

# THE INITIAL INCREASE OF A NEW GENE UNDER POSITIVE ASSORTATIVE MATING

P. A. PARSONS

Department of Genetics, University of Cambridge

Received 1.xii.61

## 1. INTRODUCTION

POSITIVE assortative mating, which is the tendency of similar phenotypes to mate with each other, has been studied both theoretically (Jennings, 1916; Fisher, 1918; Wright, 1921) and in the wild in animals such as the Blue Snow Goose (Cooch and Beardmore, 1959) and the Arctic Skua (O'Donald, 1959). In man, Pearson and Lee (1903) have demonstrated positive assortative mating for various physical characters such as stature and forearm length. Fisher (1930) has argued that assortative mating is potentially an important agent promoting evolutionary change. Its principle biometric effects seem to be to increase the genetic variance; so increasing in this way the effect of selection on human stature by more than 20 per cent. Some experimental studies with mutant genes indicate the existence of assortative mating. Rendel (1944) in *Drosophila subobscura* found that yellow males tended to mate with yellow females far more readily than with females of other genotypes. Mating preferences of various types have been described by Merrell (1949) and Rendel (1951) in *D. melanogaster* for various mutant stocks, and by Mather and Harrison (1949) for two wild type stocks of flies and selected lines derived from them. Many other examples exist, especially between closely related species.

The examination of equilibria attained under systems of mixed random mating and positive assortative mating has been carried out by O'Donald (1960) but he did not include parameters to take into account the selective values of the various genotypes. Inclusion of such parameters in the model will not permit a simple algebraic expression for the determination of the equilibrium positions and the conditions for their stability, but would require considerable computation. In such circumstances a good guide to some of the conditions for equilibria can be obtained by considering the conditions under which new genes will increase in a population, which, because of various approximations that can be made, can usually be found relatively easily. Furthermore, such conditions could in many circumstances be as important as the conditions for equilibria. Classical population genetic studies have concentrated on equilibria, but with the rapid changes of environment which are occurring at the present time due to the elimination of disease and the migration of humans, animals and plants to new areas the initial conditions are clearly relevant. Recently, various genetic systems have been examined

from this point of view (Bodmer, 1960*a*, *b*; Bodmer and Parsons, 1960, 1962; Parsons and Bodmer, 1961; Parsons, 1961*a*, *b*).

In this paper we shall examine the increase of new genes under mixed positive assortative mating and random mating, and compare these conditions with those derived for other genetic systems, in particular mixed selfing and random mating.

## 2. ASSORTATIVE MATING WITH ONE ALLELE DOMINANT

Let a proportion  $\alpha$  mate with like phenotypes and  $1-\alpha$  mate at random. Let the frequencies of the 3 genotypes  $AA$ ,  $Aa$  and  $aa$  be  $u$ ,  $v$  and  $w$  respectively. Gene  $A$  is assumed to be completely dominant so that  $AA$  and  $Aa$  are not distinguishable. Then  $\alpha(u+v)$  of  $A$  phenotypes and  $\alpha w$  of  $a$  phenotypes will mate with like phenotypes. Of the  $A$  phenotypes, there will be three mating types showing assortative mating, namely  $AA \times AA$ ,  $AA \times Aa$  and  $Aa \times Aa$  which will occur in proportions  $\frac{\alpha u^2}{(u+v)}$ ,  $\frac{2uv\alpha}{(u+v)}$  and  $\frac{\alpha v^2}{(u+v)}$ , a total of  $\alpha(u+v)$ . Then listing the 6 possible mating types we get the following genotypic frequencies in the progeny assuming the relative viabilities of  $AA$ ,  $Aa$  and  $aa$  to be  $a$ ,  $h$  and  $b$  respectively:

Mating type	Frequency	Proportion of each genotype taking viabilities into account		
		$AA$	$Aa$	$aa$
$AA \times AA$	$\frac{\alpha u^2}{u+v} + (1-\alpha)u^2$	$a$	...	...
$AA \times Aa$	$\frac{2\alpha uv}{u+v} + 2(1-\alpha)uv$	$\frac{1}{2}a$	$\frac{1}{2}h$	...
$AA \times aa$	$2(1-\alpha)uw$	...	$h$	...
$Aa \times Aa$	$\frac{\alpha v^2}{u+v} + (1-\alpha)v^2$	$\frac{1}{2}a$	$\frac{1}{2}h$	$\frac{1}{2}b$
$Aa \times aa$	$2(1-\alpha)vw$	...	$\frac{1}{2}h$	$\frac{1}{2}b$
$aa \times aa$	$\alpha w + (1-\alpha)w^2$	...	...	$b$

The above table without viabilities was given by O'Donald (1960). Thus letting  $u'$ ,  $v'$  and  $w'$  represent the genotypic frequencies in the next generation we get:

$$\begin{aligned}
 Tu' &= a \left[ \frac{\alpha(u^2 + uv + \frac{1}{2}v^2)}{u+v} + (1-\alpha)(u^2 + uv + \frac{1}{2}v^2) \right] \\
 &= a(u + \frac{1}{2}v)^2 \left[ \frac{\alpha}{u+v} + (1-\alpha) \right] \quad (2.1)
 \end{aligned}$$

$$\begin{aligned}
 Tv' &= h \left[ \frac{\alpha}{u+v} (uw + \frac{1}{2}v^2) + (1-\alpha)(uv + 2uw + \frac{1}{2}v^2 + vw) \right] \\
 &= h(u + \frac{1}{2}v) \left[ \frac{\alpha v}{u+v} + (1-\alpha)(v + 2w) \right] \quad (2.2)
 \end{aligned}$$

$$Tw' = b \left[ \alpha w + \frac{\alpha v^2}{u+v} + (1-\alpha)(\frac{1}{2}v + w)^2 \right] \quad (2.3)$$

where  $T$  is a total such that  $u' + v' + w' = u + v + w = 1$ . Now let us assume that the population is all  $AA$  and that gene  $a$  is introduced by mutation. Then  $v$  and  $w$  will be small, so ignoring second order terms in  $v$  and  $w$  we get:

$$\left. \begin{aligned}
 Tu' &= a(1-v-2w+\alpha w) \\
 Tv' &= h(v+2w-2w\alpha) \\
 Tw' &= b\alpha w
 \end{aligned} \right\} \quad (2.4)$$

Thus

$$\left. \begin{aligned}
 v' &= \frac{h}{a}v + \frac{2(1-\alpha)}{a}w \\
 w' &= \frac{b\alpha}{a}w
 \end{aligned} \right\} \quad (2.5)$$

To find the conditions under which the new gene  $a$  will increase, the 2 equations (2.5) may be represented as a  $2 \times 2$  matrix and the dominant (largest) latent root  $\lambda_b$  will determine whether or not the gene will increase in frequency. For an increase in frequency  $\lambda_b > 1$ .

We get two roots

$$\begin{aligned}
 \lambda_1 &= \frac{h}{a} \\
 \text{and } \lambda_2 &= \frac{b\alpha}{a}
 \end{aligned}$$

Now  $\lambda_1 - \lambda_2 > 0$  when  $h > b\alpha$ . Thus  $\lambda_1$  is the dominant latent root when  $h > b\alpha$  and  $\lambda_2$  when  $h < b\alpha$ . Unless  $\alpha$  is close to unity  $\lambda_1$  will be the dominant latent root giving

$$h > a \quad (2.6)$$

or the fitness of the heterozygote  $Aa > AA$  the original wild homozygote for the new gene  $a$  to increase in frequency. This is the condition for a new gene to increase in frequency in a random mating population (Bodmer, 1960b; Bodmer and Parsons, 1960).

If  $\alpha$  is close to unity then  $\lambda_2$  may be the dominant latent root, so that for a new gene to increase

$$b\alpha > a \quad (2.7)$$

For complete positive assortative mating when  $\alpha = 1$  this becomes

$$b > a \quad (2.8)$$

or the fitness of the new homozygote  $aa > AA$  the old homozygote. This is the condition derived under a system of 100 per cent. selfing (Bodmer and Parsons, 1960). The other condition derived for complete selfing is  $h > 2a$ , which is not likely to be relevant since this implies that the heterozygote  $Aa$  must have a viability over twice that of the original homozygote  $AA$  for the introduction of  $a$ . Under complete assortative mating, however, the corresponding condition  $h > a$  is clearly likely to be more important as it merely implies heterozygote advantage over the original homozygote.

We can also examine the situation where gene  $A$  is introduced into a population initially all  $aa$ . This will be necessarily different from the case just considered since complete dominance of gene  $A$  over  $a$  is being assumed. Thus we assume  $u$  and  $v$  to be small, and ignoring second order terms in  $u$  and  $v$  we get:

$$\left. \begin{aligned} u' &= \frac{a\alpha}{2b}u + \frac{a\alpha}{4b}v \\ v' &= \frac{h(1-\alpha)}{b}u + \frac{\left(1 - \frac{\alpha}{2}\right)h}{b}v \end{aligned} \right\} \quad (2.9)$$

assuming  $\frac{u}{v}$  also to be small, which is reasonable since  $u$  represents the genotype  $AA$  which will be initially far less frequent than  $Aa$ . The characteristic equation formed from the  $2 \times 2$  matrix made up by these equations is

$$\lambda^2 - \lambda \left( \frac{a\alpha}{b} + h - \frac{h\alpha}{2} \right) + \frac{ha\alpha}{4b^2} = 0. \quad (2.10)$$

At  $\alpha = 0$ , or random mating  $h > b$ , or the fitness of the heterozygote  $Aa > aa$  the old homozygote as expected for the new gene  $A$  to increase in frequency. If  $\alpha = 1$ , then  $\lambda_1 = \frac{h}{2b}$ ,  $\lambda_2 = \frac{a}{2b}$ , so that  $\lambda_1$  will be the dominant latent root if  $h > a$  and  $\lambda_2$  if  $a > h$ . Thus either we must have extreme heterozygote advantage ( $Aa > 2aa$ ), or the new homozygote  $AA$  must be more than twice as fit as the old homozygote  $aa$ . The conditions therefore differ according to whether gene  $A$  or gene  $a$  is introduced which is due to the "asymmetry" of the problem because of the complete dominance of gene  $A$ .

Thus comparing the 2 situations:

	$\alpha = 0$	$\alpha = 1$	
gene $A$ introduced	$h > b$	$h > 2b$ or $a > 2b$	(2.11)
gene $a$ introduced	$h > a$	$h > a$ or $b > a$	

we see that for  $\alpha = 1$  the advantage of the heterozygote or new homozygote which is necessary is greater when gene  $A$  is being introduced. This is no doubt because when gene  $a$  is introduced it

is initially in an  $Aa$  genotype and can spread quite appreciably before the creation of  $aa$  genotypes with positive assortative mating helping its spread, whereas no such initial spread can occur when gene  $A$  is introduced.

Returning to equation (2.10), if we put  $\alpha$  small and ignore terms in  $\alpha^2$  we obtain

$$\lambda_D = \frac{1}{b} \left( \frac{a\alpha}{4} + h - \frac{h\alpha}{2} \right) \quad (2.12)$$

which is  $> 1$  when

$$\frac{\alpha}{2} \left( \frac{a}{2} - h \right) + h > b. \quad (2.13)$$

Thus for a small amount of assortative mating, it is possible to have  $h < b$  provided  $a > h$ , *i.e.* we may have some heterozygote disadvantage compared with the original homozygote provided that the new homozygote is fitter than the heterozygote. An example would be where  $h = 1$ ,  $a = 2.8$ ,  $\alpha = 0.10$  giving

$$\frac{\alpha}{2} \left( \frac{a}{2} - h \right) + h = 1.02$$

so that  $h$  can be  $< b$  and the inequality (2.13) satisfied. Thus, as with selfing, under a small amount of positive assortative mating, a gene may increase with no heterozygote advantage over the original homozygote.

To examine the situation for all values of  $\alpha$  we can put  $h = 1$  and  $a = 1 + a_1$  where  $a_1$  is small into equation (2.10) which gives two latent roots

$$\lambda_1 = \frac{1}{2b} \left[ 1 + \frac{a_1\alpha}{2} - \sqrt{1 - \alpha} \right]$$

and

$$\lambda_2 = \frac{1}{2b} \left[ 1 + \frac{a_1\alpha}{2} + \sqrt{1 - \alpha} \right]$$

The latter is the dominant latent root and is  $> 1$  when

$$1 + \frac{a_1\alpha}{2} + \sqrt{1 - \alpha} > 2b.$$

To compare the effects of different values of  $\alpha$ , values of  $b$  the viability of  $aa$  are given below for various values of  $a_1$  and  $\alpha$  which must not be exceeded for a new gene  $A$  to increase in frequency:

		$a_1$		
		-0.1	0	0.1
$\alpha$	0	1	1	1
	0.5	0.84	0.85	0.87
	1.0	0.47	0.50	0.52

Thus as expected, the fitter the new homozygote  $AA$ , the fitter can the original homozygote  $aa$  be for gene  $A$  to be introduced.

### 3. ASSORTATIVE MATING WITHOUT DOMINANCE

It is possible that gene  $A$  may not be completely dominant to  $a$ , and we will examine here the situation where each genotype has the same probability of mating with a like genotype. This system was considered from the point of view of equilibria without selection by O'Donald (1960). Let the genotypic proportions of  $AA$ ,  $Aa$  and  $aa$  be  $u$ ,  $v$  and  $w$  respectively, let  $a$ ,  $h$  and  $b$  be the relative viabilities of the three genotypes respectively and  $\alpha$  the probability of assortative mating for each genotype. We then get frequencies as follows:

Mating type	Frequency	Proportion of each genotype taking viabilities into account		
		$AA$	$Aa$	$aa$
$AA \times AA$	$\alpha u + (1-\alpha)u^2$	$a$	...	...
$AA \times Aa$	$2(1-\alpha)uw$	$\frac{1}{2}a$	$\frac{1}{2}h$	...
$AA \times aa$	$2(1-\alpha)uw$	...	$h$	...
$Aa \times Aa$	$\alpha v + (1-\alpha)v^2$	$\frac{1}{2}a$	$\frac{1}{2}h$	$\frac{1}{2}b$
$Aa \times aa$	$2(1-\alpha)vw$	...	$\frac{1}{2}h$	$\frac{1}{2}b$
$aa \times aa$	$\alpha w + (1-\alpha)w^2$	...	...	$b$

If we assume the population to be initially all  $AA$  and introduce gene  $a$ , then we can assume  $v$  and  $w$  to be small, and ignoring second order terms in  $v$  and  $w$  we obtain

$$\left. \begin{aligned} Tu' &= a[\alpha(1-\frac{1}{2}v-w) + (1-\alpha)(1-v-2w)] \\ Tv' &= h[(1-\alpha)(v+2w) + \frac{1}{2}\alpha v] \\ Tw' &= b\alpha(\frac{1}{2}v+w) \end{aligned} \right\} \quad (3.1)$$

Hence:

$$\left. \begin{aligned} v' &= \frac{(1-\frac{1}{2}\alpha)h}{a} v + \frac{2(1-\alpha)h}{a} w \\ w' &= \frac{b\alpha}{4a} v + \frac{b\alpha}{a} w \end{aligned} \right\} \quad (3.2)$$

The  $2 \times 2$  matrix formed by these equations gives the characteristic equation:

$$\lambda^2 - \lambda \left[ \frac{h}{a} (1 - \frac{1}{2}\alpha) + \frac{b\alpha}{a} \right] + \frac{hb\alpha}{2a^2} = 0. \quad (3.3)$$

At  $\alpha = 0$  or random mating this gives  $h > a$  when  $\lambda > 1$  as expected. To investigate the situation where  $\alpha$  is small let us write

$$\lambda = \frac{h}{a} + \alpha x$$

the terms in  $\alpha$  representing deviations from the root at  $\alpha = 0$ . Putting  $\lambda$  in the characteristic equation and assuming terms in  $\alpha^2$  to be negligible gives

$$x = \frac{b-h}{2a}$$

Therefore

$$\lambda = \frac{h}{a} \left[ 1 + \frac{\alpha(b-h)}{2h} \right]. \quad (3.4)$$

If  $b < h$  then  $\frac{\alpha(b-h)}{2h} < 0$ , so that for a new gene  $a$  to increase  $h > a$ .

If, however,  $b > h$  then it is possible to have  $h < a$  for the introduction of the new gene. Hence as in the case of assortative mating with dominance, a new gene can increase without heterozygote advantage over the original homozygote for  $\alpha > 0$ .

At  $\alpha = 1$  or 100 per cent. assortative mating the two roots of the characteristic equation (3.3) are

$$\frac{h}{2a} \text{ and } \frac{b}{a} \quad (3.5)$$

giving  $h > 2a$  and  $b > a$  when  $\lambda > 1$ , which are precisely the conditions for the increase of new genes under complete selfing. The latter inequality  $b > a$  is likely to be the more important as the former implies extreme heterozygote advantage.

In the case of assortative mating without dominance there will be no difference in the form of the result when gene  $A$  is introduced into a population originally all  $aa$ . The conditions can be obtained merely by substituting  $b$  for  $a$  and  $a$  for  $b$  in equation (3.3) and in solutions derived from this equation. Any form of dominance will destroy this "symmetry". In this paper we have examined the extremes of no dominance and of complete dominance. Intermediate degrees of dominance will no doubt lead to results intermediate between those obtained.

#### 4. DISCUSSION

The conditions for the initial progress of a newly occurring gene are clearly of critical importance since they will precede any conditions that may be necessary for the establishment of the gene in a polymorphism. In an outbreeding population it has been shown (Bodmer

and Parsons, 1960) that for a new gene to become established, initial heterozygote advantage over the prevailing homozygote is necessary, and this may be followed by the "evolution of overdominance" *i.e.* superiority of the heterozygote over both homozygotes (Parsons and Bodmer, 1961).

Under mixed positive assortative mating and random mating initial heterozygote advantage may not be necessary for a new gene to become established, and the viability of the new homozygote may become relevant. In this respect positive assortative mating is qualitatively analogous to mixed selfing and random mating although there may be some quantitative differences as shown in this paper. However, in the case of assortative mating the dominance relations of the two alleles will be relevant whereas under inbreeding this will not be so. This may lead to conditions differing according to the allele being introduced which is not the case under a mixed inbreeding and random mating system.

In some species different conditions of survival and reproduction may occur at opposite ends of their geographical range. Under such a situation it is probable that "certain of the genes which exist as alternatives will be favoured at one extreme, and will tend there to increase, while at the other extreme they will be disadvantageous and tend to diminish in frequency, the intermediate zones being divisible into a series of zones in which the advantage increases, from a negative value at one extreme, through zero at a region in which the selective advantage is exactly balanced, to a certain positive advantage at the other extreme" (Fisher, 1930). Such a situation may lead to an equilibrium, and the longer it persists, the more numerous will the genetic differences between the types inhabiting the extreme regions become. This may lead to a tendency for assortative mating between like genotypes. If new genes arise they will then not necessarily need to be advantageous in the heterozygote to increase and may tend to increase especially if they happen to be linked to the main locus, or loci differentiating the two extreme populations.

Such a situation of two extremes favoured within a population has been termed disruptive selection, and the outcome can be a stable polymorphism (Mather, 1955; Thoday and Boam, 1959). It would be of interest to ascertain whether mating preferences evolve during the production of a polymorphism produced experimentally by means of disruptive selection. In the experiments of Thoday and Boam which were done with *Drosophila melanogaster* gene flow between the two halves of the population was high, namely 50 per cent., so that there was no spatial isolation of the type which may occur in many polymorphic clines such as described by O'Donald for the Arctic Skua. Fifty per cent. gene flow is at least twice as much as is usually considered in natural populations, since random mating involves a maximum of 25 per cent. gene flow. A further experiment



with 25 per cent. gene flow with positive assortative mating gave divergence considerably more rapidly, whereas under negative assortative mating this divergence was reduced (Millicent and Thoday, 1960, 1961). Thus as expected, artificial positive assortative mating will aid divergence, but the opposite question of whether divergence within a population will induce positive assortative mating is perhaps more difficult to decide, although it seems likely on the basis of the evidence quoted in the introduction.

## 5. SUMMARY

The conditions under which new genes will increase in frequency under mixed positive assortative mating and random mating are discussed for populations in which two alleles are segregating at a locus. Whereas heterozygote advantage is essential under random mating for a new gene to become established, this is no longer a prerequisite with a proportion of positive assortative mating. Generally, the results are analogous to a system of mixed selfing and random mating, but the dominance relations of the two alleles are relevant under assortative mating. A system of complete dominance of gene *A* over *a* is considered and conditions differ according to the gene being introduced, but if the probability of assortative mating is identical for each genotype the model becomes "symmetrical". The possibility of evolving mating preferences during the production of a polymorphism by disruptive selection is discussed.

## 6. REFERENCES

- BODMER, W. F. 1960a. The genetics of homostyly in populations of *Primula vulgaris*. *Phil. Trans. Roy. Soc. B*, 242, 517-549.
- BODMER, W. F. 1960b. Discrete stochastic processes in population genetics. *J. Roy. Stat. Soc. B*, 22, 218-244.
- BODMER, W. F., AND PARSONS, P. A. 1960. The initial progress of new genes with various genetic systems. *Heredity*, 15, 283-299.
- BODMER, W. F., AND PARSONS, P. A. 1962. Linkage and recombination in evolution. *Advances in Genetics* (in press).
- COOCH, F. G., AND BEARDMORE, J. A. 1959. Assortative mating and reciprocal differences in the Blue-Snow Goose complex. *Nature*, 183, 1833-1834.
- FISHER, R. A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edin.*, 52, 399-433.
- FISHER, R. A. 1930. *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.
- JENNINGS, H. S. 1916. The numerical results of diverse systems of breeding. *Genetics*, 1, 53-89.
- MATHER, K. 1955. Polymorphism as an outcome of disruptive selection. *Evolution*, 9, 52-61.
- MATHER, K., AND HARRISON, B. J. 1949. The manifold effect of selection. *Heredity*, 3, 131-162.
- MERRELL, D. J. 1949. Selective mating in *Drosophila melanogaster*. *Genetics*, 34, 370-389.
- MILLICENT, E., AND THODAY, J. M. 1960. Gene flow and divergence under disruptive selection. *Science*, 131, 1311-1312.

- MILLICENT, E., AND THODAY, J. M. 1961. Effects of disruptive selection. IV. Gene-flow and divergence. *Heredity*, 16, 199-217.
- O'DONALD, P. 1959. Possibility of assortive mating in the Arctic Skua. *Nature*, 183, 1210-1211.
- O'DONALD, P. 1960. Assortive mating in a population in which two alleles are segregating. *Heredity*, 15, 389-396.
- PARSONS, P. A. 1961a. The initial progress of new genes with viability differences between sexes and with sex linkage. *Heredity*, 16, 103-107.
- PARSONS, P. A. 1961b. The establishment of new genes in populations under various genetic systems. *Proc. 2nd Int. Conference of Human Genetics*.
- PARSONS, P. A., AND BODMER, W. F. 1961. The evolution of overdominance: natural selection and heterozygote advantage. *Nature*, 190, 7-12.
- PEARSON, K., AND LEE, A. 1903. On the laws of inheritance in man. I. Inheritance of physical characters. *Biometrika*, 2, 357-462.
- RENDEL, J. M. 1944. Genetics and cytology of *Drosophila subobscura*. II. Normal and selective matings in *Drosophila subobscura*. *J. Genet.*, 46, 287-302.
- RENDEL, J. M. 1951. Mating of ebony vestigial and wild type *Drosophila melanogaster* in light and dark. *Evolution*, 5, 226-230.
- THODAY, J. M., AND BOAM, T. B. 1959. Effects of disruptive selection. II. Polymorphism and divergence without isolation. *Heredity*, 13, 205-218.
- WRIGHT, S. 1921. Systems of mating. III. Assortative mating based on somatic resemblance. *Genetics*, 6, 144-161.