

# SELFING UNDER CONDITIONS FAVOURING HETEROZYGOSITY

P. A. PARSONS

*Department of Genetics, University of Cambridge*

Received 16.iii.57

## 1. INTRODUCTION

A LOCUS may be maintained in a heterozygous state experimentally, or by various outbreeding devices such as heterostyly or self-sterility. If one or both of the homozygotes are of lowered viability when compared with the heterozygotes, progress towards homozygosity under any inbreeding programme will be slower. This effect will be enhanced by the unconscious selection of the fitter heterozygotes to continue such a programme. The extreme case is where either or both of the homozygotes are lethal. Bennett (1956) has examined the situation where one of the homozygotes is lethal. Bartlett and Haldane (1935) have examined the progress towards homozygosity of a locus linked to loci maintained permanently heterozygous (*i.e.* both homozygotes are lethal). Hayman and Mather (1953) have discussed various situations where the homozygotes are at a disadvantage.

In the first part of this paper (section 2), using selfing as the inbreeding system, the progress towards homozygosity for a locus linked to a locus maintained permanently heterozygous will be considered, and extended to the case of a locus situated in between two such loci maintained in a permanently heterozygous condition. In the second part of this paper (section 3) selfing of tetrasomics where the homozygotes are at a disadvantage (neglecting linkage) will be considered.

Any delay due to the homozygotes maturing slower, or of unconscious selection for the more vigorous heterozygotes will be ignored. This is a problem in the setting up of inbred lines and will have the effect of slowing the approach of a line to homozygosity. We shall assume that a single mating will continue an inbreeding system.

## 2. INBREEDING OF A LOCUS LINKED TO LOCI MAINTAINED PERMANENTLY HETEROZYGOUS

Here we shall consider the approach to homozygosity of a locus *B* linked to locus *A* with recombination value  $\gamma$ , in which only the heterozygotes *Aa* are viable. This situation was discussed very briefly by Bartlett and Haldane (1935) and it is proposed to extend it here.

There are four possible genotypes :— $AB/ab$ ,  $AB/aB$ ,  $Ab/ab$  and

*Ab/aB*. Denoting their initial frequencies by  $u_0$ ,  $v_0$ ,  $w_0$  and  $x_0$  respectively, the generation matrix for selfing is:—

	$u_0$	$v_0$	$w_0$	$x_0$
$u_1$	$(1-y)^2$			$y^2$
$v_1$	$y(1-y)$	1		$y(1-y)$
$w_1$	$y(1-y)$		1	$y(1-y)$
$x_1$	$y^2$			$(1-y)^2$
Column divisor	1	1	1	1

If  $\lambda$  represents the latent roots of this matrix, an equation in  $\lambda$  (the characteristic equation) may be obtained by equating the determinant of the  $\lambda$  matrix to zero:—

$$[(1-y)^2 - \lambda]^2 (1-\lambda)^2 - (1-\lambda)^2 y^4 = 0$$

This equation yields two roots of unity which are trivial as they represent the homozygous genotypes, and a dominant latent root

$$\lambda_D = 1 - 2y + 2y^2$$

and also a fourth root

$$\lambda = 1 - 2y.$$

In terms of the recombination value  $y$ ,  $\lambda_D$  takes the following values (table 1):—

TABLE 1

$y$ per cent.	$\lambda$
0	1
0.1	0.9980
0.5	0.9901
1.0	0.9802
5.0	0.9050
10.0	0.8200
25.0	0.6250
50.0	0.5000

For two closely linked loci, locus *B* will approach homozygosity very slowly. For a single gene where one homozygote is lethal,  $\lambda_D = \frac{2}{3}$  (Bennett, 1956). In the situation discussed here  $\lambda_D \geq \frac{2}{3}$  for  $y \leq 21.13$  per cent. As  $y \rightarrow 50$  per cent.,  $\lambda_D \rightarrow \frac{1}{2}$  which is the dominant latent root obtained when a normal locus, say *B*, unaffected by the presence of *A* is inbred by selfing.

Table 1 is only true if the homozygotes and heterozygotes at the *B* locus are equally viable. If the homozygotes are less viable,  $\lambda_D$  will be somewhat greater, and progress towards homozygosity retarded. Considering *B* only, and giving viability factors  $m$  and  $n$  to *BB* and

*bb* respectively, and to *Bb* a factor of unity, the generation matrix for selfing is :—

	<i>B</i>	<i>Bb</i>	<i>bb</i>
<i>BB</i>	<i>m</i>	<i>n</i>	<i>m</i>
<i>Bb</i>		<i>n</i>	<i>n</i>
<i>bb</i>			<i>2</i>
Column divisors	<i>m</i>	<i>n</i>	<i>m+n+2</i>

The characteristic equation is :

$$(m - m\lambda)(n - n\lambda)[2 - (2 + m + n)\lambda] = 0$$

yielding a latent root of unity and the dominant latent root

$$\lambda_D = \frac{2}{2 + m + n}$$

Provided that both homozygotes are not lethal,  $\lambda_D < 1$  and there will be some progress towards homozygosity. If  $m = n = 1$ ,  $\lambda_D = \frac{1}{2}$  which is the latent root obtained from selfing a disomic without any complications.

Returning to the case of *B* linked to *A* which is permanently heterozygous, we give *BB*, *bb*, and *Bb* viabilities of *m*, *n*, and 1 respectively. The generation matrix will become :—

	$u_0$	$v_0$	$w_0$	$x_0$
$u_1$	$(1-y)^2$	...	...	$y^2$
$v_1$	$y(1-y)m$	<i>m</i>	...	$y(1-y)m$
$w_1$	$y(1-y)n$	...	<i>n</i>	$y(1-y)n$
$x_1$	$y^2$	...	...	$(1-y)^2$
Column divisors	$\frac{1-2y+2y^2+y(1-y)(m+n)}{y(1-y)(m+n)}$	<i>m</i>	<i>n</i>	$\frac{1-2y+2y^2+y(1-y)(m+n)}{y(1-y)(m+n)}$

The latent roots are :—

$$\lambda = 1$$

$$\lambda = \frac{1-2y}{1-2y+2y^2+y(1-y)(m+n)}$$

and

$$\lambda_D = \frac{1-2y+2y^2}{1-2y+2y^2+y(1-y)(m+n)}$$

The last two roots reduce to those given for the case of locus *B* free from viability disturbance when  $m = n = 1$ . When  $m, n < 1$ ,  $\lambda_D$  will increase to a limit of unity at  $m = n = 0$ . Poor viability of *bb* and *BB* will result in slower progress to homozygosity. If selection

for the more vigorous heterozygotes, or difficulties in maintaining the inbreeding programme are taken into account, progress will be reduced more.

An application of this system is found in *Oenothera* where some species are permanently heterozygous for a series of interchanges (Darlington, 1931; Catcheside, 1940). In these species crossing-over is suppressed in the differential segments of the interchanges and different mutations accumulate. Hence the segments no longer correspond and two kinds of gamete are formed. When similar gametes unite the zygote is lethal, due to recessive lethal factors. Normal crossing-over occurs in the pairing segments in the distal regions of chromosome arms. The preceding analysis would apply to these regions, although some rare cross-overs occur in the differential segments.

Bartlett and Haldane (1935) have considered in some detail the problem of sib-mating which will not be discussed here. The result is similar to that for selfing. The dominant latent root drops from unity at complete linkage to 0.8090 at independence, this being the root for ordinary sib-mating.

A more complex case in which there are two linked loci *A* and *C*, with *AA*, *aa*, *CC*, and *cc* lethal, will now be considered. It is required to find the progress towards homozygosity of a locus *B* situated between *A* and *C*. Let *p* represent no recombination in segments *AB* and *BC*, *q* represent recombination in segment *AB* but not *BC*, *r* represent recombination in segment *BC* but not *AB*, and *s* represent simultaneous recombination in both segments such that

$$p + q + r + s = 1.$$

There are four genotypes with *B* heterozygous.

$$\begin{array}{ll} u_0 & ABC/abc \\ v_0 & aBC/Abc \\ w_0 & ABc/abC \\ x_0 & AbC/aBc \end{array}$$

The generation matrix is:—

	$u_0$	$v_0$	$w_0$	$x_0$	
$u_1$	$p^2$	$r^2$	$q^2$	$s^2$	} $\div (p+s)^2 + (q+r)^2$
$v_1$	$r^2$	$p^2$	$s^2$	$q^2$	
$w_1$	$q^2$	$s^2$	$p^2$	$r^2$	
$x_1$	$s^2$	$q^2$	$r^2$	$p^2$	

It will be noted that this matrix is a Latin square of the system

$$\begin{array}{cccc} A & B & C & D \\ B & A & D & C \\ C & D & A & B \\ D & C & B & A \end{array}$$

The determinant of such a matrix is :—

$$(A+B+C+D)(A-B+C-D)(A+B-C-D)(A-B-C+D).$$

Equating the determinant of the  $\lambda$ -matrix to zero we obtain the dominant latent root :—

$$\lambda_1 = \frac{p^2+q^2+r^2+s^2}{(p+s)^2+(q+r)^2} = 1 - \frac{2(ps+qr)}{(p+s)^2+(q+r)^2}$$

and the other three roots :—

$$\lambda_2 = 1 - \frac{2(q^2+s^2+ps+qr)}{(p+s)^2+(q+r)^2}$$

$$\lambda_3 = 1 - \frac{2(r^2+s^2+ps+qr)}{(p+s)^2+(q+r)^2}$$

$$\lambda_4 = 1 - \frac{2(q^2+r^2+ps+qr)}{(p+s)^2+(q+r)^2}$$

It is obvious that  $\lambda_1 > \lambda_2, \lambda_3, \lambda_4$  for all  $p, q, r$  and  $s$ . For a short segment where  $q$  and  $r$  are small and  $s$  is close to zero,  $\lambda_1$  will be very close to unity. This somewhat hypothetical system can be regarded as the severest possible case of inbreeding where homozygotes  $AA, aa, CC$  and  $cc$  are of very much poorer viability than the heterozygote.

An application of this type of argument is found when heterozygous inversions are considered. In both paracentric and pericentric inversions, the only surviving gametes are the non cross-over and double cross-over gametes. All single cross-over gametes are lost as they are acentric or dicentric in paracentric inversions, and duplication-deficient in pericentric inversions. Thus we obtain the following generation matrix for the surviving gametes :—

	$u_0$	$v_0$	$w_0$	$x_0$	
$u_1$	$p^2$	...	...	$s^2$	} $\div (p+s)^2$
$v_1$	...	$p^2$	$s^2$	...	
$w_1$	...	$s^2$	$p^2$	...	
$x_1$	$s^2$	...	...	$p^2$	

The dominant latent root is :—

$$\lambda_D = 1 - \frac{2ps}{(p+s)^2}$$

where  $p+s$  is unity as  $q = r = 0$ . Hence  $\lambda_D = 1 - 2ps = 1 - 2s$  when  $s$  is small. Within a short inversion, a heterozygous locus will approach homozygosity very slowly. As an example, let  $p = 0.60, q = 0.18, r = 0.18$  and  $s = 0.04$ , the sum adding to unity. The value for  $s$  is rather large for the values of  $q$  and  $r$  considered. In a heterozygous inversion  $q = r = 0$  and  $p+s = 1$ . Thus  $p$  will be  $\frac{0.60}{0.64} = 0.9375$  and  $s = \frac{0.04}{0.64} = 0.0625$  which gives  $\lambda_D = 0.8828$ .

This is a segment of considerable length and progress towards homozygosity is slow. For  $p = 0.90$ ,  $q = 0.05$ ,  $r = 0.047$  and  $s = 0.003$  which is a fairly realistic situation, we obtain  $\lambda_D = 0.9934$ . Hence for a short inversion heterozygote progress towards homozygosity is virtually non-existent. It is known that inversions preserve favourable gene complexes. This analysis confirms this, for even in the most severe form of inbreeding, namely selfing, little progress towards homozygosity is made for heterozygous inversions of short or medium length. Long inversions, which are occasionally found, where most of the chromosome is inverted, will not, however, preserve favourable gene complexes so readily. This analysis neglects any inhibition of crossing-over near the break points which would reduce progress to homozygosity even more.

### 3. TETRASOMICS

Selfing of tetrasomics in which homozygotes are at a disadvantage, or lethal, will now be considered. Linkage and double reduction will be ignored. Firstly, a factor  $A$  will be considered in which the nulliplex genotype  $a_4$  is lethal and the other four genotypes  $A_4$ ,  $A_3a$ ,  $A_2a_2$  and  $Aa_3$  are normal. The generation matrix is:—

	$u_0$ ( $A_3$ )	$v_0$ ( $A_3a$ )	$w_0$ ( $A_2a_2$ )	$x_0$ ( $Aa_3$ )
$u_1$	1	1	1	...
$v_1$	...	2	8	...
$w_1$	...	1	18	1
$x_1$	...	...	8	2
Column divisors	1	4	35	3

The characteristic equation is :

$$(1-\lambda)(210\lambda^3-353\lambda^2+168\lambda-20) = 0 \quad . \quad . \quad (1)$$

which yields four latent roots :

$$\begin{aligned} &1 \\ &0.932,717 \text{ the dominant root} \\ &0.568,683 \\ &\text{and } 0.179,552. \end{aligned}$$

To find the principal component of frequency  $P$  the coefficients of  $a$ ,  $b$ ,  $c$  in

$$P = av + bw + cx$$

are needed, such that  $P^1 = \lambda P$  where  $P^1$  represents the next generation. Substituting the value of  $\lambda_D$  in the  $\lambda$ -matrix and treating the columns

of the matrix as linear functions of the coefficients  $a$ ,  $b$ ,  $c$  we obtain the principal component

$$P = v + 1.730,869 w + 2.168,596 x$$

the coefficients of  $v$ ,  $w$  and  $x$  being the relative complexities (Fisher, 1949).

The limiting frequencies are found by treating the rows of the  $\lambda$ -matrix as equations in  $v$ ,  $w$  and  $x$ . We obtain :

$$\begin{aligned} v &= 0.221,259 \\ w &= 0.418,874 \\ x &= 0.359,866 \end{aligned}$$

the sum being unity.

Substituting these limiting frequencies in the linear function  $P$  gives the average relative complexity = 1.726,681. Dividing  $P$  by this value will give the three absolute complexities :

$$\begin{aligned} V &= 0.579,146 \\ W &= 1.002,426 \\ X &= 1.255,933 \end{aligned}$$

As a check, the sum of the products of the frequencies with the absolute complexities should be unity (Fisher, 1949), *i.e.*

$$vV + wW + xX = 1.$$

Progress towards homozygosity for any mating type depends upon its absolute complexity. Comparing the three heterogenic mating types with an absolute complexity of unity as standard (Fisher, 1949) we find that  $u$  is 7.842 generations ahead of the standard, and  $v$  and  $w$  are 0.034 and 3.271 generations behind the standard of unity respectively. Hence mating type  $Aa_3$  is in the worst condition and mating type  $A_3a$  in the best when considering their approach to homozygosity.

This system may be compared with selfing of tetrasomics without disturbance for which  $\lambda_D = \frac{5}{8}$  (neglecting double reduction). One generation of selfing of normal tetrasomics is equivalent to

$$(\log_e 0.8333)/(\log_e 0.9327) = 2.618$$

generations of selfing of tetrasomics with  $a_4$  lethal. The presence of this lethal genotype therefore retards progress towards homozygosity considerably. For disomics Bennett (1956) gives  $\lambda_D = \frac{2}{3}$  for the situation in which one of the two homozygotes is lethal. One generation of selfing of normal disomics is equivalent to

$$(\log_e 0.5000)/(\log_e 0.6667) = 1.709$$

generations of selfing of disomics with one homozygote lethal. Hence, lethality of one of the homozygotes in tetrasomics reduces the progress of inbreeding more than for disomics.

A second situation in tetrasomics where homozygotes are at a

disadvantage will now be considered. Let the duplex genotypes have a viability of unity, simplex and triplex  $1-t$ , and nulliplex and quadriplex  $1-2t$ . The generation matrix is :---

	$A_4$	$A_3a$	$A_2a_2$	$Aa_3$	$a_4$
$A_4$	$1-2t$	$1-2t$	$1-2t$	...	...
$A_3a$	...	$2(1-t)$	$8(1-t)$	...	...
$A_2a_2$	...	1	18	1	...
$Aa_3$	...	...	$8(1-t)$	$2(1-t)$	...
$a_4$	...	...	$1-2t$	$1-2t$	$1-2t$
Column divisor	$1-2t$	$4-4t$	$36-20t$	$4-4t$	$1-2t$

Isomorphic pairs  $(A_4, a_4)$  and  $(A_3a, Aa_3)$  may be combined to obtain :

	$A_4$	$A_3a$	$A_2a_2$
$A_4$	$1-2t$	$1-2t$	$1-2t$
$A_3a$	...	$2(1-t)$	$8(1-t)$
$A_2a_2$	...	1	9
Column divisor	$1-2t$	$4-4t$	$18-10t$

A latent root  $\lambda = 1$  is obtained, and two roots from

$$\lambda^2(36-20t) - \lambda(36-10t) + 5 = 0 \quad (2)$$

In terms of  $t$ , the dominant latent root takes the following values.

$t$	$\lambda$
0	0.8333
0.01	0.8357
0.05	0.8453
0.1	0.8580
0.25	0.9018
0.5	1.0000

At  $t = 0$  the characteristic equation gives the two roots  $\frac{1}{6}$  and  $\frac{5}{6}$ , the values obtained for selfing without disturbance. At  $t = 0.5$  the two roots are 1 and  $\frac{5}{28}$  at which stage the homozygotes  $A_4$  and  $a_4$  will be lethal, and progress towards homozygosity halted. Various other systems where heterozygotes are at an advantage can be postulated, but the results should conform in a general way to those given.

This type of problem may be generalised for all cases of variable viabilities and lethality (zero viability). Let the viabilities of  $A_4$ ,



$A_3a, A_2a_2, Aa_3$  and  $a_4$  be  $a, b, c, d$  and  $e$  respectively. Then the generation matrix becomes :

	$A_4$	$A_3a$	$A_2a_2$	$Aa_3$	$a_4$
$A_4$	$a$	$a$	$a$	...	...
$A_3a$	...	$2b$	$8b$	...	...
$A_2a_2$	...	$c$	$18c$	$c$	...
$Aa_3$	...	...	$8d$	$2d$	...
$a_4$	...	...	$e$	$e$	$e$
Column divisor	$a$	$a+2b+c$	$a+8b+18c+8d+e$	$c+2d+e$	$e$

This yields two latent roots of unity and three from the equation :

$$[2b-\lambda(a+2b+c)][18c-\lambda(a+8b+18c+8d+e)] \\ [2d-\lambda(1+2d+e)]-8cd[2b-\lambda(a+2b+c)] \\ -8bc[2d-\lambda(c+2d+e)] = 0.$$

If we let  $a = e = 1-2t, b = d = 1-t$  and  $c = 1$  which is the viability problem discussed above, the root  $\lambda = \frac{1}{2}$  is obtained, and the equation

$$\lambda^2(36-20t)-\lambda(36-10t)+5 = 0$$

from which  $\lambda_D$  is obtained. This is the equation (2) already given for this problem.

Similarly, considering the situation where  $a_4$  is lethal, we can let  $a = b = c = d = 1$  and  $e = 1$ . The general equation reduces to

$$(1-\lambda)(210\lambda^3-353\lambda^2+168\lambda-20) = 0$$

which is the equation (1) given for this problem.

In conclusion, as for disomics, progress towards homozygosity will be slower if the heterozygotes are at an advantage. In the severest case, where one or both of the homozygotes are lethal for inbreeding progress will be very slow, or even halted.

#### 4. DISCUSSION

Although some of the examples selected may appear to be of little direct application, it is evident from this analysis, and from those of other authors, that the maintenance of heterozygosity in any inbreeding system (selfing is used here for simplicity) will slow down the rate of approach to homozygosity. Where homozygotes are of poor viability, the progress towards homozygosity will be retarded, this being manifest by a higher value of  $\lambda_D$  than normal. The effect of enforced heterozygosis does not only affect the locus concerned, but also affects neighbouring loci dependent upon the recombination value with the locus at which homozygotes are disadvantageous. In the extreme case, where both homozygotes are lethal, as in some species of *Cenothera*,

$\lambda_D$  for that locus is unity, and we can only study the approach to homozygosity of a linked locus.

Progress towards homozygosity within a short inversion heterozygote is almost negligible, and here we have enforced heterozygosity of the severest form. The stability of the gene complex within an inversion allows favourable gene complexes to be preserved. In some species of *Drosophila*, notably *D. willistoni*, heterozygous inversions are particularly common (Dobzhansky, Burla and Da Cunha, 1950). The frequency of specific inversion heterozygotes appears to vary according to environment. There are regular seasonal cycles in the frequency of various inversions, these cycles being repeated annually (Dobzhansky, 1943). Such studies show beyond doubt that inversion heterozygotes are of advantage, and the present analysis shows the difficulty of inducing any change within the inversion heterozygote for the severest possible form of inbreeding. Recombination is also restricted to a certain extent outside an inversion heterozygote (Dobzhansky and Epling, 1948), so preserving additional gene complexes.

It is likely that favourable gene complexes could be preserved to a certain degree if such a complex were situated between two relatively closely linked loci for which the heterozygotes are at a considerable advantage compared with homozygotes. The extreme case, where homozygotes are lethal, is analysed in this paper. In this extreme case  $\lambda_D$  is close to unity for short segments and progress towards homozygosity is very slow.

For the selfing of tetrasomics,  $\lambda_D$  is close to unity for loci where the homozygotes are near to lethality. Such a locus will enforce a certain degree of heterozygosity on linked loci, the magnitude of the enforced heterozygosity decreasing with distance. Progress towards homozygosity for tetrasomics is generally slower than for disomics. Lethal genes in tetrasomics reduce the approach to homozygosity considerably and enforce some degree of heterozygosity. A measure of the rate of elimination of the lethal gene, in the situation where  $a_4$  is lethal is given by  $-\log_e \lambda_D = 0.070$  compared with  $-\log_e(\frac{2}{3}) = 0.405$  (Bennett, 1950) for disomics with one homozygote lethal. Tetraploids therefore shelter lethal genes, these in turn cause enforced heterozygosity.

Generally, therefore, the presence of lethals, or of deleterious homozygotes will slow down any theoretical inbreeding system in practice, and will result in enforced heterozygosity. Furthermore, this applies to linked but unaffected loci, the extreme case being a locus situated in an inversion heterozygote.

Throughout this discussion, delay due to inadequate or slowly maturing progeny has been ignored. It is also assumed that there is no unconscious selection for the more vigorous heterozygotes. If these considerations had been taken into account, the progress of inbreeding would be slower.

## 5. SUMMARY

The progress towards homozygosity is considered for situations where homozygotes are at a disadvantage or lethal. Selfing, although impossible in some organisms, provides a basis for comparison and has the advantage of relatively simple algebra. Matrix methods are used in the calculations.

Where homozygotes are of poor viability the approach to homozygosity is slower than for homozygotes and heterozygotes of the same viability. Loci linked to those in which heterozygosis is favoured approach homozygosity slower than if they were independent, the extreme case being the heterozygous inversion.

*Acknowledgment.*—I would like to thank Dr A. R. G. Owen for many helpful discussions.

## 6. REFERENCES

- BARTLETT, M. S., AND HALDANE, J. B. S. 1935. The theory of inbreeding with forced heterozygosis. *J. Genet.*, 31, 327-340.
- BENNETT, J. H. 1956. Lethal genes in inbred lines. *Heredity*, 10, 263-270.
- CATCHESIDE, D. G. 1940. Structural analysis of *Cenothera* complexes. *P.R.S.*, B, 128, 509-535.
- DARLINGTON, C. D. 1931. The cytological theory of inheritance in *Cenothera*. *J. Genet.*, 24, 405-474.
- DOBZHANSKY, TH. 1943. Genetics of natural populations IX. Temporal changes in the composition of populations of *Drosophila pseudoobscura*. *Genetics*, 28, 162-186.
- DOBZHANSKY, TH., BURLA, H., AND DA CUNHA, A. B. 1950. A comparative study of chromosomal polymorphism in sibling species of the *willistoni* group of *Drosophila*. *Amer. Nat.*, 84, 229-246.
- DOBZHANSKY, TH., AND EPLING, C. 1948. The suppression of crossing-over in inversion heterozygotes of *Drosophila pseudoobscura*. *Proc. Nat. Acad. Sci.*, 34, 137-141.
- FISHER, R. A. 1949. *The Theory of Inbreeding*. Oliver and Boyd, Edinburgh.
- HAYMAN, B. I., AND MATHER, K. 1953. The progress of inbreeding when homozygotes are at a disadvantage. *Heredity*, 7, 165-183.