

The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. V. Cross-classification of the S-alleles of samples from three natural populations

M. J. LAWRENCE,* M. D. LANE, S. O'DONNELL & V. E. FRANKLIN-TONG

Wolfson Laboratory for Plant Molecular Biology, School of Biological Sciences, University of Birmingham, Birmingham B15 2TT, U.K.

The results obtained from a comprehensive cross-classification of the self-incompatibility alleles from three natural populations of *Papaver rhoeas* (R102, R104 and R106) indicate that no less than 23 of a total of 45 alleles occur in at least two samples and that 15 of these occur in all three. This suggests that the overlap between the complements of alleles that these populations contain is substantial. These results also show, however, that, in general, different alleles occur at a relatively high frequency in different populations which, it is argued, rules out the possibility that the unequal allele frequencies in these populations is caused by an extra effect of selection acting on the locus via a pleiotropic effect. The analytical procedures used to identify incompatibility genotypes and S-alleles in full-sib families that are raised from the seed obtained by crossing plants of known incompatibility genotype in this strictly annual species are also described.

Keywords: cross-classification of alleles, *Papaver rhoeas*, population genetics, self-incompatibility.

Introduction

A preliminary survey of the distribution of self-incompatibility (S-) alleles between samples taken from three natural populations of *Papaver rhoeas* revealed that most of the 32 alleles identified were confined to one or other of these samples and that only 10 occurred in more than one (Campbell & Lawrence, 1981a). Analysis of these data, however, showed that the differences between these samples were not significantly greater than those between samples drawn at random from a single population containing a finite number of alleles. Taken at face value, therefore, these results suggest that these natural populations contain essentially the same complement of alleles, despite the fact that the distance between the closest pair is no less than 43 km and that these populations are known to differ significantly for a number of metrical characters (Ooi, 1970). None of these samples was very large, however, so it is possible that the power of the test employed (Fisher, 1961) was insufficient to detect anything but the most extreme

differentiation between populations in respect of the alleles they contain.

Since this preliminary survey, we have resampled the same three natural populations to determine the number and frequency of the S-alleles they contain (Campbell & Lawrence, 1981b; Lawrence & O'Donnell, 1981), as a result of which larger samples of alleles from each of these populations became available. In the present paper, we report the results obtained from a comprehensive cross-classification of the alleles of each of these samples against those of the others to re-examine this question of the apparent lack of differentiation between populations.

Material

Since *P. rhoeas* is a strictly annual species, it is not possible to retain plants whose incompatibility phenotype has been identified from one year to the next. The only way to preserve the S-alleles that have been identified in one experiment for future investigation is via seed. The advantage of this method is that because the seed of this species remains viable for many years, and

*Correspondence.

each capsule can contain as many as 2000 seeds, one successful cross is capable of providing material for many further experiments. The disadvantage of this method is, of course, that the incompatibility genotypes that occur in the progeny of a cross have to be re-identified in every experiment in which the family is raised.

The parents of the families used in the present investigation were the plants whose incompatibility genotype was determined in the experiments described by Campbell and Lawrence (1981b) and Lawrence and O'Donnell (1981). It is necessary to make crosses for seed early in the flowering period of plants raised in the open. These crosses were made, therefore, before the incompatibility genotype of the parents had been determined, so that they were effectively made at random. Within each of the three samples, one from each of the populations being assayed, plants were paired off at random and crossed for seed reciprocally. Twenty families were produced in this way from the plants of the R102 (Wellesbourne, Warwickshire) sample, 19 from the R104 (Broad Oak, Herefordshire) sample and 29 families from the R106 (Hackmans Gate, Worcestershire) sample. In addition, each of the plants crossed for seed was also self-pollinated to test its self-incompatibility. With the exception of plant number 54 of the R106 sample (Campbell & Lawrence, 1981b), all plants turned out to be highly self-incompatible. Thus, two-thirds of them produced no seed at all when self-pollinated whereas the remaining third produced only a very small quantity of poor quality seed which, when tested, nearly always failed to germinate.

Methods

Provided that the parents of a cross have been correctly classified, the composition of their progeny can be predicted. It is necessary, however, to carry out a series of systematic crosses between the individuals of a family and then between these and those of other families before it is possible to assign an incompatibility genotype to each of them. This classification can be accomplished in two stages: an intra-family and an inter-family stage.

Intra-family classification

The usual way of classifying the members of a family is to cross each to every other in diallel fashion. This is an inefficient procedure, however, because with only two or four genotypic classes in a family, a full diallel matrix contains much redundant information. An alternative, more efficient procedure that requires only one-third

of the number of crosses as the diallel procedure is as follows. A first round of crosses is made between, say, five members of each family which links them in an endless chain (e.g. 1×2 , 2×3 , 3×4 , 4×5 , 5×1), the purpose of which is to find a pair of half-compatible plants. Every other member of the family is then pollinated by each of these half-compatible testers. In the case of two class families (those whose parents were half-compatible), every plant in the family is expected to be incompatible with one of these testers and half-compatible with the other; with four class families (those whose parents were fully compatible), every plant in the family is expected to be half-compatible with one tester and either incompatible or fully compatible with the other. The purpose of this second round of pollinations is to assign each member of a family to one or other of two groups. A third round of pollinations is then made to link plants of the same class directly, to verify the provisional classification of the second round. The outcome of this first, intra-family stage of classification is that each plant in a family is assigned to one or other of either two or four incompatibility classes, depending on the type of family in question.

Inter-family classification

The objective of this second stage of analysis is to assign genotypes to plants of the classes identified in the first stage. For two class families, this is most easily accomplished by crossing one plant from each of the two classes of a family to one plant from each class of the family raised from the seed of the reciprocal cross, because the same non-parental incompatibility genotype is expected to occur in each family. A 2×2 crossing scheme of this type gives a complete classification of any two class family (Lawrence *et al.*, 1978).

Four class families, on the other hand, cannot be so easily classified because their composition is the same irrespective of the direction of the cross which produced them. Classification in this case can be achieved by crossing one plant from each class from one family with one plant from each class of a second family that contains one or more alleles that also occur in the first. If the pair of families have three alleles in common, this 4×4 crossing scheme gives a complete classification of both families (Table 1a). With fewer alleles in common, however, this procedure yields only a partial classification although, provided each family is complete (each containing at least one plant in each of their four classes), it will generally allow the identification of the genotypes of a fully compatible pair of plants in each. For a complete classification, it is necessary to use plants from a third family which contains alleles that

Table 1 Two examples of inter-family analysis of four class families from the R106 population. (a) A pair of four class families which have three alleles and one genotype in common. Class A= Class A', as when crossed, plants of these classes are incompatible (-), so that their genotype must be S₃S₇. Hence the plants of class D' in family 21, being fully compatible (+ +) with those of class A', must be S₁₀S₁₇; and those of class D in Family 28 must be, for the same reason, S₁₀S₂₂. One array of family 21 contains nothing but half-compatible (+) entries so that individuals of class B' must be S₃S₁₀ and those of class C', S₇S₁₇; by a similar argument, the genotype of plants of class C of family 28 is S₇S₁₀ and of class B, S₃S₂₂. Note that prior to their complete classification, the classes are indicated by letters such that plants of class A are fully compatible with those of D and those of B are fully compatible with those of C, all other crosses between classes being half-compatible. Although these data are presented retrospectively (after classification) any alternative alphabetic labelling of these classes, within the constraints defined, would lead to the same final classification. (b) A pair of four class families that have two alleles and one genotype in common. A ≡ C', so that plants of these classes are S₁₄S₁₇; hence those of D must be S₂₁S₂₄ and of B', S₃S₂₇. Without further information, however, B is interchangeable with C, and A' is interchangeable with D'. Two tester genotypes from the two class family 29, however, can be used to complete the classification of this pair of four class families by revealing that, of the two unresolved classes in family 12, individuals of A' contain S₃, so that their genotype must be S₃S₁₇ and those of D', S₁₄S₂₁; and, in the case of family 15, plants of class B contain S₁₄ and are, therefore S₁₄S₂₁, so that those of class C must be S₁₇S₂₄.

(a) (i) Expected composition of families

Family	Cross	Progeny			
21	S ₃ S ₁₇ × S ₇ S ₁₀	S ₃ S ₇	S ₃ S ₁₀	S ₇ S ₁₇	S ₁₀ S ₁₇
28	S ₃ S ₁₀ × S ₇ S ₂₂	S ₃ S ₇	S ₃ S ₂₂	S ₇ S ₁₀	S ₁₀ S ₂₂

(ii) Expected outcome of inter-family crosses

Family	21					
	Class	A'	B'	C'	D'	
	S _i S _j	S ₃ S ₇	S ₃ S ₁₀	S ₇ S ₁₇	S ₁₀ S ₁₇	
28	A	S ₃ S ₇	-	+	+	+
	B	S ₃ S ₂₂	+	+	+	+
28	C	S ₇ S ₁₀	+	+	+	+
	D	S ₁₀ S ₂₂	+	+	+	+

also occur in the first two families (Table 1b). In planning experiments which require the complete classification of four class families, considerable care has to be taken to ensure an appropriate set of families is raised if this objective is to be achieved.

(b) (i) Expected composition of families

Family	Cross	Progeny			
12	S ₃ S ₁₄ × S ₁₇ S ₂₇	S ₃ S ₁₇	S ₃ S ₂₇	S ₁₄ S ₁₇	S ₁₄ S ₂₇
15	S ₁₇ S ₂₁ × S ₁₄ S ₂₄	S ₁₄ S ₁₇	S ₁₄ S ₂₁	S ₁₇ S ₂₄	S ₂₁ S ₂₄
29A	S ₁₁ S ₁₄ × S ₃ S ₁₁	S ₃ S ₁₁	S ₃ S ₁₄	-	-
29B	S ₃ S ₁₁ × S ₁₁ S ₁₄	-	S ₃ S ₁₄	S ₁₁ S ₁₄	

(ii) Expected outcome of inter-family crosses

Family	Class	12				29A
		A'	B'	C'	D'	-
	S _i S _j	S ₃ S ₁₇	S ₃ S ₂₇	S ₁₄ S ₁₇	S ₁₄ S ₂₇	S ₃ S ₁₄
15	A	S ₁₄ S ₁₇	+	+	-	+
	B	S ₁₄ S ₂₁	+	+	+	+
	C	S ₁₇ S ₂₄	+	+	+	+
	D	S ₂₁ S ₂₄	+	+	+	+
29A	-	S ₃ S ₁₁	+	+	+	+

Inter-population classification

Because the number of pollinations required to completely cross-classify the alleles of each population against those of the others was very large, this task had to be divided up over a number of separate experiments. In each experiment, a number of families were raised which, between them, contained the desired subset of alleles from each of two of the populations in question. Using the procedures described earlier, two testers were identified in each family which between them contained all of the alleles segregating in that family, a requirement which is met if those from four class families are fully compatible and those from two class families are half-compatible. Initially, the incompatibility genotypes of these testers were determined before they were used for inter-population crosses. Later, however, this was not done because most intra-population crosses are expected to be fully compatible, even if the populations from which they are derived contain the same complement of S-alleles. In these circumstances, it is necessary to know only that between them a pair of testers contain all of the alleles of their family. Hence, these testers could be identified by intra-family analysis alone. As soon as a cross between a tester from one population turned out to be other than fully compatible with one from another population, indicating that they possessed an allele in common, it became necessary to carry out inter-family

analysis to identify this common allele. Thus, the most efficient procedure is to carry out intra-family analysis to identify a pair of appropriate testers from each family in the experiment, to cross these testers systematically with those from the other population and to carry out inter-family analysis only for those testers that turn out to carry alleles that occur in both of the populations in question. If n_1 families are raised from one population sample and n_2 families from a second, the inter-population stage of an experiment involves $2n_1 \times 2n_2$ pollinations in all.

In the early years of this investigation, plants were raised in the open which limited the number of families that could be handled within the natural flowering period of the species (5–6 weeks) and hence the number of alleles from each population that could be cross-classified against those of another. Later, however, it became possible to extend this flowering period to nearly 5 months by raising plants in the glasshouse during winter, in which case it became possible to raise two sets of families, one from each of the populations under investigation in the experiment, which between them contained all or very nearly all of the S-alleles identified in previous experiments. Hence, although these later experiments were very large, they were also very efficient. The re-identification of the genotypes of the plants raised in these later experiments also became more efficient because of the use of pollen stored from plants of known genotype from one experiment to the next (Franklin-Tong *et al.*, 1988).

All other technical details were as previously described (Lawrence, 1975; Lawrence *et al.*, 1978), except that in later experiments stigmas harvested after pollination for classification were placed directly into aniline blue stain without fixation or softening.

Results

Verification of parental genotype

The inter-family classification procedure described in the previous section will, of course, succeed in identifying the incompatibility genotypes of family members if and only if the parents of these families had been correctly classified in previous experiments (Campbell & Lawrence, 1981b; Lawrence & O'Donnell, 1981). Inter-family analysis thus provides a valuable opportunity to check the accuracy of the original classification of the parents of the families used in the present investigation, and where errors are discovered, to correct them. During the course of this investigation, 18 of the 20 R102 families were used, including all 16 of the completely classified (both parents classified) and two of the partially classified (only one parent

classified) families; 18 of the 19 R104 families, including all 17 completely and one partially classified family; and 22 of the 29 R106 families available, all of which were completely classified. Between them, these families involved 33 of the 36 R102 parents; all 36 of the R104 parents and 42 of the 51 R106 parents. All of the R106 parents examined in this way turned out to be correctly classified. Five of the R102 parents and six of those from the R104 population, however, were misclassified in respect of one of the pair of S-alleles they contained. In the majority of cases, these errors were caused by a failure to recognize, in the previous experiments, that four alleles (three in the R102 and one in the R104 sample) were the same as others of their sample, that is, were duplicates of the latter. Three of these errors, however, were due to straightforward misclassification. Two further alleles of the R102 sample and one in the R104 sample also turned out to be misclassified in the previous experiments; we were, however, unable to re-classify these alleles before this investigation was terminated, so that they have been excluded from the data presented here. Other sources of error, such as that due to pollen or seed contamination, or to self-pollination, have not been detected in this investigation. The revised genotypes of the 11 misclassified plants are shown in Table 2.

These revisions of parental genotype make it necessary to recalculate the values of the statistics and esti-

Table 2 Plants whose genotype has been revised from those shown in Table 3 (R102) and Table 1 (R104) of Lawrence and O'Donnell (1981). The revised classification of the R102 plants involved four independent revisions and that of the R104, three (revisions shown in bold). S_k in the genotype of the R104 plant no. 32 represents an allele that remained unclassified at the end of the investigation; it is possible that it is identical to one of the other alleles in this population. Further details about the basis of these revisions are given in Lane (1990)

Population	Plant no.	Previous S_iS_j	Revised S_iS_j
R102	11	$S_{13}S_{14}$	S_4S_{14}
	19	$S_{13}S_{15}$	S_4S_{15}
	26	$S_{17}S_{18}$	$S_{10}S_{17}$
	52	$S_{27}S_{28}$	$S_{12}S_{27}$
	58	$S_{10}S_{23}$	S_5S_{23}
R104	26	S_3S_{13}	$S_{13}S_{26}$
	32	$S_{11}S_{19}$	$S_{19}S_k$
	38	S_5S_{22}	S_5S_{14}
	43	S_3S_{23}	$S_{23}S_{26}$
	50	$S_{10}S_{22}$	$S_{10}S_{14}$
	68	$S_{22}S_{27}$	$S_{14}S_{27}$

Table 3 Revised values of the number of alleles found (n) and the number of alleles examined (m) in the three samples investigated and the repeatability (R), dissimilarity (D), estimated number of alleles in the population (N) and the chi-squared test of the hypothesis of equal allele frequency. Previous values and estimates are shown in parentheses and revised values in bold. The entries in rows 1–5 for R102 and R104 replace those given in Table 5 of Lawrence and O'Donnell (1981) and those in the last row replace the corresponding entries shown in Table 3 of O'Donnell and Lawrence (1984). The estimates of N for the poppy populations are those obtained from the E2 estimator of O'Donnell and Lawrence (1981); that for the *Oenothera organensis* population (Emerson, 1939) is the maximum likelihood estimate. Although unchanged, the R106 and *Oe. organensis* values are shown for comparison

Statistic	R102	R104	R106	Oenothera
No. of alleles found (n)	27 (30)	25 (26)	31	28
No. of alleles examined (m)	72	71 (72)	102	74
Repeatability (R)	0.65 (0.61)	0.68 (0.67)	0.72	0.65
Dissimilarity (D)	1.690 (2.069)	1.689 (2.446)	3.656	–1.240
χ^2	39.000* (46.262*)	36.507* (44.618*)	64.260***	18.243
\hat{N}	35 (42)	32 (34)	38	30

mates derived from the R102 and R104 data; the revised values are shown in Table 3. Although the new values are generally smaller than those they replace, they are not by amounts which make it necessary to revise the chief conclusions that can be drawn from them, namely, that the allele frequencies in these populations appear to be significantly unequal and that these populations contain, perhaps, 40 different S-alleles, although probably not very many more than this number.

Cross-classification results

The most convenient way of presenting the results of a cross-classification experiment is in the form of a matrix whose dimensions are determined by the number of alleles in each of the two samples which have been cross classified. If the number of alleles in the first sample is n_1 and that in the second n_2 , then the matrix has n_1 rows and n_2 columns and the product of these numbers, $n_1 n_2$, is the number of pair-wise comparisons that have to be made between these alleles for a complete cross-classification. We are concerned, in the present investigation, with samples from three populations, R102, R104 and R106, so that the results obtained by cross-classifying the alleles of each sample against those of the others can be presented in the form

of three cross-classification matrices, R102 \times R104, R102 \times R106 and R104 \times R106. We have seen earlier (Table 3) that the number of alleles found in these samples were 27, 25 and 31, respectively. The dimensions of these cross-classification matrices are, therefore, 27 \times 25, 27 \times 31 and 25 \times 31 and the number of pair-wise comparisons involved in each is 675, 837 and 775, respectively, giving a total of 2287 comparison in all. A complete survey (100 per cent) was achieved with the R102 \times R104 cross-classification. With the R102 \times R106 and the R104 \times R106 cross-classifications, on the other hand, only 59.9 per cent and 76.1 per cent of these comparisons were made directly. Provided that the alleles of each pair of samples have been correctly classified, so that each is known to be different from every other within its sample, any one row or column of a cross-classification matrix can contain no more than one identity, indicating the presence of an S-allele that occurs in both samples. It follows, therefore, that all other pair-wise comparisons in a row or column containing an identity must involve alleles that are not identical. Hence, in assessing the thoroughness of cross-classification, we may add the number of comparisons inferred in this way to the number of those made directly. If this is done, the proportion of comparisons made one way or the other rises to 84.9 per cent for the R102 \times R106

S	3	17	10	11	7	14	27	1	16	28	15	18	19	20	21	24	4	5	8	12	22	25	26	2	6	9	13	23	29	30	31
7	✓	✓	*	✓	✓	✓	✓	✓	○	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	○	✓	○	✓	○	✓	✓	✓	○
3	✓	✓	✓	✓	✓	✓	✓	*	○	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	○	✓	○	✓	○	✓	✓	○	
10	✓	✓	✓	✓	○	✓	✓	✓	○	✓	✓	✓	✓	✓	✓	*	✓	✓	✓	✓	✓	✓	✓	○	✓	○	✓	○	✓	○	
8	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
9	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
14	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
20	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
26	✓	✓	✓	✓	✓	✓	✓	○	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	○	✓	*	○	✓	○	✓	○	○	
4	✓	✓	✓	✓	○	✓	○	○	✓	✓	✓	✓	✓	✓	✓	✓	*	✓	✓	✓	✓	○	○	○	○	○	○	○	○	○	
11	✓	✓	✓	✓	○	*	○	○	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	○	○	○	○	○	○	○	○	○	
19	✓	✓	✓	✓	○	○	○	*	○	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	○	○	○	○	○	○	○	○	○	
5	*	✓	✓	✓	✓	✓	✓	○	○	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	○	○	○	○	○	○	○	○	○	
6	✓	✓	✓	✓	✓	✓	✓	○	○	✓	✓	✓	✓	*	✓	✓	✓	✓	✓	✓	✓	○	○	○	○	○	○	○	○	○	
13	✓	✓	✓	✓	○	✓	○	○	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	○	○	○	○	○	○	○	○	○	
17	✓	✓	✓	✓	*	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	
18	✓	✓	✓	✓	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	
25	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
1	✓	✓	✓	✓	*	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	
2	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
12	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
21	✓	✓	✓	✓	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	
23	✓	✓	○	✓	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	
24	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	○	○	○	○	○	○	○	○	○	
27	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	*	○	○	○	○	○	○	○	○	○	○	○	○	

Fig. 3 Results obtained from the cross-classification of the 25 alleles of the R104 sample (left) against the 31 alleles of the R106 sample (top). Although, again unlikely, as many as eight further identities might have been detected had those comparisons neither made nor inferred (blank cells) been made. Other details as for Figs 1 and 2.

Table 4 A list of the 53 identities revealed by cross-classifying the S-alleles of each of the three samples against those of the others and the number of pollinations on which these identities are based. Identities not revealed directly but by the best fit procedure (see text) are indicated by a single asterisk, those found indirectly via their identity with an S-allele of the third sample are shown with a double asterisk and identities that have been detected but could not be fully resolved are indicated by a triple asterisk. The number of times that an allele occurred in its sample is shown in superscript. For brevity, S₁ is indicated by 1; S₂ by 2 and so on for the other alleles

R102 × R104	No. of pollinations	R102 × R106	No. of pollinations	R104 × R106	No. of pollinations
15 ⁸ ≡ 9 ⁴	35	16 ⁷ ≡ 17 ¹¹	48	3 ⁶ ≡ 1 ⁴	20
1 ⁷ ≡ 10 ⁶	26	1 ⁷ ≡ 24 ³	30	7 ⁹ ≡ 10 ⁷	19
12 ³ ≡ 4 ³	10	10 ³ ≡ 3 ¹¹	11	5 ² ≡ 3 ¹¹	14
5 ³ ≡ 7 ⁹	9	4 ³ ≡ 7 ⁵	9	4 ³ ≡ 4 ²	12
10 ³ ≡ 5 ²	9	14 ³ ≡ 14 ⁵	9	10 ⁶ ≡ 24 ³	10
20 ² ≡ 20 ⁴	6	5 ³ ≡ 10 ⁷	6	1 ¹ ≡ 7 ⁵	4
8 ² ≡ 23 ¹	4	12 ³ ≡ 4 ²	4	11 ³ ≡ 27 ⁵	3
4 ³ ≡ 1 ¹	3	21 ¹ ≡ 28 ⁴	2	23 ¹ ≡ 30 ¹	3
27 ³ ≡ 18 ²	3	24 ¹ ≡ 18 ³	1	6 ² ≡ 20 ³	2
2 ³ ≡ 11 ³	2	2 ³ ≡ 27 ⁵	1	19 ³ ≡ 28 ⁴	2
11 ¹ ≡ 26 ⁴	2	8 ² ≡ 30 ¹	1	26 ⁴ ≡ 2 ¹	1
14 ³ ≡ 17 ²	2	17 ² ≡ 20 ³	1	24 ¹ ≡ 12 ²	1***
17 ² ≡ 6 ²	2	11 ¹ ≡ 2 ¹	0*	or 13 ¹ or 15 ³	
21 ¹ ≡ 19 ³	1	7 ⁵ ≡ 5 ²	1*	12 ¹ ≡ 12 ²	1***
25 ¹ ≡ 21 ¹	1	3 ³ ≡ 1 ⁴	0*	or 13 ¹ or 15 ³	
30 ¹ ≡ 24 ¹	1	26 ¹ ≡ 16 ⁴	0***	17 ² ≡ 14 ⁵	0**
		or 26 ² or i		14 ⁴ ≡ 16 ⁴	4***
				or 26 ² or j	
				16 ¹ ≡ 6 ¹	2***
				or 16 ⁴ or 26 ²	
				or k	
3 ³ ≡ 3 ⁶	6*	30 ¹ ≡ 12 ²	0***	27 ¹ ≡ 5 ²	1*
		or 4 ³ or 15 ³			
7 ⁵ ≡ 27 ¹	3*				
26 ¹ ≡ 14 ⁴	3*				

Table 5 Fifteen alleles that occurred in all three samples (* = or 26² or j and ** = 13¹ or 15³; see Table 4)

R102		R104		R106
1 ⁷	≡	10 ⁶	≡	24 ³
7 ⁵	≡	27 ¹	≡	5 ²
2 ³	≡	11 ³	≡	27 ⁵
3 ³	≡	3 ⁶	≡	1 ⁴
5 ³	≡	7 ⁹	≡	10 ⁷
10 ³	≡	5 ²	≡	3 ¹¹
12 ³	≡	4 ³	≡	4 ²
4 ³	≡	1 ¹	≡	7 ⁵
14 ³	≡	17 ²	≡	14 ⁵
8 ²	≡	23 ¹	≡	30 ¹
17 ²	≡	6 ²	≡	20 ³
11 ¹	≡	26 ⁴	≡	2 ¹
21 ¹	≡	19 ³	≡	28 ⁴
26 ¹	≡	14 ⁴	≡	16 ^{4*}
30 ¹	≡	24 ¹	≡	12 ^{2**}

complement of alleles of each sample and those of the others. The status of the identities shown in these matrices, however, varies considerably. A few have been confirmed many times, both in the same experiment and independently in different experiments because they involve alleles which occurred at a relatively high frequency in the population in which they were found and, hence, are well represented in the families used in this investigation. Most identities, however, necessarily involve alleles that occurred at lower frequencies in their population and are based on fewer pollinations because these alleles occur in only one family and cannot be confirmed independently. Two of the identities shown in Table 4 were found indirectly because in each case the alleles in question were involved in identities elsewhere (e.g. S₁₁ of R102 is the same as S₂ of R106 because each is the same as S₂₆ of the R106 sample). Six of the identities shown in Table 4 were established by a best fit procedure, in which the most probable pairing was determined from a group of pollinations known to contain an identity but which could not otherwise be unambiguously located (Lane, 1990). Also, six further identities could not be fully resolved before the investigation was terminated and, hence, are not shown in Figs 2 and 3. Nevertheless, while the confidence attached to some of these identities must be less than that attached to others, taken overall, these results leave little doubt that many alleles occur in more than one sample.

The second point worth making about these results is that, in general, alleles occur at different frequencies in different samples and, hence, presumably in different populations (Tables 4 and 5). For example S₁₅,

which occurred eight times in the R102 sample, occurred only four times in the R104 sample, where it was labelled S₉, and not at all in the R106 sample. Again, although S₁₆, an allele that occurred seven times in the R102 sample, turned out to be the same as S₁₇, one of the two most frequent alleles in the R106 sample, it has not been found in the R104 sample. This evidence is particularly instructive because, for reasons given earlier, it is considerably easier to cross-classify the alleles of one sample against those of another when one of them occurs at a high frequency in its sample. It follows, therefore, that the absence of an allele from one sample that occurs at a relatively high frequency in another is unlikely to be due to incomplete cross-classification.

Discussion

Hitherto the only information available about the distribution of alleles between natural populations of species of flowering plant in which self-incompatibility is determined by a single, multi-allelic gene whose expression in the pollen is gametophytic, is that provided by our preliminary survey (Campbell & Lawrence, 1981a). Although only 10 of a total of 32 alleles found in this earlier survey occurred in more than one locality, differences between these localities turned out to be, on Fisher's (1961) test, no greater than those expected between samples drawn at random from a single population containing a finite number of alleles. Hence, while there must be some doubt about the power of a test applied to a relatively small body of data, the results of this preliminary survey provided, somewhat surprisingly, no evidence of any differentiation between these populations in respect of the alleles they contain.

The samples of the present investigation are at least three times the size of those of the previous ones so that any test of the null hypothesis of no differentiation carried out on the present data should be more powerful than one applied to the data of the previous experiment. The alleles examined in the latter, however, were completely cross-classified across samples whereas those of the present investigation are not. Hence Fisher's test, which requires complete cross-classification, cannot be applied easily to the present data. It is, nevertheless, instructive to summarize these data in the way required by this test (Table 6). Of a total of 45 alleles identified in this investigation, no less than 23 occur in more than one sample and the majority of these, 15, occur in all three. Furthermore, had cross-classification been complete some of the 12 alleles, which apparently occur only in the R106 sample, would almost certainly be redistributed to classes 5, 6

Table 6 Distribution of the 45 alleles between localities

Locality combination	Observed frequency
R102 only	6
R104 only	4
R106 only	12
R102 and R104	4
R102 and R106	2
R104 and R106	2
R102, R104 and R106	15

and 7 because the cross-classification of the R106 alleles against those of the R102 and the R104 samples was less complete than that of the R102 alleles against those of the latter. A redistribution of some of the alleles of class 3 would also, of course, reduce the total number of alleles found in the experiment. While the information given in Table 6 is probably incomplete, there is little doubt that the overlap between the three sets of alleles is extensive. This being the case, our chief interest in these data is to use them to estimate the overlap between the complements of the alleles of the populations from which these samples were taken rather than asking whether these populations are differentiated in this respect; this is a matter which is taken up in the following paper.

The second point worth making about these results arises from the observation that, in general, the alleles which occur at a relatively high frequency in one sample are not the same as those which occur at a relatively high frequency in other samples. Assuming that this is also true of the natural populations from which these samples were drawn, we may ask whether this observation tells us anything about the cause of the significantly unequal allele frequencies that previous experiments have revealed in these samples (Campbell & Lawrence, 1981b; Lawrence & O'Donnell, 1981).

In the first of these papers, we advanced two alternative hypotheses to account for these unequal allele frequencies. The first of these supposed that some of the alleles are subject to natural selection over and above that associated with the incompatibility effect due either to a pleiotropic effect of the S-gene or to close linkage with another gene that affects fitness. Pollen carrying certain S-alleles, for example, might grow more quickly through compatible stigmas than pollen carrying other alleles. Clearly, if this were to be the case the frequencies of the former in the population would be greater than those of the latter and neither would be equal to the frequencies expected on the assumption that selection is limited to the incompatibility effect.

The second hypothesis supposed that the unequal allele frequencies were due to sampling effects caused,

for example, by some disturbance of the habitat which does not at the same time reduce the size of the population. In these circumstances, any allele could occur at a relatively high frequency in the population and would continue to do so until the population had reattained equilibrium. In particular, if each of two populations were subject to this kind of disturbance, we would not expect the same alleles to occur at a relatively high frequency in each.

The observation that different alleles occur at different frequencies in a sample and, by inference, in the population from which these samples were drawn, is of course consistent with this second hypothesis. These data are also consistent with one version of the first hypothesis. Thus, if the extra effect of selection on the S-gene is due to its close linkage to a second gene, which is the chief target of this extra effect of selection, there is no obvious reason why we should expect the same S-allele to be associated with this extra effect of selection in different populations. Indeed, any association of this kind is unstable because, in time, either linkage equilibrium will be established between the alleles of the two genes if the effect of this extra selection is to maintain polymorphism at the second locus or the advantageous allele of this second locus will become fixed in the population. In the first case, once linkage equilibrium has been attained, every S-allele is as likely to be associated with the alleles at the second locus as every other and in the second case, once fixation of the advantageous allele has occurred, each and every S-allele will have become associated with it. The chief point here is that this version of the first hypothesis also supposes that the population is out of equilibrium and that the circumstances which cause unequal S-allele frequencies are transient, although if linkage between the two genes is tight, it could take many generations before equilibrium is attained. Hence, the observation that different alleles occur at different frequencies in different samples does not, by itself, allow us to distinguish, in general, between these alternative hypotheses.

This observation does allow us, on the other hand, to virtually rule out the possibility that the unequal allele frequencies are due to an extra effect of selection that acts pleiotropically on the rate of pollen tube growth, for if this were the case, we would expect the same S-alleles to occur at a relatively high and stable frequency in different populations provided we assume, not unreasonably, that this extra effect of selection is not modified in any major way by the external environment.

Acknowledgements

We are grateful to acknowledge the assistance we received in the early years of this investigation from Jo Moriarity, Paul Driver, Davinda Arora and John Booth, each of whom were MSc students in receipt of an SERC Advanced Course Studentship. In addition, one of us (S.O'D.) was supported by a SERC Research Studentship, two others (M.D.L. and V.E.F-T.) by an NERC Research Studentship and the fourth (M.J.L.) by an NERC Research Grant, which we are grateful to acknowledge. Lastly, this work could not have been carried out without the very able assistance of our horticultural staff, John Martin, Charlie Osgood and Michael Robertson, who raised and cared for the many hundreds of plants required for this investigation.

References

- CAMPBELL, J. M. AND LAWRENCE, M. J. 1981a. The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. I. The number and distribution of S-alleles in families from three locations. *Heredity*, **46**, 69–79.
- CAMPBELL, J. M. AND LAWRENCE, M. J. 1981b. The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. II. The number and frequency of S-allele in a natural population (R106). *Heredity*, **46**, 81–90.
- EMERSON, S. 1939. A preliminary survey of the *Oenothera organensis* population. *Genetics*, **24**, 528–537.
- FISHER, R. A. 1961. Possible differentiation in the wild population of *Oenothera organensis*. *Aust. J. Biol. Science*, **14**, 76–78.
- FRANKLIN-TONG, V. E., LAWRENCE, M. J. AND FRANKLIN, F. C. H. 1988. An *in vitro* bioassay for the stigmatic product of the self-incompatibility gene in *Papaver rhoeas* L. *New Phytologist*, **110**, 109–118.
- LANE, M. D. 1990. *The population genetics of the self-incompatibility polymorphism in Papaver rhoeas* L. Ph.D. Thesis, University of Birmingham.
- 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.
- LAWRENCE, M. J. AND O'DONNELL, S. 1981. The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. III. The number and frequency of S-alleles in two further natural populations (R102 and R104). *Heredity*, **47**, 53–61.
- O'DONNELL, S. AND LAWRENCE, M. J. 1984. The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. IV. The estimation of the number of alleles in a population. *Heredity*, **53**, 495–507.
- OOI, S. C. 1970. Variation in wild populations of *Papaver rhoeas* L. Ph.D. Thesis, University of Birmingham.