Heredity (1981), 46 (1), 69-79

© 1981. The Genetical Society of Great Britain

THE POPULATION GENETICS OF THE SELF-INCOMPATIBILITY POLYMORPHISM IN PAPAVER RHOEAS. I. THE NUMBER AND DISTRIBUTION OF S-ALLELES IN FAMILIES FROM THREE LOCALITIES

J. M. CAMPBELL and M. J. LAWRENCE Department of Genetics, University of Birmingham, Birmingham B15 2TT

Received 30.vi.80

SUMMARY

Though the self-incompatibility polymorphisms are as well-known as any and have been the subject of much theoretical investigation, little is known about their characteristics in natural populations.

A preliminary survey of the extent of the polymorphism in *P. rhoeas* found 32 different *S*-alleles among eighteen full-sib families derived from natural populations growing in three widely separated localities in the West Midlands. Though most of these alleles occurred in one or other of these populations only, an analysis of their distribution between localities failed to show any significant differentiation between them.

These results are compared with those from other species and reasons are given for the apparent lack of differentiation between localities.

1. INTRODUCTION

ONE of the most attractive features of the self-incompatibility polymorphisms is that it is possible to deduce most of their basic properties from a knowledge of the inheritance of the system alone. Thus in those species in which self-incompatibility is determined by a single, multi-allelic gene, S, and in which control of the pollen phenotype is gametophytic (the Nicotiana system),

(i) a population cannot persist with less than three alleles;

(ii) because all individuals are heterozygous and the alleles act independently in the style or stigma, the number of different genotypes or morphs where k alleles are present is k(k-1)/2;

(iii) the selective advantage of an allele is negatively related to its frequency in the population, that is, the polymorphism is maintained by frequency dependent selection;

(iv) from (iii), it follows that the number of different alleles in a population is potentially very large;

(v) provided that the effect of selection on the locus is limited to that associated with incompatibility, the equilibrium frequency of each of the k alleles present in the population is 1/k; and

(vi) the selective advantage of a new allele that appears in the population by mutation or migration is negatively related to the number of alleles already present, a deduction which in practice may limit the total number of alleles that can be maintained even in a very large population.

Deductions (i), (iii), (v) and (vi) were first made by Wright (1939) and (ii) is formal.

In any actual population, sampling effects will affect both the number and frequency of the S-alleles present. Thus:

(vii) since there cannot be more than 2N different alleles in a population of size N (*i.e.*, $k \leq 2N$), any natural population that is established from a small number of individuals or is greatly reduced in size from time to time will contain only a sub-set of the alleles of the species; indeed, because of drift, the number of alleles that can be maintained in a population is very much less than N (see Wright, loc. cit., fig. 2); and

(viii) while, again, because of drift, the allele frequencies at equilibrium will not be equal to 1/k, according to Wright they will nevertheless be close to this value; however, in any population whose habitat is subject to gross disturbance from time to time, allele frequencies will depart quite considerably from approximate equality in the short term at least;

(ix) lastly, if the effect of selection on the locus is not, in fact, limited to that associated with incompatibility, because of linkage or pleiotropy, the frequencies of advantageous alleles will clearly be greater than 1/k at equilibrium and vice versa. There is some evidence from *Nicotiana* (East and Mangelsdorf, 1925; East and Yarnell, 1929) that pollen carrying certain *S*-alleles may grow more quickly down the style than that carrying other alleles. If this phenomenon should occur in natural populations, its effect on allele frequencies could be similar to that of chance in (viii).

These deductions and arguments leave no doubt that any experimental investigations of the self-incompatibility polymorphism has at the outset rather more information than is generally available for other polymorphisms. Yet despite this advantage, little is known about the characteristics of the polymorphism in natural populations. Indeed, such knowledge as we have is almost wholly derived from the pioneering study of Emerson (1938; 1939; 1940; 1941) on the number and distribution of S-alleles in Oenothera organensis, a species which appears to be restricted to the Organ Mountains of New Mexico and which Emerson estimated consisted "of less than one thousand and very likely less than five hundred" individuals. Yet ultimately, Emerson found no less than 45 different alleles in this species. In considering the quite remarkable extent of the polymorphism in this species, Wright (loc. cit.) showed that a population of 500 individuals, assuming that the population was in equilibrium, that the individuals it contained bred at random and that the effective mutation rate was of the order of 10^{-6} per generation, could be expected to maintain only about 13 alleles. Of the three hypotheses put forward by Wright to account for the excess in the number of alleles, two, those proposing a high mutation rate at the locus and a sub-division of the population into a number of more or less isolated sub-populations, are ruled out by the evidence (Emerson, 1939; Lewis, 1948; 1951; Fisher, 1961). The third hypothesis, which proposed a recent decline in the size of the population and that the loss of alleles had occurred at too slow a rate to have reached equilibrium (Wright showed that a population of 5000 individuals would be required to maintain 45 alleles), while on general grounds the most likely of the three, is nevertheless the weakest because it cannot be put to the test in the same way as the other hypotheses. Thus, until very recently, no completely satisfactory explanation has been given for the extent of the polymorphism in this species. Indeed, most of the literature on the population genetics of the self-incompatibility polymorphism in *Oe. organensis* is devoted to criticisms

of and arguments about the mathematical validity of Wright's analysis, first by Fisher (1958), to which Wright replied in 1960, and secondly by Moran (1962), to which Wright replied in 1964. This matter was finally settled by computer simulation of the problem independently by Crosby (1966), Ewens and Ewens (1966), Mayo (1966) and Kimura and Maruyama (1966), each of whom showed that Wright's conclusions were essentially correct. However, very recently, Levin, Ritter and Ellstrand (1979) have published new evidence on the *Oe. organensis* population which leaves little doubt that the species behaves as a panmictic unit and that the population probably consists of some 5000 plants, a number which is sufficiently large to maintain 45 *S*-alleles. Thus after 40 years, the problem of the number of selfincompatibility alleles in the *Oenothera* population appears at last to have been solved.

Two further investigations of the polymorphism that deserve mention are Atwood's (1944) study of two natural populations of *Trifolium repens* in Pennsylvania in which he found 36 different alleles out of a total of 49 examined in one and 39 out of 49 different S-alleles in the other population; and Williams and Williams' (1947) analysis of two non-pedigree strains of *T. pratense* in which he found 41 out of 48 and 35 out of 38 different alleles respectively. In both studies, compatibility was assessed on seed-set, which, as Emerson showed, is less reliable than the observation of pollentube growth on the stigma or in the style, and in neither of these studies was any attempt made to cross-classify the alleles from one population with those of the other. These data are less satisfactory for our present purpose, therefore, than Emerson's on *Oenothera*.

The work on these three species represents the sum total of our knowledge of the polymorphism in natural populations of those species which have the *Nicotiana* system of self-incompatibility. It is clear, therefore, that theoretical investigation of the polymorphism is in advance of its experimental investigation and that more information is required from the latter if the theory is to be developed on a sound basis.

It was with this argument in mind that an investigation of the selfincompatibility polymorphism in the Field Poppy, *Papaver rhoeas* was begun some years ago. Previous work has shown that the genetical control of self-incompatibility in this species is the same as that in *Oenothera* and *Trifolium* (Lawrence, 1975; Lawrence, Afzal and Kenrick, 1978). In the present paper we consider the number and distribution of *S*-alleles in eighteen families of the species whose ancestors were obtained from three localities in the West Midlands.

2. MATERIALS AND METHODS

The eighteen full-sib families used in the investigation originated from three localities, these being Wellesbourne, Warwickshire (R102; 6 families), Broad Oak, Herefordshire (R104; 5 families) and Hackmans Gate, Worcestershire (R106; 7 families). Broad Oak is some 53 miles from Wellesbourne and 42 miles from Hackmans Gate, the distance between the latter pair of localities being 26 miles. Each family has been obtained by from one to four generations of sib-mating among the descendants of a single non-inbred family of full-sibs that had been chosen at random from the families of an intra-locality, North Carolina Design 2 experiment (Comstock and Robinson, 1952) which was originally grown for another purpose (Ooi, 1970). Though none of the semi-inbred families had originated from the same pair of ancestors as another, some were drawn from the same male or female array in which case their ancestors were related as half-sibs.

The data with which we are concerned were obtained over a period of three years during the winter seasons of 1974 to 1977 from plants grown under supplementary illumination in the glasshouse. In each season, the task of identifying the genotypes that were present in the families concerned was tackled in two stages. In the first pollinations were made between the plants in each family in a chain-wise fashion (Lawrence *et al.*, 1978) in order to determine their compatibility class. Once this had been done, pollinations were then made between plants from each class of a family with those of every other family in the experiment in order to ascertain their genotype. Compatibility was determined by observing the reaction of pollen on the stigma twenty-four hours after pollination by the aniline-blue fluorescence method of Martin (1959) as modified by Lawrence (1975) and Lawrence *et al.*, (loc. cit.).

Because P. rhoeas is a strictly annual species it was necessary to make crosses for seed so that the alleles identified in one season could be unambiguously re-identified in the next. This was done by making reciprocal crosses between two plants that were known to be half-compatible. Suppose that the parental genotypes are S_1S_2 and S_1S_3 . Then if the direction of the cross is as written, half of the offspring are expected to be S_1S_3 (class II) and half S_2S_3 (class III); for the reciprocal cross, half of the offspring are expected to be S_1S_2 (class I) and half S_2S_3 (class III). The identity of plants of the third class, which is expected in both families, is revealed if pollinations are made between a plant from each class of the first family with one from each class of the second. Of the four pollinations, only one, that between an S_2S_3 plant from the first family and one of the same genotype from the second family, is expected to be incompatible, the remaining three being half-compatible. In this way, plants of class III are identified directly in each family and those of classes I and II are identified indirectly, via their half-compatible relationship with the former. As many crosses for seed were made as were necessary to carry forward all the S-alleles identified in one season to the next. Such families are referred to as tester families. Further details of both materials and methods can be found in Campbell (1977).

3. Results

(i) Number of alleles found and repeatability, R

A summary of the results obtained from the eighteen families examined, which are based on over 2350 pollinations, is shown in table 1. In all, 32 different S-alleles were found in these families. However, since 11 of these families were of the two-class and 7 of the four-class type, the total number of alleles examined and hence the number of different alleles that could have been identified in the experiment is $(11 \times 3) + (7 \times 4) = 61$. This number is nearly twice the number of different alleles actually found. It is clear, therefore, that many of these alleles occurred in more than one family.

TABLE 1

Summary of the results obtained over three seasons showing the S-alleles identified in the eighteen families. Four alleles, S_5 , S_{20} , S_{22} and S_{31} do not appear in the list because they were shown to be identical to others at a late stage in the analysis. $(15^* = S_{15} \text{ or } S_{17}, \text{ this being an ambiguity in the classification that was not resolved when the experiment was terminated.}$

Locality	Family		S-all	eles	
R 102	1	15	17	21	
	2	4	16	19	
	7	4	13	17	
	8	3	7	9	21
	9	23	24	33	
	10	19	25	26	
R 104	13	27	28	34	
	18	4	27	29	30
	21	4	12	27	
	23	4	19	35	36
	24	13	15*	27	32
R 106	26	1	3	4	
	27	3	12	13	14
	28	11	15	16	
	29	15	16	17	18
	31	1	6	7	8
	34	1	2	3	
	36	9	10	11	

It is convenient to summarise this finding in the form of what may be called a repeatability statistic, R. Let n be the number of different alleles found and m the number examined $(n \le m)$. Then allowing for the fact that any population must contain a minimum of 3 S-alleles if it is to persist (*i.e.*, $N \ge 3$).

$$R = 1 - \frac{n-3}{m-3} = \frac{m-n}{m-3}.$$

If the number of alleles found is the same as the number examined, no repeats have occurred and R = 0. If, on the other hand, only 3 alleles are found, R = 1, irrespective of the number examined. More particularly, as $R \rightarrow 1$, $n \rightarrow N$. Thus R provides a measure of the thoroughness with which an analysis has been carried out.

Now in the present case, n = 32 and m = 61, so that R = 0.5. However, it will be recalled that some of the ancestors of the families that we have examined (those from the same locality) had a half-sib relationship because they originated from the same array of a North Carolina Design 2 Experiment. Unfortunately, because this experiment was raised for another purpose, no record was kept of the origins of these ancestors. It is thus possible (though perhaps unlikely) that all repeats over families from the same location are due to this cause alone. If repeats within locations are ignored, m = 45 and R = 0.31.

(ii) The distribution of S-alleles between localities

While there may be some doubt about the suitability of these data in respect of the distribution of alleles within localities, their distribution between localities is beyond suspicion, because the ancestors of families from different locations were drawn from different North Carolina Design 2 experiments. Thus any repeat occurrences of S-alleles in this case must be due to natural causes alone. The distribution of S-alleles between localities is shown in table 2 (second column). While most of the alleles appear to be confined to one or other of the three localities, 7 occur in two and 3 in all three localities. The problem here is to decide whether the observed differences between localities are no greater than can be expected between samples drawn at random from a single, panmictic population containing a finite number of S-alleles; or whether, alternatively, the differentiation between localities is sufficiently great to justify the conclusion that the samples have been drawn from different independent populations. Fisher (1961), who considered the same problem in connection with Emerson's (1939) Oe. organensis data, devised an ingenious test of the first hypothesis which involved the estimation of the number of alleles present in the population which did not occur in any of the locality samples. Once this has been done, the expected frequencies for each of the seven observed locality classes can be calculated and compared with the observed frequencies in the usual way by means of a χ^2 goodness of fit test. The expected frequencies for the present case are shown in the third column of table 2 and $P(\chi^2_{(3)} \ge$ 5.380 = 0.146 which is not significant. The observed differentiation between localities is, therefore, no greater than would be expected by chance alone.

Lastly, we note that if the estimate of the number of alleles present in the population, but which were absent in the locality samples, is added to the number found, we obtain an estimate of $\hat{N} = 46$ for the total number of different S-alleles in the West Midlands population of poppies.

4. DISCUSSION

We can now compare the present results with those from other species that have been similarly investigated and which have the same system of self-incompatibility. The species shown in table 3 fall into two broad groups, with the less thoroughly investigated *Trifolium repens* and *T. pratense* occupying the first and *Nicotiana sanderae*, *Oenothera organensis* and *T. hybridum* the second. Thus the number of alleles found in the species of the first group is generally higher (n = 24-41) and the repeatability correspondingly lower (R = 0.04-0.28) than the number (n = 13-34) and repeatabilities (R = 0.44-0.76) found in those of the second group. With n = 32 and R = 0.5, *P. rhoeas* appears to be intermediate to these two groups, unless the more conservative estimate of the repeat ratio is preferred (R = 0.31), in which case it lies closer to the clovers of the first group. Be this as it may, the chief point that emerges from this comparison is that the extent of the incompatibility polymorphism in *P. rhoeas* is broadly similar to that in other species.

The second point worth making about the present results concerns our inability to detect any real evidence of differentiation between the three localities in respect of the S-alleles they contain, which is, on the face of it, rather surprising in a colonising species whose populations must be greatly reduced in size from time to time. It is true that the probability we have

2	
Щ	
BL	
2	
F	

Distribution of S-alleles between localities $(15^* = S_{15} \text{ or } S_{17}; 17^* = S_{17} \text{ or } S_{15})$

				2		•		5			
	Locality	Observed	Expected								
Class	Combination	frequencies	frequencies				S-all	eles			
0	None	Ì	14.0817								
-	R102 only	9	6.7958	21	23	24	25	26	33		
5	R104 only	80	5.5337	27	28	29	30	32	34	35	36
ŝ	R106 only	œ	8.2316	1	6	9	×	10	11	14	18
4	R102 and R104	1	3.2346	19							
ŝ	R102 amd R106	ŝ	3.9724	ę	2	6	15*	16			
9	R104 and R106	1	2.6704	12							
7	R102. R104 and R106	e	1.5610	4	13	17^{*}					
		32	46.0817								

7117 711 7	n of the polymorphism in her speed	is at firmering printing mineri in	are me same means	I. I CH MARCA MANN	110043
			Number of	Number of	
			alleles found	alleles examined	Repeatability
Species	Reference	Nature of material	(u)	(m)	(R)
Nicotiana sanderae	East and Yarnell (1929)	Commercial varieties	15	34	0.61
Oenothera organensis	Emerson (1939)	Natural population	34	134	0.76
Trifolium repens	Atwood (1942)	Material of diverse	25	26	0.04
		origin			
		Breeders stocks	28	34	0.19
	Atwood (1944)	Natural population (1)	36	49	0.28
		Natural population (2)	39	49	0.22
Trifolium pratense	Williams and Williams (1947)	Natural population (1)	41	48	0.16
		Natural population (2)	35	38	0.09
Trifolium hybridum	Williams (1951)	Breeders stocks	13	21	0-44

TABLE 3

The extent of the polymorphism in five species of flowering plants which have the same incompatibility system as P. rhoeas

obtained is low (P = 0.146), certainly much lower than that obtained by Fisher (loc. cit.) in his analysis of a portion of Emerson's data $(\chi^2_{(3)} = 0.935, P = 0.90-0.80)$. There is also some evidence that the locality most distantly removed from the others (R104) contains a higher than expected number of alleles that occur only in this sample (class 2 of table 2) and a lower than expected number of alleles that also occur in the other localities (classes 4 and 6). Nevertheless, on the present evidence, the null hypothesis of no differentiation between localities must stand, a conclusion which raises the question of whether the size of the experiment is too small to detect anything but the most gross differentiation between localities; or, whether, alternatively, the biology of the species is, despite first impressions to the contrary, such as to impede divergence between populations in respect of the S-alleles they contain.

On the first question, although overall the scale of our investigations is second only to Emerson's, the number of alleles found and the number examined in each locality is, nevertheless, quite small (15 out of 19 in R102; 13 out of 18 in R104; and 17 out of 24 in R106). Though the use of full-sib families in investigations of this type is analytically convenient, in terms of the number of alleles examined they are less efficient for an equivalent effort than the analysis of a random sample of plants from each location, in which any differences in allele frequency must be due to natural causes alone.

There are three aspects of the biology of the species which may act to impede divergence between localities in respect of the S-alleles they contain (see McNaughton and Harper (1964), for a general account of the ecology of the species). Firstly, seed from natural populations possesses a pronounced innate dormancy, which enables the seed to persist for many years in the soil if conditions for germination are unfavourable (this dormancy has to be broken by artificial means if a satisfactory germination percentage is to be obtained in laboratory experiments). Thus although the species is a strict annual, the dormancy of its seed must cause generations to overlap, thus buffering any response by the population to either secular or random changes in the environment. Secondly, because stands of poppies are found growing only on recently disturbed and uncolonised soil, the species being apparently unable to compete with more persistent forms, the interval between successive generations may be considerable. For this reason, too, any change that is taking place in the population will, on a per annum basis, be slow. Lastly, though historically poppies are well-known as a weed of arable crops, they are now so easily controlled by herbicides that they are rarely found in any numbers except in roadside and waste-places. However, because of the longevity of their seed and the slow turnover and overlapping nature of their generations, the present genetical structure of their populations may reflect very much more their historical position as arable weeds rather than their contemporary position in the flora. For these reasons, there may have been insufficient time for any detectable divergence between populations to have taken place in respect of the S-alleles they contain irrespective of whether this is due to selection or to drift. If these arguments are correct, we should not expect to detect differences between localities even in very large experiments, unless the differences between the latter are very much greater than in present circumstances. The value of this prediction is, of course, that it is testable and experiments are currently under way with this purpose in mind.

Acknowledgments.—We are indebted to Mrs D. Edgerton for technical assistance, to Dr Paul Davies of the Department of Mathematical Statistics for help with Fisher's χ^2 test of differentiation and to Dr J. S. Gale for much useful discussion. One of us (J.M.C.) was in receipt of an S.R.C. studentship whilst the work described in this paper was being carried out.

5. References

- ATWOOD, S. S. 1942. Oppositional alleles causing cross-incompatibility in *Trifolium repens*. *Genetics*, 27, 333-338.
- ATWOOD, S. S. 1944. Oppositional alleles in natural populations of *Trifolium repens. Genetics*, 29, 428-435.
- CAMPBELL, J. M. 1977. Population genetics of self-incompatibility in *Papavar rhoeas L*. Ph.D thesis, University of Birmingham, England.
- COMSTOCK, R. E., AND ROBINSON, H. F. 1952. Estimation of average dominance of genes. In *Heterosis*, ed. J. M. Gowen. Iowa State College Press, Ames, Iowa.
- CROSBY, J. L. 1966. Self-incompatibility alleles in the population of *Oenothera organensis*. Evolution, 20, 567-569.
- EAST, E. M., AND MANGELSDORF, A. J. 1925. A new interpretation of the hereditary behaviour of self-sterile plants. *Proc. Nat. Acad. Sci.*, 11, 166-171.
- EAST, E. H., AND YARNELL, S. M. 1929. Studies on self-sterility. VIII. Self-sterility allelomorphs. *Genetics*, 14, 455-487.
- EMERSON, S. 1938. The genetics of self-incompatibility in Oenothera organensis. Genetics, 23, 190-202.
- EMERSON, S. 1939. A preliminary survey of the Oenothera organensis population. Genetics, 24, 524-537.
- EMERSON, S. 1940. Growth of incompatible pollen tubes in Oenothera organensis. Bot. Gat., 101, 890-911.
- EMERSON, S. 1941. Linkage relationships of two gametophytic characters in Oenothera organensis. Genetics, 26, 469-473.
- EWENS, W. J., AND EWENS, P. M. 1966. The maintenance of alleles by mutation—Monte Carlo results for normal and self-sterility populations. *Heredity*, 21, 371-378.
- FISHER, R. A. 1958. The Genetical Theory of Natural Selection, 2nd edition. Dover Publications, New York.
- FISHER, R. A. 1961. Possible differentiation in the wild population of Oenothera organensis. Aust. J. Biol. Science, 14, 76-78.
- KIMURA, M., AND MARUYAMA, T. 1966. In Mayo, O. 1966.
- LAWRENCE, M. J. 1975. The genetics of self-incompatibility in Papaver rhoeas. Proc. R. Soc. Lond. B., 188, 275-285.
- LAWRENCE, M. J., AFZAL, M., AND KENRICK, J. 1978. The genetical control of self-incompatibility in Papaver rhoeas. Heredity, 40, 239-253.
- LEVIN, D. A., RITTER, K. AND ELLSTRAND, N. C. 1979. Protein polymorphism in the narrow endemic Oenothera organensis. Evolution, 33, 534-542.
- LEWIS, D. 1948. Structure of the incompatibility gene. I. Spontaneous mutation rate. Heredity, 2, 219-236.
- LEWIS, D. 1951. Structure of the incompatibility gene. III. Types of spontaneous and induced mutation. *Heredity*, 5, 399-414.
- MARTIN, F. W. 1959. Staining and observing pollen tubes in the style by means of fluorescence. Stain Technology, 34, 125-128.
- MAYO, O. 1966. On the problem of self-incompatibility alleles. Biometrics, 22, 111-120.
- McNAUGHTON, I. M., AND HARPER, J. L. 1964. Biological Flora of the British Isles. No. 99. Papaver L. J. Ecol., 52, 767-793.
- MORAN, P. A. P. 1962. The Statistical Processes of Evolutionary Theory. Clarendon Press, Oxford.
- OOI, S. C. 1970. Variation in wild populations of *Papavar rhoeas L*. Ph.D thesis, University of Birmingham, England.
- WILLIAMS, R. D., AND WILLIAMS, W. 1947. Genetics of red clover (*Trifolium pratense L*) compatibility. III. The frequency of incompatibility S alleles in two non-pedigree populations of red clover. J. Genet., 48, 67-79.
- WILLIAMS, W. 1951. Genetics of incompatibility in alsike clover, Trifolium hybridum. Heredity, 5, 51-73.

- WRIGHT, S. 1939. The distribution of self-sterility alleles in populations. Genetics, 24, 538-552.
- WRIGHT, S. 1960. On the number of self-incompatibility alleles maintained in equilibrium by a given mutation rate in a population of given size: a re-examination. *Biometrics*, 16, 61-85.
- WRIGHT, S. 1964. The distribution of self-incompatibility alleles in populations. *Evolution*, 18, 609-619.