1 in the expansion formulae should both be unity; if, for example q_1 had been taken to be unity in each case the constant would have been reduced by

$$\frac{\log \sqrt{2}}{\log \lambda} + \frac{\log (\sqrt{3} + 1)}{\log \mu}.$$

Indeed, a number of equally suitable conventions might have been adopted.

The natural logarithms of λ and μ represent the initial and final values of the selective intensity (a); these are:

	Formula	Value
Initial selective intensity	$\log\left(1 + \dfrac{1}{2\sqrt{2}}\right)$	0·30273
Final selective intensity	$\log 2(\sqrt{3} - 1)$	0·38124

Between these extreme values the selective intensity steadily increases as the gene **g** is replaced by **G**.

A second quantity which shows a similar steady increase is the ratio of the average effect α, to the average excess a. These are always of the same sign, as is necessarily the case where only a single factor is at work. In this case α is always the smaller; the ratio, easily calculable from the tabulated values of P, Q and R, rises from a limiting initial value $\sqrt{\tfrac{1}{2}} = 0\cdot70711$, to a limiting final value $\tfrac{1}{6}(3 + \sqrt{3}) = 0\cdot78868$. Thus both a and α increase as the population changes, the latter the more rapidly.

SUMMARY

The definitions of the primary quantities a and α arising in quantitative inheritance are elaborated.

It is not permissible to assume that a is equal to α except for groups in which random mating is and has for long been the rule.

Apart from some technical difficulties of interpretation, Wright's formula for natural selection

$$\Delta p = \frac{pq}{2\overline{W}} \frac{d\overline{W}}{dp},$$

seems intended merely to assert the equality of these two quantities for any factor influencing survival. If it were true, then selective intensities could in general be derived from a function of the species as a whole, without regard to the competition among individuals.

An example is examined in detail of a hypothetical factor affecting self-fertilization. Intense selective activity is shown to be compatible with an entire absence of change in the average survival value of the population. Throughout the process, the selective intensity, a, is greater than the average effect of the gene substitution, α, and changes at a different rate.

REFERENCES

R. A. FISHER (1930). *The Genetical Theory of Natural Selection.* Oxford: Clarendon Press.

SEWALL WRIGHT (1937). 'The distribution of gene frequencies in populations.' *Proc. nat. Acad. Sci.* **23**, 307–20.

—— (1940 a). 'Breeding structure of populations in relation to speciation.' *Amer. Nat.* **74**, 232–48.

—— (1940 b). 'The statistical consequences of Mendelian heredity in relation to speciation.' *The New Systematics*, pp. 161–82. Oxford: Clarendon Press.