The Genetics of Homostyly in Populations of Primula vulgaris

W. F. Bodmer

Phil. Trans. R. Soc. Lond. B 1960 **242**, 517-549
doi: 10.1098/rstb.1960.0001

THE GENETICS OF HOMOSTYLY IN POPULATIONS OF *PRIMULA VULGARIS*

By W. F. BODMER

Department of Genetics, University of Cambridge

(*Communicated by Sir Ronald Fisher, F.R.S.—Received* 3 *April* 1959—*Revised* 16 *July* 1959)

CONTENTS

Natural populations of primroses containing homostyles were found by Crosby (1940) near Sparkford, Somerset, and in the Chilterns. Homostyle frequencies up to 80% have often been recorded, but few populations containing higher frequencies have been found. In explaining observed configurations of homostyle frequencies Crosby (1949) assumed that they always self-fertilized. Bodmer (1958), however, obtained evidence for the occurrence of crossing between the homostyles under natural conditions. The estimated amount of crossing increases with lower viability for the homostyles and the data indicate a frequency of crossing of at least 80%.

Analysis of counts of populations containing homostyles in Somerset made by Professor Sir Ronald Fisher, F.R.S., over the years 1941–54 shows heterogeneity within populations, but no evidence of an upward trend in homostyle frequency. The only significant change detected is a downward trend in the homostyle frequency. There is also a marked deficiency of thrums relative to pins for high homostyle frequencies.

Investigation of theoretical models which take into account crossing between homostyles shows that the homostyles will increase in frequency so long as there are no viability disturbances. However, a method for calculating an effective initial selective value for the homostyle gene shows that when the frequency of crossing is high only moderate viability disturbances are needed to prevent the homostyles from increasing in frequency. Numerical investigation of the models indicates that only a slight decrease in viability is needed to decrease the proportion of homostyles in populations where they are already abundant. Moreover, if the decreased viability affects particularly the mature plants it is possible to account for the thrum deficiency and also a possible very rapid decrease in homostyle frequency.

With high initial selective values for the homostyle gene, such as are obtained when the homostyles only self-fertilize, even when they have considerably reduced viability, only a few occurrences are needed to make them almost certain to gain a hold and increase in frequency. It does not seem possible, therefore, to explain why the homostyles have occurred only in a few restricted areas, unless the effective initial selection for them is very small or negative, which can only occur when there is a considerable amount of crossing between the homostyles. It thus seems that the tendency of the homostyles, and possibly also of pins and thrums, to cross-fertilize is an essential buffering mechanism against the possible reversal of an outcrossing system to inbreeding.

1. Introduction

In natural populations of *Primula vulgaris*, only the pin and thrum forms are commonly found. However, naturally occurring homostyles, having both stigma and anthers at the high level, have been found in certain populations near Sparkford, Somerset, and in the Chilterns (Crosby 1940, 1949). The compatibility relationships of the three forms of flower and their genotypes are illustrated in figure 1. The arrows connecting anthers and stigmas indicate the only possible compatible fertilizations. If the homostyles are self-pollinated and are at no disadvantage relative to the pins and thrums, with respect to their fitness and fertility, it is easy to see that their fertility relationships place them at an unconditional advantage over the pins and thrums. They may thus be expected to increase in frequency, completely replacing the pins and thrums. An analysis of this situation has been given by Fisher (1949), who considered in detail the rate and progress of approach to equilibrium.

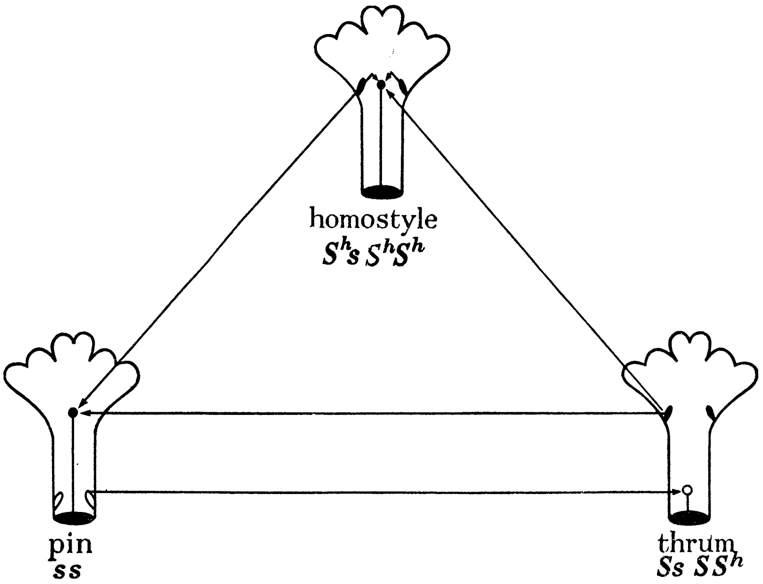

Figure 1. Compatibility relationships and genotypes of the three forms of flower of *Primula vulgaris*.

In counts on natural populations homostyle frequencies up to 80 % have often been recorded, but few populations containing higher frequencies have been found. It is therefore clear that some modification of the model examined by Fisher (1949) is necessary to explain the observed proportions. Crosby (1949) examined, in detail, the possible explanations on the assumption that all the homostyles self-fertilize. The excess of pins over thrums in populations containing no homostyles, led him to assume that approximately 10 % of pins self-fertilize illegitimately under natural conditions. In order then to obtain a final equilibrium of about 80 % homostyles and 20 % pins, the thrums being absent from populations with high homostyle counts, he had to assign a viability of 65 % to the homozygous homostyles (S^hS^h) relative to the other genotypes. Crosby showed that even if the S^hS^h homostyles had a relative viability of 81·5 % the homostyles would still reach 100 %.

Bodmer (1958) presented evidence from open-pollination data and observations of homostyle flowers during growth, for the occurrence of crossing between the homostyles under natural conditions. It is the purpose of this paper to consider the dynamics of the changes in the *Primula* populations taking into account the possibility of cross-fertilization between the homostyles.

2. A DETAILED ANALYSIS OF THE OPEN-POLLINATION DATA

Bodmer (1958) obtained an estimate of crossing from the open pollination data assuming that the ratio of the genotypes $\boldsymbol{S^hs}:\boldsymbol{S^hS^h}$ was constant at 2:1 and that there were no viability disturbances. He showed that if s is the proportion of homostyles which result from self-fertilization and the frequencies with which the genotypes $\boldsymbol{S^hs}$ and $\boldsymbol{S^hS^h}$ are present among the parents are in the ratio $v/(1-v)$, then the proportion of pins in segregating progenies is $\frac{1}{4}(s+v(1-s))$. As s decreases from 1 to 0, the proportion of pins decreases from $1/4$ to $v/4$. However, cross-fertilization affects the proportion v, which does not therefore remain constant.

The technique of growing the progenies of homostyles from segregating progenies implies in fact a progress to an equilibrium in the proportion of pins and of segregating progenies. This equilibrium will depend on the viability disturbances and the amount of crossing.

Let the frequency of the genotypes $\boldsymbol{ss}$, $\boldsymbol{S^hs}$, $\boldsymbol{S^hS^h}$ in any one generation be u, w and z, respectively, and their survival values from zygote to maturity, or relative viabilities be 1, c and e. The expected progenies of the segregating homostyles will be made up as follows:

mating	contribution
$\boldsymbol{S^hs}$ selfed	$sw[\frac{1}{4}\boldsymbol{ss}+\frac{1}{2}c\boldsymbol{S^hs}+\frac{1}{4}e\,\boldsymbol{S^hS^h}]$
$\boldsymbol{S^hs}\times\boldsymbol{S^hs}$	$\dfrac{(1-s)\,w^2}{w+z}[\frac{1}{4}\boldsymbol{ss}+\frac{1}{2}c\,\boldsymbol{S^hs}+\frac{1}{4}e\,\boldsymbol{S^hS^h}]$
$\boldsymbol{S^hs}\times\boldsymbol{S^hS^h}$	$\dfrac{(1-s)\,zw}{w+z}[\frac{1}{2}c\,\boldsymbol{S^hs}+\frac{1}{2}e\,\boldsymbol{S^hS^h}]$

where s is the proportion of selfing and $u+w+z=1$. Hence the expected proportions u', w', z' in the following generation will be given by

$$\left.\begin{aligned} Tu' &= \frac{s}{4}+\frac{(1-s)\,w}{4(1-u)},\\ \frac{Tw'}{c} &= \frac{s}{2}+\frac{(1-s)\,w}{2(1-u)}+\frac{(1-s)\,z}{2(1-u)} = \frac{1}{2},\\ \frac{Tz'}{e} &= \frac{s}{4}+\frac{1-s}{4}+\frac{(1-s)\,z}{4(1-u)} = \frac{1}{4}+\frac{(1-s)\,z}{4(1-u)},\end{aligned}\right\} \tag{2·1}$$

where T is such that $u'+w'+z'=1$. Now for equilibrium $u=u'$, etc., hence at equilibrium, from (2·1), we have

$$\frac{cz}{2}=Twz=\frac{we}{4}\left(1+\frac{(1-s)\,z}{w+z}\right)$$

and hence

$$e\left(\frac{w}{z}\right)^2+\frac{w}{z}[2(e-c)-es]-2c=0.$$

If we put $w/z=v/(1-v)$, so that $v:1-v$ is the proportion of genotypes $\boldsymbol{S^hs}:\boldsymbol{S^hS^h}$, this equation becomes

$$v^2[e(s-1)]+v[2(e+c)-es]-2c=0. \tag{2·2}$$

We have further that

$$Tu = \frac{s}{4} + \frac{(1-s)\,v}{4} = u\left[\frac{1}{4} + \frac{e}{4} + \frac{c}{2} - \frac{(1-s)\,(1-e)\,(1-v)}{4}\right]$$

giving

$$s = \frac{v[1-u(1-e)]-2u(e+c)}{(1-v)\,[u(1-e)-1]}.$$

If we equate this to the value for s given by (2·2) we obtain the equation for v at equilibrium,

$$v = \frac{c[1-u(1-e)]}{(e+c)\,(1-u)} = \frac{1-u(1-e)}{(1+\alpha)\,(1-u)}, \qquad (2{\cdot}3)$$

where $\alpha = e/c$. From (2·2)

$$s = \frac{2v(1+\alpha)-\alpha v^2-2}{\alpha v(1-v)} \qquad (2{\cdot}4)$$

and since $s \geqslant 0$ we must have that

$$\alpha \geqslant \frac{2(1-v)}{2v-v^2}.$$

Table 1 gives the lower limiting values of α for given values of v.

TABLE 1. LOWER LIMITING VALUES OF α, THE VIABILITY OF S^hS^h RELATIVE TO S^hs

$v =$ 0·5	0·55	0·6	0·65	0·7	0·8
$\alpha \geqslant$ 1·3333	1·2853	0·9524	0·7978	0·6593	0·4167

In order to obtain limits for e given u and α, we must express the inequality $s \geqslant 0$ in terms of u, e and α only. Thus from (2·3) and (2·4) we must have

$$\frac{2(1-u(1-e))}{1-u} - \alpha\frac{(1-u(1-e))^2}{(1+\alpha)^2\,(1-u)^2} - 2 \geqslant 0$$

which reduces to

$$e^2 - 2e\frac{1-u}{u}\left(1+\alpha+\frac{1}{\alpha}\right) + \left[\frac{1-u}{u}\right]^2 \leqslant 0.$$

The roots e_1, e_2 of this quadratic are real and positive, and when $u < \frac{1}{2}$ only one of them, e_1 say, is less than one. Hence, since $e \leqslant 1$, for $s \geqslant 0$ we must have $e \geqslant e_1$, where

$$e_1 = \frac{1-u}{u}\left[1+\alpha+\frac{1}{\alpha} - \sqrt{\left\{\left(1+\alpha+\frac{1}{\alpha}\right)^2 - 1\right\}}\right]. \qquad (2{\cdot}5)$$

The analysis of the open pollination data, for the years 1946–54, showed that there was no significant variation in the observed proportion, u, of pins. The estimated value of u is 0·1823, with 95 % fiducial limits 0·1598 and 0·2048. We may take this estimate for u as the equilibrium value determined by the equations (2·1). When $\alpha = 0{\cdot}65$, the situation proposed by Crosby (1949), and $u = 0{\cdot}1823$, $e_1 = 0{\cdot}714$ and we have $c > 1$. Hence the observed proportion of pins will not tolerate so low a viability for the homozygous homostyles, even when there is no selfing. In order that $c \leqslant 1$ or $e_1 \leqslant \alpha$ we must have

$$u \geqslant \frac{F(\alpha)}{\alpha+F(\alpha)}, \quad \text{where} \quad F(\alpha) = 1+\alpha+\frac{1}{\alpha} - \sqrt{\left\{\left(1+\alpha+\frac{1}{\alpha}\right)^2 - 1\right\}}$$

which gives $u \geqslant 0{\cdot}1985$ when $\alpha = 0{\cdot}65$. When $u = 0{\cdot}1985$ and $\alpha = 0{\cdot}65$, we have $e_1 = 0{\cdot}65$, $c_1 = 1$, $v = 0{\cdot}7036$ and no selfing can occur.

In order to obtain reasonable limits for the proportion of crossing and the relative viabilities from the open pollination data, it is necessary to have some estimate of the quantity v. The observed proportions of segregating progenies for the years 1948–55 are given in table 2.

TABLE 2. OBSERVED PROPORTIONS OF SEGREGATING PROGENIES

year	segregating	non-segregating	total
1948	20	12	32
1949–50	10	6	16
1951	8	3	11
1952	3	3	6
1953	9	4	13
1954	8	4	12
1955	20	4	24
	78	36	114

A heterogeneity χ^2 gives 4·307 for 5 d.f., which is not significant. We thus obtain an estimate for v of 0·6842 with 95 % fiducial limits 0·5992 and 0·7692. Now when $v = 0{\cdot}77$ we must have $\alpha \geqslant 0{\cdot}486$ for positive s. At this value of α however $u \geqslant 0{\cdot}229$ if $c \leqslant 1$. This is greater than the upper fiducial limits for u obtained from the data, which therefore will not very plausibly tolerate such high values of v and hence low possible α. When u is 0·2 the lower limit for e is about 0·7. This therefore seems to be the minimum viability that it is reasonable to assign to the homozygous homostyles. In order that $c \leqslant 1$ when $u = 0{\cdot}2$ we must have $\alpha \geqslant 0{\cdot}7$. Hence 0·7 seems to be the lowest reasonable value for α tolerated by the data. In general, for a low value of e, c must be close to unity and the proportion of crossing very high. For example, when $u = 0{\cdot}19$, $v = 0{\cdot}685$, $c = 1$ and $e = 0{\cdot}7$, the proportion of crossing is 99 %. If there are no viability disturbances, $e = c = \alpha = 1$ and from (2·3) and (2·4) we have $s = \{8u(1-u)-1\}/(1-2u)$. This gives $s = 0{\cdot}296$ when $u = 0{\cdot}1823$. Thus 70 % is certainly about the lowest frequency of crossing that the data would tolerate.

Since 1955 progenies have been grown only from homostyles which were forced to self-fertilize. The total proportion of pins observed is 223/893 and there is no evidence of heterogeneity between the years. These data do not therefore show any indication of a viability disturbance.

It should be pointed out that there is no evidence for any increase in the proportion of pins observed successively in the four generations up to 1954 derived from the original plants, and grown in open-pollination plots. Of the progenies derived from homostyles which were forced to self-fertilize, only those grown in 1957 were derived from the 1955 progenies which showed an unexplained excess of pins over the previous years. The proportion of pins observed for 1956 and 1958 together is 123/499, and these are derived entirely from progenies which showed the characteristic deficiency of pins on open-pollination. The progeny size has varied between 16 and 25. Hence if we assume the expected proportion of pins is 1/6, the average number of progenies from heterozygous homostyles which have been overlooked due to their not segregating is certainly less than 5·4 %. Crosby (1959) has suggested that the observed increase in the proportion of pins on forced self-fertilization of the homostyles may actually be due to selection for a higher pin proportion as a result of breeding only from those progenies which during the previous years had segregated. This excludes progenies of heterozygous homostyles which did not segregate and so might have given, on average, lower pin proportions. However, it seems difficult to suppose that such a lax selection differential as that indicated could, after only four generations, result in an extremely abrupt response to the selection of what must admittedly be a fairly complex character.

The total number of progenies grown is as yet inadequate to detect any departure from the 2:1 ratio expected for segregating to non-segregating progenies when there are no viability disturbances. It should be pointed out that the number of progenies needed, to distinguish for example a 2:1 from a 2:0·8 ratio corresponding to a viability of 80 % for the homozygous homostyles, is 1006 for a minimum certainty of 5 %.

It is clear that data on viability disturbances collected at Cambridge are not necessarily applicable to the natural populations in Somerset. Apart from straightforward differences in environment, the careful supervision given under experimental conditions may be expected to reduce any viability disturbances. Moreover, as indicated above, in order to detect relatively small but significant viability disturbances, a large number of progenies must be grown over several years. There is thus likely to be heterogeneity in the data at least of the same order of magnitude as the disturbances to be detected. This implies an indeterminacy in the detection of the small variations in fitness necessary for evolutionary progress by natural selection, the techniques for detection having coarser errors than the disturbances to be detected. More confidence can be placed on the estimates of cross-fertilization. The data are remarkably homogeneous and there is little reason to suppose that this will be much influenced by the experimental conditions of growth.

3. An analysis of some population counts

Since 1941 Professor Sir Ronald Fisher, F.R.S., has been collecting counts of *Primula* populations around Sparkford in Somerset. He has very kindly made his records available to me.

The constitutions of counts totalling more than about fifty are plotted on a triangular graph in figure 2. The use of the triangular graph enables the three frequencies of the three recognizable forms of the flower to be plotted as a single point. The vertices H, P and T correspond to 100 % of homostyles, pins and thrums, respectively. The frequencies of homostyles, pins and thrums represented by any point on the graph are given by the respective perpendicular distances from the point onto the lines $H = 0$, $P = 0$ and $T = 0$. The lines superimposed on the data are two of the theoretically calculated curves representing the population changes, which will be discussed in § 5.

The counts seem to follow a consistent pattern and show a similar configuration to those charted by Crosby (1949). Counts less than 100 are represented by points, those between 100 and 200 by circles, and counts greater than 200 by triangles. There is a concentration of points in the high and low homostyle regions. This is to be expected if the data represent a consistent change in the homostyle frequencies caused by their fertility relationships. For, as will be shown in the next two sections, these regions will be the ones in which the rate of change is slowest.

In order to gain information about the progress of such changes as those in the *Primula* populations, it is important to have counts from populations extending over a number of years. Counts collected over several years, from populations in the Sparkford area, are given in tables 3 to 7. For tables 4 and 6 the heterogeneity χ^2's were calculated on grouping the pin and thrum frequencies. It is difficult to detect any consistent trend in these data. The heterogeneity of the Sparkford north-west counts is due largely to the differences in the thrum frequencies, particularly for 1952, which between them contribute 20·6 to the overall heterogeneity χ^2. The heterogeneity of the Laurel Copse hedgerow counts stems from the sudden apparent change in the population constitution between 1943 and 1944, the data for the years 1942 and 1943, and the years 1944, 1948 and 1949 forming two homogeneous groups. Four counts made near Laurel Copse in 1943 are given in table 8 and these also show considerable heterogeneity. It is characteristic for such population data to show heterogeneity even over very small regions. This can be interpreted as indicating a very short average breeding range and Crosby (1948) has investigated this possibility in some detail. The heterogeneity of the Sparkford north-west data might, therefore, easily be due to small variations from year to year in the exact location of the counts. It is, however, difficult to account for the sudden change in the constitution of the Laurel Copse hedgerow population in this way. It has a clearly defined habitat which allows for no such variation in location of counts. The heterogeneity of the Sparkford south data is on the verge of significance at 5 %. This is the only set of counts which shows any clear trend in the homostyle frequency and contrary to expectation the trend is downward. A χ^2 for 1 d.f. can be calculated to test the significance of a linear trend (Fisher 1925–54). For this data the χ^2_1 is 7·3238, which is highly significant with a probability between 1 and 0·5 %. The residual χ^2_2, obtained

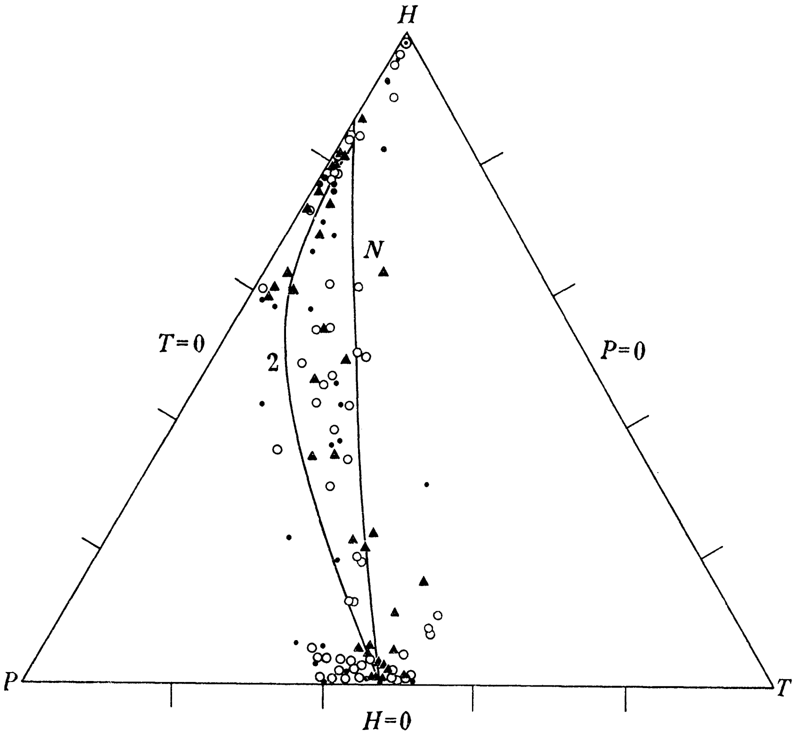

FIGURE 2. Constitutions of observed populations of *Primula vulgaris* with two calculated curves of change. Counts totalling: < 100, ●; 100–200, ○; > 200, ▲. *N* = Progress on data *N*. 2 = Regress 2 with partial replacement. *Note.* Only counts totalling more than fifty are included. For discussion of the two calculated curves see § 5.

TABLE 3. *PRIMULA* COUNTS FROM WILFORD COPSE

year	homostyles	pins	thrums	total	%		
					H	*P*	*T*
1942	4	102	82	188	2·1	54·3	43·6
1943	15	159	128	302	5·0	52·6	42·4
1944	46	505	437	978	4·7	51·6	43·7
1948	7	103	66	176	4·0	58·5	37·5

Heterogeneity $\chi^2_6 = 5{\cdot}5016$, $0{\cdot}5 > P > 0{\cdot}3$.

TABLE 4. *PRIMULA* COUNTS FROM SPARKFORD WOOD, SOUTH

year	homostyles	pins	thrums	total	%		
					H	*P*	*T*
1941	257	52	3	312	82·4	16·7	0·9
1943	195	50	2	247	78·9	20·3	0·8
1944	117	31	2	150	78·0	20·7	1·3
1948	148	56	1	205	72·3	27·3	0·5

Heterogeneity $\chi^2_3 = 7{\cdot}636$, $P \simeq 5{\cdot}1\,\%$. χ^2_1 for linear downward trend $= 7{\cdot}3238$.
$P \simeq 0{\cdot}75\,\%$ leaving residual $\chi^2_2 = 0{\cdot}3122$.

TABLE 5. *PRIMULA* COUNTS FROM SPARKFORD WOOD, NORTH-WEST

year	homostyles	pins	thrums	total	%		
					H	*P*	*T*
1941	64	41	15	120	53·3	34·2	12·5
1943	156	67	27	250	62·4	26·8	10·8
1948	107	68	25	200	53·5	24·0	12·5
1952	183	112	10	305	60·0	36·7	3·3
1954	57	39	4	100	57·0	39·0	4·0

Heterogeneity $\chi^2_8 = 27{\cdot}6$, $P < 0{\cdot}001$.

TABLE 6. *PRIMULA* COUNTS FROM EDWARD LEAR COPSE

year	homostyles	non-homostyles	total	%		
				H	*P*	*T*
1943	69	14	83	83·1	14·5	2·4
1944	167	33	200	83·5	16·5	—
1948	201	49	250	80·4	19·2	0·4
1952	299	79	378	79·1	19·8	1·1
1954	184	53	237	77·6	22·4	—

Heterogeneity $\chi^2_4 = 3{\cdot}0855$, $0{\cdot}7 > P > 0{\cdot}5$.

TABLE 7. *PRIMULA* COUNTS FROM LAUREL COPSE HEDGEROW

year	homostyles	pins	thrums	total	%		
					H	*P*	*T*
1942	148	191	88	427	34·7	44·7	20·4
1943	81	98	56	235	34·5	41·7	23·8
1944	137	302	242	681	20·1	44·3	35·6
1948	77	163	117	357	21·6	45·7	32·7
1949	37	89	69	195	19·0	45·6	35·4

Heterogeneity χ^2 gives $P \leqslant 0{\cdot}001$.

by subtraction from the overall heterogeneity χ^2, is 0·3122 for 2 d.f. and is not significant. Hence the heterogeneity of the Sparkford south data can be entirely accounted for by a systematic decrease in the frequency of homostyles. This lends added weight to the reality of the sudden decrease in the frequency of homostyles in the Laurel Copse hedgerow population. The counts from Edward Lear's Copse also show a consistent downward trend, but this is clearly not significant. However, it may be pointed out, as was indicated in the previous section, that a large amount of data is needed to detect even a change of 10 %.

Table 8. *Primula* counts near and around Laurel Copse from 1943

				%		
homostyles	pins	thrums	total	*H*	*P*	*T*
53	78	46	177	29·9	44·1	26·0
73	59	18	150	48·7	39·3	12·0
22	54	24	100	22·0	54·0	24·0
81	98	56	235	34·5	41·7	23·8

Heterogeneity $\chi^2_6 = 26·3045$, $P < 0·001$.

Table 9. Counts of populations in Somerset containing no homostyles

pins	thrums	total	χ^2_1 for deviation from 1:1
72	75	147	0·0612
95	94	169	2·6094
63	62	125	0·0080
40	42	82	0·0487
112	98	210	0·9333
107	88	195	1·8512
50	41	91	1·3296
187	158	345	2·4376
122	104	226	1·4336
74	50	124	4·6451
162	143	305	1·1836
72	52	124	3·2258
49	34	84	0·3012
76	71	147	0·1700
1281	1092	2373	15·0530

Heterogeneity $\chi^2_{13} = 5·2174$, $0·98 > P > 0·95$.

Crosby (1948) also observed a decrease in homostyle frequencies in a number of his counts. He explained these as due to variation in the sampling area from year to year as was suggested for the heterogeneity of the Sparkford north-west data. However, it is unlikely that such systematic downward changes in the homostyle frequency would be caused by chance variation in the exact location of the counts. Variation from this source would be expected to occur randomly in all possible directions of change of the three frequencies.

A number of counts from populations in the Sparkford area not containing homostyles are given in table 9. The data are remarkably homogeneous, and as a whole differ very significantly from a 1:1 ratio. The estimated frequency of pins is 53·98 % compared with that assumed by Crosby (1949) of 52·7 %. The homogeneity of the data indicate that pins and thrums in this area are at a stable equilibrium, and that the breeding conditions throughout the area are probably uniform. That breeding conditions for *Primula* may vary in different regions is indicated by the counts of pins and thrums from some populations in Cornwall, given in table 10. These show considerable heterogeneity. However, the counts with pins less frequent than thrums are homogeneous, as are those with pins more frequent than thrums. The data from the populations containing pins and thrums only are summarized in table 11.

It is clearly very difficult to base any detailed analysis on such population data as these, although a number of major characteristics of the change are clearly indicated by the data. There is a marked deficiency of thrums relative to pins as the frequency of homostyles increases. The lack of any upward trend in the homostyle frequency over the years in which counts have been made indicates a slow approach to equilibrium, in terms of years, if possibly not generations. There is no clear evidence for a final equilibrium of about 80 % homostyles and 20 % pins as Crosby supposes. In fact there is a definite indication of a downward trend in homostyle frequency for populations with a high proportion of homostyles.

TABLE 10. COUNTS OF POPULATIONS IN CORNWALL CONTAINING NO HOMOSTYLES

pins	thrums	total	χ_1^2 for deviation from 1:1
178	204	382	1·7650
239	225	464	0·4224
26	21	47	0·5319
89	70	159	2·2704
26	45	71	5·0845
40	49	89	0·9101
47	33	80	2·4500
29	45	74	3·4594
32	33	65	0·0153
706	725	1431	0·2522

Heterogeneity $\chi_8^2 = 16{\cdot}6585$, $0{\cdot}05 > P > 0{\cdot}025$.

TABLE 11. SUMMARY OF COUNTS FROM POPULATIONS WITH NO HOMOSTYLES

					%	
	pins	thrums	total	χ^2 for deviation from 1:1	P	T
Somerset	1281	1092	2373	15·0530	53·98	46·02
Cornwall $P > T$	401	349	750	3·6053	53·5	46·5
Cornwall $P < T$	305	376	681	7·4023	44·8	55·1

4. A THEORETICAL ANALYSIS OF THE POPULATION CHANGES

4·1. *The construction of models and investigation of their equilibria*

We shall consider first the situation in populations containing only pins and thrums. Assume that the pollen and ovule fertilities of pins and thrums are the same, but that the relative proportion of pin and thrum zygotes surviving to maturity are a and b. Assume further that a proportion p of pins self-fertilize illegitimately, but that only pin pollen fertilizes the thrums. Then if the frequencies of pins and thrums in one generation are u and v, where $u+v = 1$, the expected frequencies u', v' in the succeeding generation are given by

$$\frac{Tu'}{a} = (1-p)\,\tfrac{1}{2}u+\tfrac{1}{2}v+up = \tfrac{1}{2}(1+up),$$

$$\frac{Tv'}{b} = (1-p)\,\tfrac{1}{2}u+\tfrac{1}{2}v = \tfrac{1}{2}(1-up),$$

where T is such that $u'+v' = 1$. The equilibrium value of u is given by equating u to u', and is a solution of

$$u^2p(1-b/a)+u[1+b/a-p]-1 = 0.$$

If $b/a < 1$, the product of the two roots of this equation is negative, and so there exists only one valid equilibrium frequency, u_0, say. To investigate the stability of this equilibrium put $u = u_0+x$, where x is small, then

$$T(u_0+x') = \tfrac{1}{2}a[1+u_0p+xp].$$

Now

$$\begin{aligned} T &= \tfrac{1}{2}a(1+up)+\tfrac{1}{2}b(1-up) \\ &= \tfrac{1}{2}[a+b+p(a-b)\,(u_0+x)] = T_0+\tfrac{1}{2}\{xp(a-b)\} \end{aligned}$$

and hence

$$u_0+x' = \frac{a(1+u_0p)}{2T_0}+x\left[\frac{ap}{2T_0}-\frac{p(a-b)\,a(1+u_0p)}{4T_0^2}\right]+O(x^2).$$

Since u_0 is the equilibrium frequency

$$u_0 = \frac{a(1+u_0p)}{2T_0}$$

and so

$$x' = \frac{(ap-p(a-b)\,u_0)}{2T_0}x = kx, \quad \text{say.}$$

Now $p < 1$, and $b > 0$, so that $k < 1$ and thus the equilibrium is stable.

If there are no disturbances and so $a = b$ and $p = 0$, the equilibrium populations contain 50 % each of pins and thrums. Hence, in order to explain the observed frequency of pins and thrums in natural populations, it is necessary to assume some disturbance of the sort considered above. Crosby (1949) assumed that the disturbed ratio was due entirely to illegitimate selfing of the pins. Thus when $a = b$, $u_0 = 1/(2-p)$, and with $p = 0{\cdot}1$, u_0 is 52·63 % which is the situation proposed by Crosby. If, on the other hand, $p = 0$,

$$u_0 = \frac{1}{1+\alpha}, \quad \text{where} \quad \alpha = \frac{b}{a}, \tag{4·1}$$

and to obtain the same equilibrium frequency as above we must have $\alpha = 0{\cdot}89$. The high observed frequency of cross-fertilization between the homostyles suggests that in natural conditions where there is available large amounts of either thrum or homostyle pollen which can legitimately fertilize the pin ovules, illegitimate selfing of the pins is unlikely to occur. Moreover, Crosby (1948) found that when a mixture of pin and thrum pollen was placed on a pin stigma the thrum pollen always reached the ovules first even when the pin pollen had a day's start. It is of course possible to obtain the proposed equilibrium of 52·63 % pins by a mixture of illegitimate self-fertilization of the pins and a viability deficiency for the thrums. For example, if $\alpha = 0{\cdot}95$, $p = 0{\cdot}065$ and when $\alpha = 0{\cdot}975$, $p = 0{\cdot}089$.

In considering the model for the general situation, in which all the types of flower are present, we shall again, at first, assume equal pollen and ovule fertility for all the genotypes. This implies that of the types of pollen available to fertilize legitimately one of the ovules, none is at any selective advantage or appreciably more abundant. It also implies that the average number of fertilized zygotes, produced by each genotype as seed parent, is the same. We shall further assume that a proportion p of pins self-fertilize illegitimately, but that no other illegitimate fertilizations occur, and that a proportion s of the homostyles self-fertilize. The thrum genotype ***SS*** cannot then occur. The frequencies of the five genotypes that occur, together with their relative survival values from zygote to maturity, or viabilities, are illustrated in table 12.

The expected frequencies of the genotypes in a generation, in terms of the frequencies in the previous generation, can be obtained by taking each genotype in turn as seed parent, considering the possible types of pollen available to fertilize it and tabulating the results of each cross. Consider, for example, the cross with an $\boldsymbol{S^h s}$ homostyle as seed parent and an $\boldsymbol{SS^h}$ thrum as pollen parent. There are four possible sources of pollen for the $\boldsymbol{S^h s}$ homostyle seed parent, and that from the $\boldsymbol{SS^h}$ thrum constitutes a proportion $x/v+w+x+z$ of the total available. Moreover, a proportion $(1-s)$ of progeny from the homostyle is produced by cross-fertilization so that the cross $\boldsymbol{S^h s} \times \boldsymbol{SS^h}$ gives progeny with the frequencies

$$\frac{(1-s)\,wx}{v+w+x+z}\left(\frac{b}{4}\boldsymbol{Ss}+\frac{c}{4}\boldsymbol{S^h s}+\frac{d}{4}\boldsymbol{SS^h}+\frac{e}{4}\boldsymbol{S^h S^h}\right).$$

TABLE 12. FREQUENCIES AND VIABILITIES OF THE FIVE OCCURRING GENOTYPES

flower form	genotype	frequency	viability
pin	$\boldsymbol{ss}$	u	a
thrum	$\boldsymbol{Ss}$	v	b
homostyle	$\boldsymbol{S^h s}$	w	c
thrum	$\boldsymbol{SS^h}$	x	d
homostyle	$\boldsymbol{S^h S^h}$	z	e

The results from all the possible crosses can then be combined to give the equations (4·2) below, for the expected frequencies u', etc., in one generation in terms of those in the preceding generation:

$$\left.\begin{aligned}
\frac{Tu'}{a} &= (1-p)\,u\left[\frac{v+w}{2(1-u)}-1\right]+(1-s)\,w\left[\frac{v+w}{4(1-u)}-\frac{1}{4}\right]+u+\tfrac{1}{2}v+\tfrac{1}{4}w, && (a)\\
\frac{Tv'}{b} &= \frac{(1-p)\,u(v+x)}{2(1-u)}+\frac{(1-s)\,w(v+x)}{4(1-u)}+\tfrac{1}{2}v+\tfrac{1}{2}x, && (b)\\
\frac{Tw'}{c} &= \frac{(1-p)\,u(x+w+2z)}{2(1-u)}+(1-s)\left[\frac{w(v+2w+x)}{4(1-u)}+\frac{z(2w+v)}{2(1-u)}-\frac{w}{2}\right]+\tfrac{1}{2}w+\tfrac{1}{2}x, && (c)\\
\frac{Tx'}{d} &= \frac{(1-s)\,(v+x)\,(w+2z)}{4(1-u)}, && (d)\\
\frac{Tz'}{e} &= (1-s)\left[\frac{(w+x+2z)\,(w+2z)}{4(1-u)}-\tfrac{1}{4}w-z\right]+\tfrac{1}{4}w+z. && (e)
\end{aligned}\right\}\quad (4{\cdot}2)$$

Here $u+v+w+x+z=1$, and T is such that $u'+v'+w'+x'+z'=1$. When

$$a=b=c=d=e=1, \quad s=1 \quad \text{and} \quad p=0$$

these equations reduce to those given by Fisher (1949), and when

$$a=b=c=d=1, \quad s=1 \quad \text{and} \quad p=\tfrac{1}{10}$$

they reduce to those given by Crosby (1949). A general analytic solution of these equations does not seem possible, and discussion of a numerical treatment of these equations with the aid of an electronic digital computer, will be given in the next section.

Consider first the case of 100 % crossing between the homostyles, no selfing of the pins, and no viability disturbances, i.e.

$$a=b=c=d=e=1, \quad s=0 \quad \text{and} \quad p=0$$

The equations reduce to

$$\left.\begin{aligned}
u' &= \frac{u(v+w)}{2(1-u)}+\tfrac{1}{2}v+\frac{w(v+w)}{4(1-u)}, && (a)\\
v' &= \frac{u(v+x)}{2(1-u)}+\tfrac{1}{2}v+\frac{w(v+x)}{4(1-u)}+\tfrac{1}{2}x, && (b)\\
w' &= \frac{u(x+w+2z)}{2(1-u)}+\frac{w(v+2w+x+2z)}{4(1-u)}+\tfrac{1}{2}x+\frac{z(v+w)}{2(1-u)}, && (c)\\
x' &= \frac{(v+x)\,(w+2z)}{4(1-u)}, && (d)\\
z' &= \frac{(w+x+2z)\,(w+2z)}{4(1-u)}. && (e)
\end{aligned}\right\}\qquad (4\cdot3)$$

For equilibrium we equate u to u', etc. Equations (4·3) $(b)+(d)$ give

$$(v+x)=\frac{v+x}{2}\left[\frac{u}{1-u}+1+\frac{w}{1-u}+\frac{z}{1-u}\right],$$

or if $v+x\neq 0, 1=2u+w+z=u+v+x+z$ and so $u=v+x$. Hence from (b) we have $v=\{u(w+2)\}/\{4(1-u)\}$, and so (a) gives

$$w^2(4-3u)+4uw(3-2u)-4u(3-2u)\,(1-2u)=0 \qquad (4\cdot4)$$

an equation determining w in terms of u. It is easily shown that using equations (b), (d) and (e) we obtain the same equation for w in terms of u. Thus, since there exists the relation $u+v+w+x+z=1$ connecting the five quantities, only three of the equations (4·3) are independent. Hence there must be an infinite linear continuum of equilibrium frequencies, determined in terms of u by equation (4·4). Solving for w in terms of u we obtain

$$\begin{aligned}
w &= \frac{-2u(3-2u)+\surd\{4u^2(3-2u)^2+4(4-3u)\,(1-2u)\,(3-2u)\,u\}}{4-3u}\\
&= \frac{-2u(3-2u)+4(1-u)\surd\{u(3-2u)\}}{4-3u} \quad \text{since} \quad w>0.
\end{aligned}$$

If we put $u=\frac{3}{2}\sin^2\theta$, then $\quad \surd\{u(3-2u)\}=\dfrac{3\sin\theta\cos\theta}{\surd 2}$

and the five frequencies can be expressed as rational functions of $\sin\theta$ and $\cos\theta$, the parameter θ varying from 0 to $\sin^{-1}\surd\frac{1}{3}$ (when $u=\frac{1}{2}$). We have, thus, at equilibrium

$$\left.\begin{aligned}
u &= \tfrac{3}{2}s^2,\\
w &= \frac{6sc(c-s\surd 2)}{2\surd(2)\,c-s},\\
v &= \frac{u(w+2)}{4(1-u)},\\
z &= 1-2u-w,\\
\text{and}\qquad x &= u-v,
\end{aligned}\right\}\qquad (4\cdot5)$$

where $s\equiv\sin\theta$ and $c\equiv\cos\theta$. I am grateful to Professor Sir Ronald Fisher, F.R.S., for suggesting this transformation to me as a convenient way of representing the infinite linear continuum of equilibrium frequencies in terms of a single parameter. The values of w, v, x and z are plotted against u in the graph given in figure 3. The frequencies of pins and thrums are always equal at equilibrium, and so this clearly does not represent the situation in the

Somerset populations. If plotted on a triangular graph the line of change would be the perpendicular from *H* onto *PT*.

A theoretical consideration of the stability of equilibria along the continuum was not feasible, but a numerical investigation was carried out with the aid of the electronic digital computer, EDSAC II at the Mathematical Laboratory in Cambridge. The set of equilibrium frequencies was calculated for a small interval in the parameter θ. Each set of frequencies was then rounded off to about five decimal places, the machine's accuracy being eleven places, in order to provide a set of frequencies slightly disturbed from equilibrium.

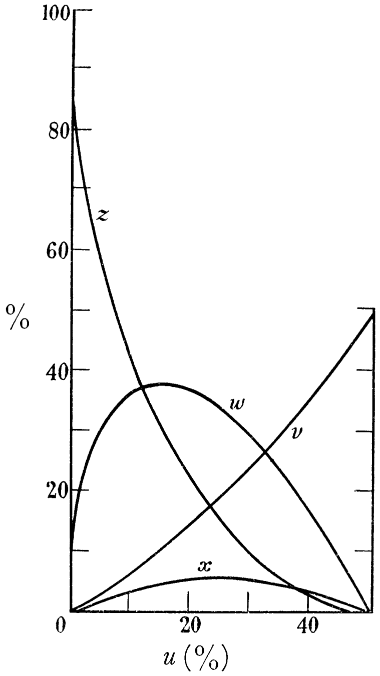

FIGURE 3. The infinite linear continuum of possible equilibria resulting from the model in which only crossing occurs and there are no viability disturbances, is illustrated by plotting the equilibrium values of v, w, x and z against the value of u.

The equilibrium frequencies attained, by starting with this rounded set of frequencies, and calculating frequencies in successive generations from the equations (4·3), were then recorded over the whole range of θ. These calculated frequencies corresponded with the initial theoretically calculated frequencies to about four places of decimals, thus indicating stability along the whole infinite continuum determined by θ.

We shall now consider the more general situation, in which a proportion s of selfing occurs, but still taking $p = 0$ and $a = b = c = d = e = 1$. As many of the observed populations with high homostyle frequencies lacked thrums, we consider first what happens when $v = x = 0$. The relevant equations describing the expected changes are

$$\left.\begin{aligned} u' &= \frac{w}{4(1-u)}\,[s+u(2-s)+w(1-s)], && (a)\\ w' &= \frac{(w+2z)}{2(1-u)}\,[u+(1-s)\,w]+\tfrac{1}{2}sw, && (b)\\ z' &= \frac{(1-s)}{4(1-u)}\,(w+2z)^2+\tfrac{1}{4}sw+sz. && (c)\end{aligned}\right\}\qquad (4\cdot 6)$$

At equilibrium we have, from (4·6) (*b*) and (*c*), that

$$w+2z = \frac{w+2z}{2(1-u)}\,[u+(1-s)\,w+(1-s)\,(w+2z)+2s(1-u)].$$

Disregarding the trivial case $w+2z=0$ (corresponding to a population consisting entirely of pins) we must have

$$2(1-u) = u+(1-s)\,w+2s(1-u)+(1-s)\,w+2(1-s)\,(1-u-w)$$

which reduces to $u=0$. From (4·6) (*a*) this implies $w=0$, as $0<s<1$. Hence $z=1$, and the homostyles reach 100 % for all values of s. Taking into account thrums, we substitute $a=b=c=d=e=1$ and $p=0$ in equation (4·2). From (4·2) (*b*) and (*d*) we obtain, for equilibrium

$$(v+x) = \frac{u(v+x)}{2(1-u)}+\frac{(1-s)\,w(v+x)}{4(1-u)}+\frac{v+x}{2}+\frac{(1-s)\,(v+x)\,(w+2z)}{4(1-u)},$$

or

$$\frac{v+x}{2(1-u)}\,[(1-s)\,(w+z)-1+2u]=0.$$

This implies either $v+x=0$ or $(1-s)\,(w+z)-1+2u=0$. Now if the latter obtains, $v+x=1-w-z-u=1-u-(1-2u)/(1-s)$ and must be greater than zero. Hence we must have $(1-s)\,(1-u)\geqslant 1-2u$ or $u\geqslant s/(1+s)$. Now as s increases from 0, $s/(1+s)$ increases from 0 to 0·5, and so the lower limit for u increases as the proportion of selfing increases. Such a situation is not likely to be realized in practice. This implies that generally $v+x=0$, and so from the above, that the homostyles reach 100 % even when there is crossing. That this is so, has been verified numerically, by the methods to be described in the next section.

It is possible to introduce further parameters representing fertility differences. Thus we may take the relative survival values of the pollen for the five genotypes to be λ, μ, ν, ρ and τ respectively, and of the ovules to be j, k, l, m and n. These quantities refer to the relative survival of gametes up to the point of fertilization. The products of the various crosses may be considered exactly as before and the equations representing the changes in frequency obtained. For example, the cross $\boldsymbol{S^h s}\times\boldsymbol{S s}$ occurs with a frequency

$$(1-s)\,w(\mu\nu)/(\mu\nu+\nu w+\rho x+\tau z)$$

produces a relative quantity l of ovules and produces offspring in the proportions $\frac{1}{4}a\boldsymbol{ss}+\frac{1}{4}d\boldsymbol{SS^h}+\frac{1}{4}c\boldsymbol{S^h s}+\frac{1}{4}b\boldsymbol{Ss}$. The general equations will be even less manageable than the equations (4·2). As so little information on these further parameters is available, it has not seemed worthwhile to include the situation in our numerical investigation. It is easily shown, however, that if pollen and ovule survival rates are the same for each genotype, the equations reduce to (4·2) with u, etc., replaced by ju, etc.; that is, zygotic frequencies are replaced by gametic frequencies.

The general case when there are no homostyles, presents no difficulties. The equations for pin and thrum frequencies, not taking into account any illegitimate self-fertilization of pins, become

$$Tu' = \frac{aj\mu u}{2}+\frac{ak\lambda v}{2},$$

$$Tv' = \frac{bj\mu u}{2}+\frac{bk\lambda v}{2}.$$

Here, as usual, T is such that $u'+v' = 1$ and the other symbols have the meanings attached to them above. If we translate

$$j\mu u \to u, \quad k\lambda v \to v,$$

and

$$j a\mu \to a, \quad k\lambda b \to b$$

these equations are the same as before with $p = 0$, and taking T such that $j\mu u' + k\lambda v' = 1$.

4·2. *The initial progress of newly occurring homostyles*

We shall now investigate the initial stages of the progress determined by the general system of equations (4·2) after the introduction of a small number of homostyles into a population containing only pins and thrums. This is in order to obtain information about the selective pressure acting on newly occurring homostyles in predominantly pin-thrum populations. Substitute $u = u_0 - U$, $v = v_0 - V$, $w = W$, $x = X$ and $z = Z$, where U, V, W, X and Z are small and u_0, v_0 are the equilibrium frequencies of pins and thrums before the introduction of the homostyles. Then on neglecting quadratic terms in U, etc., we can obtain a set of linear equations for the initial changes when the frequency of homostyles is small.

Let $\boldsymbol{U}$ stand for the set of frequencies U, V, W, X, Z considered as vector and let $\lambda_1, \lambda_2, \ldots, \lambda_6$ be the latent roots of the linear transformation $\boldsymbol{U}' = A\boldsymbol{U}$ connecting the values of $\boldsymbol{U}$ in consecutive generations. Then each component of $\boldsymbol{U}$ in the nth generation can be expressed as a sum $\sum_{j=1}^{6} a_j \lambda_j^n$, where the a_j are determined by the initial values of the frequencies. The rate of increase of the frequencies is asymptotically determined by the largest of the latent roots, this being the initial common ratio of geometic increase for the newly occurring genotypes. This latent root λ_1, say, may thus be considered as the logarithm of the Malthusian parameter, in the sense defined by Fisher (1930), at least for the initial stages of the process. Thus if $\lambda_1 = 1+\xi$, where ξ is small, ξ can be considered as an effective selective pressure for the homostyle gene. Using the results obtained by Fisher (1930) we can estimate the chance of survival of a single homostyle gene formed either by crossing-over or mutation within the heterostyly gene complex. The relative importance of the random effects of finite population size may also be assessed.

Assume that $p = 0$ and $a = 1$, and let $b = 1-\beta$, $c = 1-\gamma$, $d = 1-\delta$ and $e = 1-\epsilon$. Then from equation (4·1) we have

$$u_0 = 1/(2-\beta) = \tfrac{1}{2}(1+\tfrac{1}{2}\beta) \quad \text{and} \quad v_0 = \tfrac{1}{2}(1-\tfrac{1}{2}\beta)$$

neglecting β^2 and higher powers of β. Now substituting for u, etc., in terms of U, etc., in (4·2) and neglecting quadratic terms in U, etc., and also in β and s we obtain

$$\begin{aligned}
T(\tfrac{1}{2}(1+\tfrac{1}{2}\beta)-U') &= \tfrac{1}{2}-\tfrac{1}{2}U-\tfrac{1}{2}V+\tfrac{1}{4}W-\tfrac{1}{2}(1+\beta)\,(X+Z),\\
T(\tfrac{1}{2}(1-\tfrac{1}{2}\beta)-V')/(1-\beta) &= \tfrac{1}{2}-\tfrac{1}{2}U-\tfrac{1}{2}V-\tfrac{1}{4}W[1+s+2\beta]+\tfrac{1}{2}X-\tfrac{1}{2}(1+\beta)\,Z,\\
TW'/(1-\gamma) &= \tfrac{1}{4}W[3+2\beta+s]+X(1+\tfrac{1}{2}\beta)+\tfrac{1}{2}Z(3+2\beta-s),\\
TX'/(1-\delta) &= \tfrac{1}{4}(1-s)\,(W+2Z),\\
TZ'/(1-\epsilon) &= \tfrac{1}{4}sW+sZ.
\end{aligned}$$

As before, T is such that $u'+v'+w'+x'+z'=1$, i.e. such that $W'+X'+Z'-U'-V'=0$. On substituting in these equations for T, and neglecting also quadratic terms in γ, δ and ϵ we obtain the following linear transformation for U', V', W', X' and Z':

$$\begin{aligned}
U' &= \tfrac{1}{2}(1+\beta)\,U+\tfrac{1}{2}(1+\beta)\,V-\tfrac{1}{8}(2+3\gamma+\delta)\,W-\tfrac{1}{2}(\gamma-\beta-1)\,X-\tfrac{1}{4}(3\gamma+\delta-4\beta-2)\,Z,\\
V' &= \tfrac{1}{2}U+\tfrac{1}{2}V+\tfrac{1}{8}(2+4\beta-3\gamma-\delta+2s)\,W-\tfrac{1}{2}(1+\gamma)\,X+\tfrac{1}{4}(2+2\beta-3\gamma-\delta)\,Z,\\
W' &= \tfrac{1}{4}(3+\tfrac{7}{2}\beta+s-3\gamma)\,W+(1+\beta-\gamma)\,X+\tfrac{1}{2}(3+\tfrac{7}{2}\beta-s-3\gamma)\,Z,\\
X' &= \tfrac{1}{4}(1+\tfrac{1}{2}\beta-\delta-s)\,W+\tfrac{1}{2}(1+\tfrac{1}{2}\beta-\delta-s)\,Z,\\
Z' &= \tfrac{1}{4}sW+sZ.
\end{aligned}$$

The characteristic equation is

$$\begin{vmatrix} \tfrac{1}{2}(1+\beta)-\lambda & \tfrac{1}{2}(1+\beta) \\ \tfrac{1}{2} & \tfrac{1}{2}-\lambda \end{vmatrix} \quad \begin{vmatrix} \tfrac{1}{4}(3+\tfrac{7}{2}\beta+s-3\gamma)-\lambda & 1+\beta-\gamma & \tfrac{1}{2}(3+\tfrac{7}{2}\beta-s-3\gamma) \\ \tfrac{1}{4}(1+\tfrac{1}{2}\beta-\delta-s) & -\lambda & \tfrac{1}{2}(1+\tfrac{1}{2}\beta-\delta-s) \\ \tfrac{1}{4}s & 0 & s-\lambda \end{vmatrix} = 0$$

or

$$\lambda(\lambda-(1+\tfrac{1}{2}\beta))=0$$

and

$$\lambda^3-\tfrac{1}{4}\lambda^2[3+\tfrac{7}{2}\beta-3\gamma+5s]-\tfrac{1}{4}\lambda[1+\tfrac{3}{2}\beta-\delta-\gamma-\tfrac{7}{2}s]+\tfrac{1}{8}s=0. \tag{4·7}$$

When $\beta=\gamma=\delta=s=0$ equation (4·7) reduces to $\lambda(\lambda-1)\,(\lambda+\frac{1}{4})=0$. Its largest root will therefore be given, approximately, by $\lambda=1+f\beta+g\gamma+h\delta+ks$, where f, g, h and k are chosen such that equation (4·7) is satisfied when quadratic terms in β, γ, δ and s are neglected. This gives

$$\lambda=1+\beta-0{\cdot}8\gamma-0{\cdot}2\delta+0{\cdot}3s. \tag{4·8}$$

Now from the linear equation for U', etc., given above it is easily seen that

$$U'+V'-W'-X'-Z'=(1+\tfrac{1}{2}\beta)\,(U+V-W-X-Z)=0.$$

The latent root $1+\frac{1}{2}\beta$ is therefore not relevant to the description of the initial changes in U, V, W, X and Z and is analogous to the latent root of unity always obtained in the complete treatment of the progress of an inbreeding system. The expression given by (4·8) is therefore the dominant latent root so that $\beta-0{\cdot}8\gamma-0{\cdot}2\delta+0{\cdot}3s$ may be considered as the effective initial selective pressure for the homostyle gene, at least to the first order in β, γ, δ, ϵ and s.

It is clear that for the situation described by the equations (4·3), when

$$\beta=\gamma=\delta=\epsilon=s=0,$$

there is no initial selective pressure for the homostyle gene which is therefore effectively neutral. It might therefore have been expected that there would be stability along the whole infinite continuum of possible equilibrium frequencies, as was in fact inferred numerically.

The homostyle gene is effectively neutral for values of β, γ, δ and s such that $\lambda_1=1$ or

$$\beta-0{\cdot}8\gamma-0{\cdot}2\delta+0{\cdot}3s=0. \tag{4·9}$$

When $\lambda_1<1$, the selection is against the homostyles so that they do not increase in frequency. Only for values of β, γ, δ and s such that $\lambda_1>1$ will the homostyles tend to increase when initially introduced into a population containing only pins and thrums. When $\beta=0{\cdot}1$, the value suggested by the population data, the viability of the $\boldsymbol{S^h s}$ homostyles c, is the dominant remaining factor in determining the threshold of homostyle increase.

A selection of values of the parameters which give the threshold, for which $\lambda_1 = 1$, is given in table 13. As s increases very low values of d are needed to give thresholds with the $\boldsymbol{S^h s}$ homostyle viability c at about 0·9. Even with values of γ at about 0·2 and of δ at 0·3 it will not be possible to have low initial selective values for the homostyle gene unless fewer than 40 % of the homostyles self-fertilize. Low initial selective values for the homostyles are, however, necessary to explain why the homostyles have arisen only now in Somerset and the Chilterns and not before, either there or elsewhere.

TABLE 13. VALUES OF THE VIABILITY PARAMETERS AND THE PROPORTION OF SELFING WHICH GIVE THE THRESHOLD OF NO HOMOSTYLE INCREASE

viability of $\boldsymbol{Ss}$ thrums	proportion of selfing	viability of $\boldsymbol{S^h s}$ homostyles $c = 1-\gamma$	viability of $\boldsymbol{SS^h}$ homostyles $d = 1-\delta$
$b = 1-\beta = 0{\cdot}9$	$s = 0{\cdot}1$	0·8875	0·8
		0·9111	0·7
		0·9375	0·6
	$s = 0{\cdot}15$	0·8688	0·8
		0·8938	0·7
		0·9188	0·6

5. A NUMERICAL INVESTIGATION OF THE POPULATION CHANGES

5·1. *The possible changes on the basis of the general equation* (4·2)

As a theoretical attack on the general equations (4·2) for the population model was not possible, it was decided to investigate them numerically, using EDSAC II, the electronic digital high-speed computer at the Mathematical Laboratory in Cambridge. The aim was to find out how variations in the various parameters affected the progress of the system and the final equilibrium, and see how the calculated curves corresponded with the observed configuration of counts discussed in § 3. The general method used was to calculate the expected frequencies over successive generations by repeated application of the equations (4·2), or modifications of them. The calculated frequencies could be tabulated at various generation intervals. An example of the form in which the results were obtained is given in table 14. The frequencies u, v, w, x and z are given in five columns for every tenth generation, for the case $a = 1$, $b = 0{\cdot}9$, $c = 1$, $d = 0{\cdot}7$, $e = 0{\cdot}7$, $s = 0{\cdot}1$ and $p = 0$, which we shall call data N. The initial frequencies are those obtained when 0·2 % of heterozygous homostyles are introduced into a population of pins and thrums in equilibrium for the relative viabilities assumed.

The case with no viability disturbance and $p = 0$ was first of all investigated. As indicated in the last section, it was discovered that a population with 100 % homozygous homostyles was obtained for all values of s, but that as s decreased the rate of approach to equilibrium decreased.

In general, the expression (4·8), giving the effective initial selective value of the homostyles indicates that the rate of approach to equilibrium will increase with s. An example of the effect of moderate variation in the value of s is given by table 15. This shows the number of generations taken for the homostyles to reach a frequency of 85 % for a selection of values of s, when the viabilities of the five genotypes are 1, 0·85, 1, 0·65 and 0·7, respectively, and

$p = 0$. The value of p was also found to affect the rate of approach to equilibrium. However, the curves of change are affected little by a change from an assumed thrum viability to the corresponding proportion of illegitimate self-fertilization needed for the pins. For example, in the model assumed by Crosby (1949) take the thrum viability $b = 0{\cdot}88$ and $p = 0$ instead of $p = 0{\cdot}1$ and $b = 1$. The corresponding curve of approach is hardly altered, but the equilibrium condition with about 21 % pins and no thrums, is reached after only thirty generations as compared with about 46 generations on Crosby's model. In subsequent models we shall always assume that $p = 0$.

TABLE 14. THE APPROACH TO EQUILIBRIUM OF THE FIVE GENOTYPE FREQUENCIES FOR DATA *N*, TABULATED AT EVERY TENTH GENERATION

genotype	ss	Ss	S^hs	SS^h	S^hS^h	
viability	1	0·9	1	0·7	0·7	$s = 0{\cdot}1$

u	v	w	x	z	generation
0·5255000	0·4725000	0·0020000	0·0000000	0·0000000	0
0·5247633	0·4709820	0·0036115	0·0005719	0·0000713	10
0·5229668	0·4678636	0·0077770	0·0012260	0·0001666	
0·5191710	0·4613048	0·0165303	0·0025809	0·0004131	
0·5114149	0·4480282	0·0342138	0·0052360	0·0011070	40
0·4965636	0·4230718	0·0673146	0·0099030	0·0031471	
0·4711762	0·3818028	0·1215756	0·0166545	0·0087909	
0·4344666	0·3251613	0·1947944	0·0238283	0·0217495	
0·3908785	0·2624400	0·2734903	0·0286726	0·0445187	
0·3476498	0·2049960	0·3423318	0·0299050	0·0751174	90
0·3099555	0·1587367	0·3944448	0·0283736	0·1084894	
0·2794236	0·1238850	0·4308864	0·0255361	0·1402690	
0·2555032	0·0982295	0·4555818	0·0224102	0·1682753	
0·2369311	0·0793192	0·4722711	0·0194830	0·1919956	
0·2224554	0·0651973	0·4836888	0·0169205	0·2117380	140
0·2110585	0·0544625	0·4916507	0·0147405	0·2280878	
0·2019720	0·0461472	0·4973240	0·0129052	0·2416517	
0·1946311	0·0395877	0·5014555	0·0113621	0·2529636	
0·1886238	0·0343254	0·5045274	0·0100605	0·2624629	
0·1836479	0·0300385	0·5068564	0·0089567	0·2705006	190
0·1794801	0·0264980	0·5086537	0·0080147	0·27735361	200

TABLE 15. THE EFFECT OF VARIATION OF s ON THE RATE OF APPROACH TO EQUILIBRIUM

genotype	ss	Ss	S^hs	SS^h	S^hS^h	
viability	1	0·85	1	0·65	0·7	and $p = 0$

value of s	no. of generations required for homostyles to reach 85 %
0·25	64
0·2	71
0·15	81
0·1	96
0·05	116

As indicated in the previous section, the rate of approach to equilibrium, more especially in the initial stages of progress, is particularly sensitive to changes in the value of the S^hs homostyle viability. For example, when $a = 1$, $b = 0{\cdot}85$, $c = 0{\cdot}93$, $d = e = 0{\cdot}75$ and $s = 0{\cdot}15$ after 200 generations the frequency of pins is 10·2 % and of thrums 0·3 %, the equilibrium being when pins are at 9·45 % and there are no thrums. However, when the value of c

is reduced from 0·93 to 0·83 the pin frequency reaches 5·3 % only after 800 generations, the equilibrium pin frequency being 4·1 %. The value for c at the threshold of no homostyle increase, calculated from equation (4·9), is 0·81 and when $c = 0{\cdot}82$ the frequency of pins changes by less than 1 % in a thousand generations. When $c = 0{\cdot}83$ the threshold is not reached until e, the $\boldsymbol{S^hS^h}$ viability, has been reduced to 0·6. This verifies that e only has a second-order effect on the threshold. However, even when $e = 0{\cdot}7$ the frequency of pins at equilibrium is 45·5 % and that of thrums 32·6 %. More generally, the value of e considerably affects the final equilibrium constitution even though it has little effect on the initial rate of progress.

It was pointed out in § 3 that the observed population counts showed a marked deficiency of thrums relative to pins for the higher homostyle frequencies. The curve of change for data N, superimposed on the observed population counts in figure 2 and tabulated in table 14, clearly does not show this deficiency. A decrease in the value of s causes an increase in the proportion of thrums relative to pins and, as was shown in the previous section, when $s = 0$ the proportion of pins and thrums are always equal. Decreasing the $\boldsymbol{Ss}$ thrum viability b tends to decrease the proportion of thrums relative to pins, but d, the viability of $\boldsymbol{SS^h}$ thrums, affects the progress very little. This is to be expected, as the $\boldsymbol{SS^h}$ thrums are produced mainly by crosses of thrums onto homostyles. As these rarely attain a frequency of more than a few per cent, they will contribute little to the population constitution. However, alterations in the values of b and s, in general, contribute only marginally to the tendency for a thrum deficiency.

In order to investigate the possibility of a large excess of pins over thrums in a population with a high homostyle frequency, the progress of the system when there were no thrums present was studied. It was found that the frequency of pins at equilibrium depended essentially on the ratio c/e of the viabilities of heterozygous and homozygous homostyles. When $c = e$ homostyles reach a frequency of 100 % so long as e is greater than about 0·5. As c/e increases so does the frequency of pins at the equilibrium. Table 16 shows the equilibrium pin frequencies for a selection of values of c and e. Variations in the value of s have no appreciable effect on the equilibrium constitution when there are no thrums present. Thus with $s = 0{\cdot}2$ and $c = 0{\cdot}95$, $e = 0{\cdot}6$ the frequency of pins at equilibrium is 20 %; with Crosby's model, for which $s = 1$, $c = 1$, $e = 0{\cdot}65$ and also 10 % of pins self-fertilize illegitimately, the frequency is 21 %. It is clear that very low values of e are needed in order to produce high pin frequencies. Values of the parameters giving such high proportions of pins at equilibrium when thrums are absent, would also be expected to be those most likely to give high pin to thrum ratios when thrums are present, at least on the basis of the general model so far considered. However, when thrums are introduced under these conditions they compete with the homostyles and it is thus not possible to reproduce an adequate thrum deficiency. For example, when $a = 1$, $b = 0{\cdot}9$, $c = 0{\cdot}93$, $d = 0{\cdot}7$, $e = 0{\cdot}7$ and $s = 0{\cdot}1$ the equilibrium constitution contains 36 % pins and 22 % thrums, whereas table 16 indicates that if thrums were absent there would only be about 14 % pins at the equilibrium.

In order to explain the fact that homostyles have attained appreciable frequencies only in Somerset and the Chilterns and not elsewhere, one must assume that in most areas where *Primula vulgaris* occur the effective initial selection for the homostyles is either negative or very small if positive. For populations which are near the threshold of homostyle increase, only a small change in the homostyle viability or the proportion of crossing could cause a critical increase in the effective initial selection for the homostyles, allowing these to increase and become relatively abundant. It is probable that variations in viability between different areas are small so that the populations in Somerset containing homostyles are still near the threshold. If this were the case, it would only need a correspondingly small change of viability in the reverse direction to lower the population to the threshold once again. That this might have happened in a few areas is indicated by the evidence for systematic decreases in homostyle frequency in some of the population counts discussed in § 3. The data from Sparkford Wood, south, given in table 4, indicate further that such a decrease might result in the sort of deficiency of thrums relative to pins which has been observed. This possibility was investigated numerically by calculating the progress from points near the equilibrium end of some of the calculated curves, after viabilities had been reduced in such a way as to give a negative initial selective value to the homostyles.

TABLE 16. THE EFFECT OF CHANGES IN c AND e ON THE EQUILIBRIUM FREQUENCY OF PINS WHEN THRUMS ARE ABSENT

equilibrium frequency of pins (%)	viability of S^hs homostyles: c	viability of S^hS^h homostyles: e
0·2	0·95	0·9
6·3	0·95	0·8
13·7	0·95	0·7
20·2	0·95	0·6
25·7	0·95	0·5
30·6	0·9	0·4
34·9	0·9	0·3

The value of s does not affect the pin frequencies to any appreciable extent.

It was shown in the previous section that the effective initial selective value for the homostyles is particularly sensitive to changes in the S^hs homostyle viability c. Results given earlier in this section indicate further that c has a considerable effect on the rate of progress to equilibrium. A small reduction in c therefore seems the most likely way in which the decrease in homostyles could be brought about. For the parameters assumed for data N, that is $s = 0·1$, $a = 1$, $b = 0·9$, $c = 1$, $d = 0·7$ and $e = 0·7$, the threshold is reached when c is reduced to 0·91. When $c = 0·9$ and a starting point near the end of the curve for data N at which there are 16 % pins and 1 % thrums is chosen, a constitution containing 26·2 % pins and 11·5 % thrums is attained after 80 generations. The curve 4 on figure 4 shows a regress from the same starting point, but with e reduced from 0·7 to 0·6. This reaches a constitution containing 37·1 % pins and 22·5 % thrums after only 40 generations. It is thus possible in this way to explain a sudden and fairly rapid decrease in the homostyle frequency. However, although showing a tendency towards the thrum deficiency, this curve still does not seem to explain adequately this aspect of the observed data.

5·2. *Modifications of the general model*

The model represented by equations (4·2) assumes that there is complete renewal of a population at each generation. Allowing for overlapping of generations is not expected in general, to give results which differ materially from those obtained on the assumption of no overlapping of generations. However, if viabilities from zygote to maturity differ appreciably from those of the mature plants, and only a relatively small proportion of the population die and are replaced in every breeding interval, the expected proportions of plants may differ considerably from that given by a simpler model. Consider, for example, the extreme situation, when no replacement by breeding is taking place. If the mature homostyles have a viability of 0·8 relative to the pins, the relative proportion of homostyles to pins would simply decrease by a factor of 0·8 in every generation, or, for example, a factor of 0·38 in five generations. So long as the rate of replacement is small, this effect of the viability of the mature plant may considerably outweigh the relative breeding advantage of the homostyles. In this case the proportion of homostyles may decrease without a corresponding increase in the proportion of thrums.

It is possible to investigate such a situation in more detail provided certain simplifying assumptions are made. We shall consider the life of the plant as consisting of two stages. The first is the time up to maturity, that is between the formation of the zygote and the time when it first produces fertile pollen and ovules. The second stage is the time after maturity. We shall further assume that the relative fertility and viability of a plant after reaching maturity are independent of its age. The time unit of change will be the interval between zygote formation and maturity and this can be considered as effectively discrete. This implies that breeding takes place only during a relatively short period and at the same time in each year, and also that a plant either contributes wholly or not at all to breeding at a given time. Then $\boldsymbol{u}_n$, the vector of the five genotype frequencies at time n, satisfies a recurrence equation of the following form

$$T\boldsymbol{u}_n = l(\alpha u_{n-1}, \beta v_{n-1}, \gamma w_{n-1}, \delta x_{n-1}, \epsilon z_{n-1}) + k\boldsymbol{S}(\boldsymbol{u}_{n-1}). \qquad (5{\cdot}1)$$

Here $\boldsymbol{S}(\boldsymbol{u}_{n-1})$ represents the effect of breeding on the population at time $n-1$, that is the vector of frequencies that would be given by the general equations (4·2), if there were complete renewal in a unit time interval. The relative viabilities a, b, etc., operating in the calculation of $\boldsymbol{S}(\boldsymbol{u}_{n-1})$ will therefore be those referring to the time up to maturity. The quantities α, β, etc., are the relative viabilities of the mature plants and l and k represent respectively average proportional survival rates and 'birth' rates. As before, T is such that $u_n + v_n + w_n + x_n + z_n = 1$. The consequences of this model can be investigated numerically in the same way as before. It is, of course, possible that the time taken by different plants to reach maturity will vary to some extent. However, this variation will only affect the model if it acts differentially on the five genotypes, and it may in general be expected to be small compared with the mature life of a plant.

The products $a\alpha$, etc., represent the average overall viabilities. Thus when $a = \alpha = \surd(a\alpha)$, etc., the model given by (5·1) should give the same curve of change as the simpler model gives with $a \equiv a\alpha$, etc. For example, when $a = \alpha = 1$, $b = \beta = 0{\cdot}95$, $c = \gamma = 1$, $d = \delta = 0{\cdot}84 = e = \epsilon$, the curve of change given by (5·1) is almost identical with that for data N. With $l = k = 0{\cdot}5$ the time taken for a change corresponding to one generation in data N, is 1·7 time units. The value of T at each stage of the progress gives a measure of the proportionate change in the size of population and is about 0·98, for example, given above.

In reproducing a sudden decrease in homostyles, only reductions in the viabilities after maturity were assumed to take place. The same starting point was used as for the simple regress described above. For example, with a, etc., and s taking the same values as for data N, and $\alpha = \beta = 1$, $\gamma = 0{\cdot}9$, $\delta = \epsilon = 0{\cdot}85$, and $l = 0{\cdot}7$ the proportion of pins present when there are about 3 % thrums increases from about 25 % when $k = 0{\cdot}4$ to 41 % when $k = 0{\cdot}1$. The corresponding values of T range from 1·1 to 0·8. The relative increase in pins is effected in only 20 units of time. The curve for the above situation when $k = 0{\cdot}2$ is shown in figure 4 as curve 3. Together with it are data N, the simple regress curve 4, and two further curves representing similar changes on the basis of equation (5·1). These curves show that after reaching the maximum pin to thrum ratio there is a sharp increase in thrums until, eventually, all the homostyles are eliminated. Clearly, it is possible on the basis of this model to represent any degree of thrum deficiency for relatively high homostyle frequencies and also to explain a sufficiently rapid decrease in homostyles to account for the observed counts in Sparkford Wood, south, given in table 4.

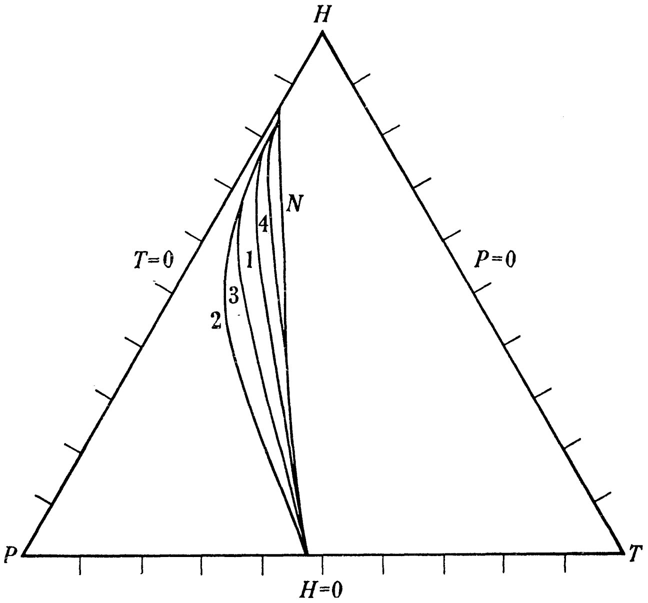

FIGURE 4. Comparison of four regress curves with progress on data N. For data N, $a = 1$, $b = 0{\cdot}9$, $c = 1$, $d = 0{\cdot}7$, $e = 0{\cdot}7$, $s = 0{\cdot}1$. 4 = simple regress with $a = 1$, $b = 0{\cdot}9$, $c = 0{\cdot}9$, $d = 0{\cdot}7$, $e = 0{\cdot}6$, $s = 0{\cdot}1$. 1, 2, 3 = regresses with partial replacement with parameters as follows:

	a	α	b	β	c	γ	d	δ	e	ϵ	l	k	s
1	1	1	0·9	1	1	0·9	0·75	0·85	0·75	0·85	0·7	0·3	0·1
2	1	1	0·9	1	1	0·9	0·8	0·85	0·8	0·85	0·6	0·1	0·1
3	1	1	0·9	1	1	0·9	0·7	0·85	0·7	0·85	0·7	0·2	0·1

where a, b, etc., represent viabilities up to maturity; α, β, etc., represent viabilities after maturity; l represents proportional survival rate; k represents proportional birth rate; and s is the proportion of homostyles which self-fertilize. All the regresses start from a point near the end of the curve for data N at which there are 16 % pins and 1·5 % thrums.

Constitutions reached after 20 time units

	pins	thrums	homostyles
1	25·6	3·8	70·6
2	36·4	2·6	61·0
3	32·2	4·4	63·4

The time unit of change is the interval between zygote formation and maturity.

It is clearly unrealistic to suppose that the viabilities and the proportions of crossing will remain constant over any appreciable period of time. Two models were therefore constructed to investigate the effects of such variation on the system. In the first model the viability of the S^hS^h homostyles was made to depend on the frequencies of the two sorts of homostyles. The relation assumed was of the form $e_n = e_0 + qw_n + rz_n$, e_n being the viability in the nth generation. Such a situation could conceivably occur, though it is on the whole more likely that the increase in viability should depend on the time as well as the change in population constitution. The curve G in figure 5 represents such a situation with viabilities 1, 0·85, 1, 0·65 and $e_n = 0{\cdot}4 + 0{\cdot}4w_n + 0{\cdot}5z_n$, $s = 0{\cdot}1$ and $p = 0$. This is compared with the curve for data N, that for Crosby's model, and the regress 2 given in figure 4. This latter curve is also superimposed on the observed population counts plotted in figure 2. For this curve the average

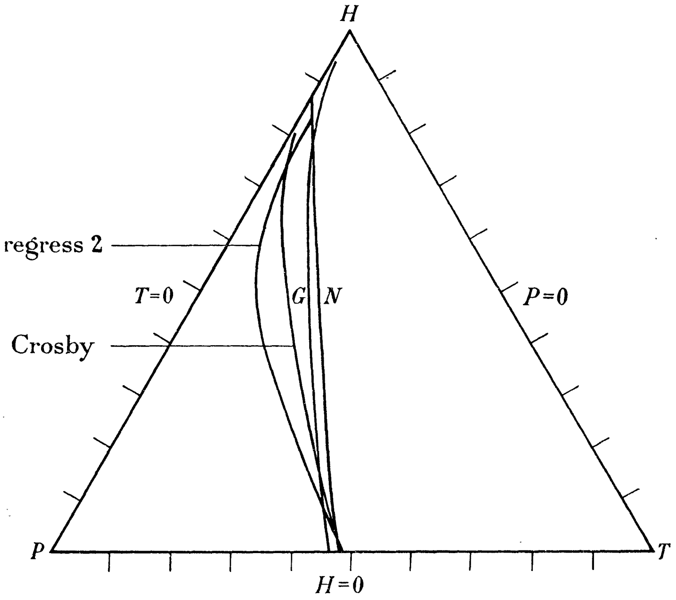

FIGURE 5. A comparison of the curves for Crosby's model and data N with two other curves. G represents a model for which the viability of the homozygous homostyles, e, increases according to the proportion of homostyles present. For this model $a = 1$, $b = 0{\cdot}85$, $c = 1$, $d = 0{\cdot}65$, $e_n = 0{\cdot}4 + 0{\cdot}4w_n + 0{\cdot}5z_n$ and $s = 0{\cdot}1$. For Crosby's model $a = b = c = d = 1$, $e = 0{\cdot}65$, $s = 1$, and 10 % of the pins self-fertilize illegitimately. The curves for data N and regress 2 are as indicated in figure 4.

value of T, which as described above represents the proportional change in size of the population, is about 0·68. This model thus represents a fairly rapidly diminishing population. The curve for Crosby's model represents a thrum deficiency intermediate between that for data N and those given by the partial replacement model described above. The G curve at first also represents a slightly greater thrum deficiency than data N, but as the value of e increases it crosses N arriving at a final equilibrium of 6·8 % pins and 93·2 % homostyles. It would, of course, be quite possible to let the homostyle viability increase further with time until the homostyles reached a frequency of 100 %.

The aim of the second model was to investigate the effect of variation from year to year in the value of s, as might be caused by differences in the weather conditions at the time of flowering. It was assumed that s was normally distributed with given mean and variance, so that in applying the equations (4·2) for each generation change, the value of s was chosen as a random deviate from a specified normal population. The populations were in fact 'condensed' at the values 0 and 1, s being taken as 0 when the chosen s was less than 0, and as 1 if the chosen $s > 1$. The results obtained, when taking s from a normal population with mean and standard error 0·1, and the other parameters as for data N, are very close to those for data N but tend very slightly toward a lower thrum frequency. The rate of approach to equilibrium is, however, a little faster. The effect of the random variation on the frequencies seems only to assert itself when the thrums have been eliminated from the population. The frequency of pins then varies by about 1 % either way about the average equilibrium value of 14·6 %. Investigation of similar models and even higher standard errors for s, confirmed the above situation, and showed that the deviation from the average expected curve was rarely more than about 1 %.

In considering the partial replacement model, described above, it was stated that overlapping of generations did not, in general, give results which differ materially from the simple discrete model. However, if the interval between breeding times is short compared with a plant's lifetime one may consider time to vary continuously. Then, provided differential viabilities do not alter appreciably with the age of a plant, the equations (4·2) with $u-u'$, etc., replaced by $\mathrm{d}u/\mathrm{d}t$, etc., will provide a good approximation to the situation in which generations overlap. The equations are now a set of first-order simultaneous differential equations, whose solutions u, etc., are functions of the time, giving the progress of the system. These equations were solved numerically on EDSAC II, by a step-by-step process, starting from a given set of initial frequencies. The initial frequencies were chosen, as before, to represent a population of pins and thrums in equilibrium for the chosen viabilities, with 0·2 % homostyles added to it. A comparison of parallel sets of results shows that the continuous time solution is remarkably close to the solution of the recurrence equations, a unit interval in the time corresponding closely to one generation. The main difference is a slight speeding up of the continuous solution during the period of most rapid increase of the homostyle frequency. It seems, therefore, that the overlapping of generations will not affect the conclusions drawn from an analysis of the discrete recurrence equations.

5·3. *The effect of random variations due to finite population size*

All the calculations described so far have been in terms of expected frequencies, and so do not take into account the random variations caused by sampling from a population of finite size. These variations can be effectively simulated by assuming that the observed proportions in a generation are chosen as a multinomial sample of size N (the population size), whose expected frequencies are given by applying equations (4·2) to the observed proportions in the previous generation. This can be done numerically by choosing N random numbers in the interval $0 < \xi < 1$ and assigning them to the five genotypes according to which of the five intervals

$$0 < \xi \leqslant u, \quad u < \xi \leqslant u+v, \quad u+v < \xi \leqslant u+v+w,$$
$$u+v+w < \xi \leqslant u+v+w+x \quad \text{and} \quad u+v+w+x < \xi \leqslant 1$$

they lie in. Here u, v, etc., are taken to be the expected proportions, calculated from the observed proportions of the previous generation. This process can easily be carried out on an electronic computer, using one of the standard methods for generating random numbers.

The effect of finite population size was studied in this way for data N and for the simple regress obtained from data N when c, the $S^h s$ homostyle viability, is reduced from 1 to 0·9. For data N a starting point on the expected curve, where the population contains about 12 % homostyles, was chosen. This was done so that the chance of losing the homostyles would be small, as the aim was to study the progress of an increase in the homostyles. Even so, with a population size of $N = 50$ the homostyles were lost in three trials out of four which extended over 100 generations. Typical examples of trials with three population sizes, 50, 200 and 500 are shown in figure 6 together with the expected curve for data N. The lines join points at intervals of ten generations. The effect of finite population size is clearly least in the middle range of increase, where the effective selection in favour of homostyles will be greatest. When the rate of increase has slowed down and the frequency of thrums is low, extinction of the thrums may be considerably hastened. The variation tends to alternate on either side of the expected curve, but may conceivably contribute to the thrum deficiency for high homostyle frequencies. However, there is no evidence for the thrum deficiency being associated with small population size, as might be expected if this were its sole cause. The rate of increase of homostyles is appreciably enhanced by the earlier extinction of the thrums caused by the random variation. For the example given in figure 6 with $N = 50$, the thrums are lost after only forty generations when their expected frequency is still about 16 %. Thereafter the frequencies oscillate about a mean of 15·1 % for the pins corresponding closely with the equilibrium frequency of 14·7 % given by data N. The range of variation in pin frequency is from 6 to 28 %.

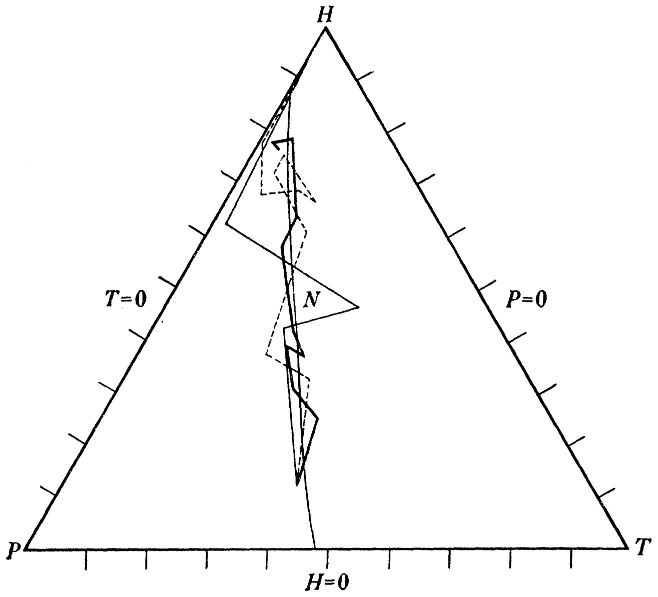

FIGURE 6. The effect of random variation due to finite population size on progress with data N. The lines represent typical examples of trials with the three population sizes indicated: ——, 50; - - - -, 200; ——, 500. They join points at intervals of ten generations.

In studying the effects of variation on the regress from a high homostyle frequency, a starting point containing 4·5 % thrums was chosen. Nevertheless, with $N = 50$ the thrums were retained and increased in frequency in only one trial out of six which extended over twenty generations. However, they were not lost in any of several trials over the same number of generations with $N = 200$ and $N = 500$. Examples of trials, with the three population sizes 50, 200 and 500, together with the expected curve for the simple regress for the case, when c is reduced from 1 to 0·9 in data N are given in figure 7. The lines join points at intervals of two generations. The general effects of the variation caused by finite population size are similar to those discussed above for the progress to high homostyle frequency with data N.

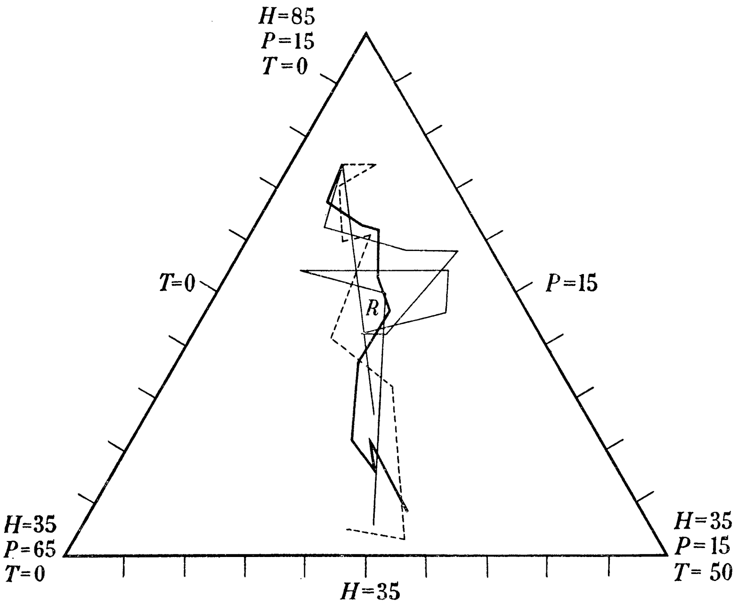

FIGURE 7. The effect of random variation due to finite population size on a simple regress from a point near the end of the curve for data N. The expected line for the simple regress R is that obtained from data N when the $S^h s$ viability c is reduced from 1 to 0·9. The lines represent typical examples of trials with the three population sizes indicated: ——, 50; - - - -, 200; ——, 500. They join points at intervals of two generations.

6. DISCUSSION

The most important differences brought out by both the theoretical and the numerical investigations are in the rate of approach to equilibrium. In other respects the curves given for the straightforward progress represented by equations (4·2) seem fairly stable even for considerable changes in the values of the relative viabilities of the five genotypes. The rate of approach to equilibrium, and more especially the effective initial selective value of the homostyle gene, are clearly factors of some importance in the proposed evolutionary process.

It is difficult to draw any satisfactory conclusions as to the rate of progress, from the counts of populations extended over several years as given in § 3, without a better knowledge of the correspondence between generations and years. The effective generation, under the experimental conditions of growth at Cambridge, is 2 years, and it is to be expected that this will be a considerable underestimate of the length of a generation under natural conditions. Supposing that counts over about 10 years represent about two generations' change, for populations containing between 10 and 70 % homostyles the homostyle frequency should have increased by about 7·5 % on Crosby's model and 0·8 % on data N, over this time. The only two counts in this middle range of homostyle frequency are those from Sparkford Wood, north-west, and Laurel Copse given in tables 5 and 7. Both show significant heterogeneity, but no upward trend in the homostyle frequency. As pointed out in § 3, this heterogeneity is probably due to slight variation from year to year in the location of the count. Such heterogeneity would be expected to be most severe for the populations in this middle range of homostyle frequency, as small differences in the stage of the evolutionary process reached by the various regions of the population would be magnified by the more rapid rate of change of the homostyle frequency. However, even a change of 7·5 % might prove difficult to detect. If the generation time is in fact much longer than 5 years, counts over a very long period of time will be needed to detect any change even for a comparatively rapid rate of approach to equilibrium.

If, as seems reasonable to suppose on the basis of Ernst's work (1933), the homostyle gene complex $\boldsymbol{S^h}$ arises from a crossover between $\boldsymbol{S}$ and $\boldsymbol{s}$ complexes, such $\boldsymbol{S^h}$ genes must have arisen many times before in the history of the Somerset and other populations, even supposing the crossover rate is of the order of a mutation rate. Now Fisher (1930) has shown that if a is the selective value for a gene, and is small, the chance of ultimate survival for a single newly occurring gene is approximately $2a/(1-e^{-4Na})$, where N is the population size. The value of N must be at least 100 for most of the populations counted, and is more likely to be of the order of 500. Thus for selective values greater than about 0·01, $Na > 5$ and the factor e^{-4Na} can be neglected. On the basis of Crosby's model, the initial selective value calculated numerically is 0·26, which gives an ultimate chance of survival for a single homostyle gene of about 0·52. Although in this case the selective value is not small, it is, according to Fisher (1930), unlikely that a more exact treatment for values of a which are not small, would give results differing essentially from those for small selective values. Assuming therefore this value for the chance of ultimate survival of a single gene, the homostyle gene need only have arisen six times for a minimum certainty of 1/50 of the ultimate survival, and hence rapid increase, of the homostyles. Even for data N, giving an initial selective value of 0·07, the homostyle gene need only have arisen about twenty-six times for a minimum certainty of survival of 1 in 50. It is clear therefore that, as pointed out in previous sections, for populations other than in the Somerset and Chiltern areas a considerably lower selective value is needed to account for the fact that such populations do not contain appreciable numbers of homostyles. When there is no crossing, and $p = 0{\cdot}1$ as Crosby assumes, the initial selective value decreases to zero with e. Thus even when the homozygous homostyles are completely eliminated the homostyle gene has the same chance of survival as that of a neutral gene in a randomly mating population. As pointed out at the end of § 4, only with fairly high frequencies of cross-fertilization can very low or negative initial selective values for the homostyle gene, reasonably be obtained.

It was shown in the previous section, that the large deficiency of thrums relative to pins for fairly high homostyle frequencies was difficult to explain on the assumption that the observed counts represent a consistent upward trend in the homostyle frequency as described by the theoretical models considered. Certainly, if such an assumption were true, a remarkably low viability for the homozygous homostyles would be indicated whether crossing occurred or not. It seems improbable that moderate disturbances not considered in the model, but of a similar nature to the viability disturbances such as fertility or compatibility differences, could adequately account for the thrum deficiency. The discussion of fertility differences in § 4 indicates that these will on the whole affect the system in a similar way to viability disturbances. In fact, if the ovule and pollen fertilities for each of the five genotypes are equal, the results obtained can be taken to represent the relative gametic frequencies, and must be multiplied throughout by the relative fertilities to obtain the zygotic frequencies. Such a disturbance could not explain the thrum deficiency occurring only for fairly high proportions of homostyles.

A low viability for the homozygous homostyles might be expected, at least in the initial stages of a process, for the reasons suggested by Crosby (1949). Thus Mather & De Winton (1941) have shown the thrum homozygote (***SS***) in *P. sinensis* to have a viability of 72·8 % relative to the ***Ss*** thrums and the pins (***ss***). They attribute this to a poor internal balance of polygenic combinations on the ***S***-carrying chromosome, which has arisen after the evolution of an efficient outbreeding system precluding the occurrence of the homozygous thrums. It is quite likely that a similar situation exists for the homozygous thrums of *P. vulgaris*. Now Ernst (1933) has shown that homostyle complexes, such as the one in *P. vulgaris*, are probably the result of a very rare recombination between the pin (***s***) and thrum (***S***) complexes. Hence the S^hS^h homostyle, being homozygous for a part of the ***S***-carrying chromosome may be expected to show a similar lack of internal balance, and so viability deficiency. However, once the S^hS^h homostyle occurs with any appreciable frequency, the situation favouring such a lack of balance will no longer hold. Selection may be expected to act in such a way as to increase the viability of the S^hS^h homostyles and a situation such as that represented by the curve *G* discussed in the previous section, and illustrated in figure 5, might be expected. The open-pollination data suggest a very high proportion of crossing when the homozygous homostyles have a low relative viability. Suppose, for example, that in the natural populations $e = 0{\cdot}7$, and the viability is increased by 10 % under artificial cultivation so that for the open pollination data we would take $c = 1$ and $e = 0{\cdot}8$. The proportion of crossing estimated by the methods of § 2 is then 93·82 %.

It is possible that some systematic bias might have affected the collection of the data. Thrums and homostyles are not immediately distinguishable, and there may, in rare cases, be a normal overlap of the two forms of flower. There is, however, no reason to expect such misclassification consistently to favour the homostyles. Alternatively, there may be differences in the flowering time of the three forms of flower, as suggested by Crosby (1949). If the thrums flowered late, this might account for such a deficiency, but it is difficult to imagine why any such bias should only affect counts from populations with high frequencies of homostyles.

A model in which homostyles were fertilized by homostyles with greater frequency than by thrums has also been investigated by the methods of the previous section. When the relative frequency of thrum to homostyle pollen fertilizing the homostyles was 0·8, the effect on the curve for data *N* was a shift of never more than 1 % towards the line representing absence of thrums. The rate of approach to equilibrium was slightly increased owing to the more rapid extinction of the thrums. However, even values of 0·5 for the relative fertilization have little effect. Data given by Crosby (1959) suggest the possible existence of such a disadvantage of the thrum pollen on homostyle stigmas. A high proportion of crossing was found by Bateman (1956) in the wallflower, which had previously been assumed to self-fertilize. He showed, by mixed pollination, that this was caused by a mild incompatibility reaction when self-pollen had to compete with foreign pollen on its own style, as under open pollination. It is not inconceivable that such a mild form of incompatibility may place thrum pollen at a disadvantage when competing with homostyle pollen on a homostyle stigma.

In the previous section it was shown that the thrum deficiency could be easily explained by a slight reduction in the relative viability of the mature homostyle plants, and a consequent reduction in the frequency of homostyles. The rates of change of frequencies for the regresses on the partial replacement model described by equation (5·1) far exceed those obtained for a simple progress to equilibrium, especially if, as seems likely, a *Primula* generation extends over a number of years. Only such a change could account for the decrease in homostyle frequency observed in Sparkford Wood, south (table 4). Moreover, the fact that neighbouring populations may differ widely in the proportion of homostyles they contain, could also be reasonably accounted for by such a delicately balanced situation.

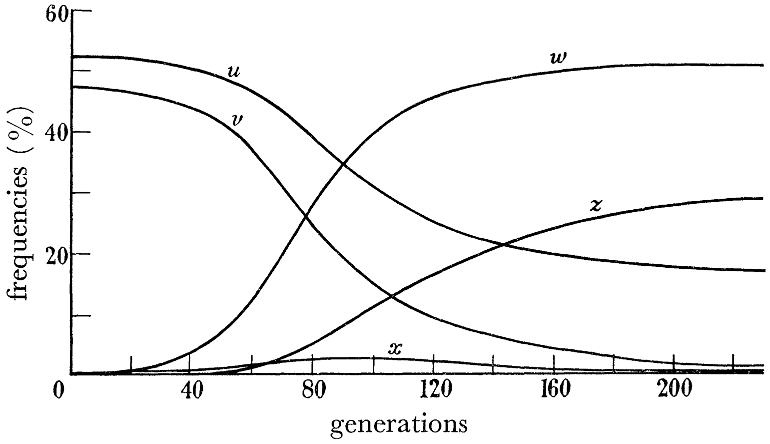

FIGURE 8. The approach to equilibrium of the five genotype frequencies for data *N*.

genotypes	ss	Ss	S^hs	SS^h	S^hS^h	
viabilities	1	0·9	1	0·7	0·7	$s = 0·1$

Crosby (1948) has described a smut fungus, *Tuburcinia primulicola*, which attacks *Primula vulgaris* and occurs in the Somerset area. The flowers of smutted plants are deformed, their anthers are full of conidia and ripe ovaries usually contain brand spores and probably no seeds. It is not improbable that the long homostyles, having both anthers and stigma in the mouth of the corolla, are more susceptible to an attack by the smut than either pins or thrums. An increase in the smut fungus might therefore be accompanied by a decrease in the relative viability of the mature homostyles, sufficient to account for their decrease in frequency and the consequent thrum deficiency.

On the assumption that the thrum deficiency is caused by such a reversal in the upward trend of the homostyle frequency, an excess of populations in the middle range of homostyle frequencies might be expected. The frequencies of the five genotypes for data *N* are plotted against the number of generations in figure 8. This shows that the middle range is the region of fastest change, and hence the region in which least populations would be expected to fall if there were no decrease in homostyle frequency. There is certainly, as pointed out in § 3, a concentration of populations near the end-points of the supposed change, although in well over a third of the counts the homostyle frequency lies between 10 and 65 %. It is difficult to compare this with any expected density without detailed knowledge of the time scale of the evolutionary process, and the range of time which the observed counts represent. However, if we assume that the counts of populations with homostyle frequencies from 10 to 65 % are approximately randomly distributed in time, a rough comparison of the observed density with expected densities of counts can be made. The number of counts expected between two homostyle frequencies will then be proportional to the number of generations required for an increase from the lower to the higher homostyle frequency. This will tend to overestimate the number of counts in the middle region if the tendency for the homostyles to increase is not of comparatively recent origin. The observed number of counts for the frequency ranges 10 to 30 % and 30 to 65 %, and the number of generations required for the changes on Crosby's model and data *N* are given in table 17. The observed proportion does in fact differ significantly from the expected proportions on Crosby's model, and is lower than that for data *N*, though not significantly so. Thus, in spite of the unsatisfactory assumptions made, there is suggestive evidence for an increased density of counts in the regions where this might be expected. Crosby's (1949) data also shows an excess of counts in the middle range of homostyle frequency, but he claims that this is because 'In deciding where to count populations I often aimed at districts where change might be expected to be most rapid'.

TABLE 17. A COMPARISON OF OBSERVED AND EXPECTED NUMBERS OF COUNTS IN THE MIDDLE RANGE OF HOMOSTYLE FREQUENCY

homostyle frequency range (%)	observed number of counts	number of generations required for the change	
		Crosby model	data N
10 to 30	13	6	23
30 to 65	35	9	48
observed proportion	0·2708	0·4	0·3239
fiducial limits	0·3990 to 0·1426	expected proportions	

Bodmer (1958) showed that a probable reason for such a high frequency of crossing amongst the self-compatible homostyles was a tendency for these to be protogynous, although a weak incompatibility of the type discovered by Bateman (1956) in the wallflower is by no means out of the question. Three possibilities were suggested with respect to this tendency, viz.

(*a*) It is the residual of a tendency that exists in conjunction with the complete illegitimacy system.

(*b*) It is a tendency which has been selected for as the proportion of homostyles increased, because of the advantage of outcrossing.

(*c*) The tendency existed with the complete illegitimacy system and has been further selected for as the proportion of homostyles increases.

The conclusions reached above strongly favour the first of these possibilities. For if the tendency to cross did not exist with the illegitimacy system, the great initial selective advantage of the homostyles even for low viabilities, should have led to the widespread occurrence of the homostyles and so the loss of the outcrossing mechanism.

Observations of the relative growth of stigma and anthers made on the 1958 progenies confirm the results of earlier experiments reported by Bodmer (1958). In the later experiments the time at which the flowers were fully open was recorded as well as the time of dehiscence. It was found that the average number of days between the flowers being fully open and the anthers dehiscing was 0·975 based on observations of 490 flowers. Similar observations were also made on 67 pin flowers and these showed an average delay in dehiscence of 1·149 days. This confirms the suggestion that the tendency to cross-fertilize existed in conjunction with the complete illegitimacy system and so should be exhibited in the pin flowers. It seems, therefore, that the tendency of the homostyles, and possibly the pins and thrums, to cross-fertilize is an essential buffering mechanism against the reversal of the outcrossing system to a return to inbreeding.

7. Conclusions

On Crosby's (1949) assumption that homostyle primroses only self-fertilize, newly occurring homostyles are at a considerable selective advantage over pins and thrums. There is not then any possible controlling mechanism which prevents homostyles from sweeping through the primrose populations as soon as they arise. The discovery by Bodmer (1958) that a high proportion of homostyles may cross-fertilize under natural conditions seems to provide the answer to the existence of such a controlling mechanism. When there is a high frequency of crossing amongst the homostyles only moderate viability disturbances are needed to prevent them from increasing in frequency. The balance between the relative viabilities of the different occurring genotypes and the proportion of crossing provides a delicate control over the occurrence of the homostyles. The evidence from population counts for decreases in homostyle frequency suggests that the homostyle populations may in fact be near the threshold of no homostyle increase. A small decrease in the relative viability of mature homostyles may then cause a very rapid decrease in their frequency and lead to the characteristic deficiency of thrums for high homostyle counts shown by the observed data.

It is probable that the tendency of the homostyles to cross-fertilize exists also in the pins and thrums. Hence it seems likely that some form of relative self-incompatibility was in existence before the final evolution of the incompatibility switch mechanism had taken place. Otherwise there would be no control over the occurrence of the homostyles in the early stages of the incompatibility system and so no control over the almost immediate loss of the newly developed outcrossing mechanism. These conclusions support the type of gradual evolution of the incompatibility system suggested by Mather (1950).

Data on the occurrence of homostyles at low frequency in normal pin-thrum populations have been given by Ford (1955). Thus when the opportunity arises extremely rapid changes in primrose population constitutions may take place. The significance for the study of population genetics of being able to follow such rapid evolutionary changes has been emphasized by Ford (1955). When coupled with more detailed ecological investigations the study of the changes in the *Primula* populations should provide valuable information on the selective forces controlling changes in incompatibility systems such as exist in the *Primula*.

I would like to thank Professor Sir Ronald Fisher, F.R.S., for his suggestions and many helpful discussions and for making his records of population counts available to me. I am also grateful to Dr A. R. G. Owen for his general advice and encouragement in the course of this investigation. I would further like to thank the Director of the Mathematical Laboratory, University of Cambridge, for permission to use the high speed electronic digital computer EDSAC II. Part of this work was done during the tenure of an Agricultural Research Council Studentship.

REFERENCES

Bateman, A. J. 1956 Cryptic self-incompatibility in the wallflower: *Cheiranthus cheiri* L. *Heredity*, **10**, 257–261.

Bodmer, W. F. 1958 Natural crossing between homostyle plants of *Primula vulgaris*. *Heredity*, **12**, 363–370.

Crosby, J. L. 1940 High proportions of homostyle plants in populations of *Primula vulgaris*. *Nature, Lond.* **145**, 672.

Crosby, J. L. 1948 Population genetics in the genus *Primula*. Dissertation for the degree of Ph.D., University Library, Cambridge.

Crosby, J. L. 1949 Selection of an unfavourable gene complex. *Evolution*, **3**, 212–230.

Crosby, J. L. 1959 Outcrossing in homostyle primroses. *Heredity*, **13**, 127–131.

Ernst, A. 1933 Weitere Untersuchungen zur Phänanalyse zum Fertilitäts problem und zur Genetik heterostyler Primeln. I. *Primula viscosa*. *Allg. Arch. J. Klaus. Stift.* **8**, 1–215.

Fisher, R. A. 1925–54 *Statistical methods for research workers*. Edinburgh: Oliver and Boyd.

Fisher, R. A. 1930 *The genetical theory of natural selection*. Oxford University Press.

Fisher, R. A. 1949 A theoretical system of selection for homostyle *Primula*. *Sankhya*, **9**, 325–342.

Ford, E. B. 1955 Rapid evolution and the conditions which make it possible. *Cold Spr. Harb. Symp. Quant. Biol.* **20**, 230–238.

Mather, K. 1950 The genetical architecture of heterostyly in *Primula sinensis*. *Evolution*, **4**, 340–352.

Mather, K. & De Winton, D. 1941 Adaptation and counter-adaptation of the breeding system in *Primula*. *Ann. Bot., Lond.* (N.S.), **5**, 297–311.