The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. VII. The number of S-alleles in the species

M. D. LANE & M. J. LAWRENCE*

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

The results obtained from the cross-classification of the S-alleles of a sample of 30 plants raised from the seed taken from a Spanish population against those of a sample taken from a British population revealed that 24 alleles of the former also occurred in the latter, giving an estimated overlap of 53 per cent between the allelic complements of these populations. These data indicate that the number of alleles in the species is unlikely to be much greater than 66. It is argued that the number of alleles in the species may be subject to a dynamic restraint on the number of alleles contained in its constituent populations and possibly, also, by a molecular restraint on the allelic diversity of the S-gene.

Keywords: number of S-alleles, *Papaver rhoeas*, population genetics, self-incompatibility.

Introduction

It was argued in the previous paper of this series that while the long-term effect of the frequency dependent selection which maintains the self-incompatibility polymorphisms might be to generate a very large number of different S-alleles in the species, any one of its constituent populations, being finite in size, is expected to contain only a subset of these alleles and that, consequently, different populations would contain different subsets of the alleles of the species. The results obtained from the cross-classification of the alleles of samples taken from three natural populations of the species indicated that, contrary to this expectation, these populations appear to contain essentially the same set of alleles (Lawrence *et al.*, 1993; O'Donnell *et al.*, 1993). In discussing these results, it was suggested that either the number of S-alleles in local populations is limited because the British Isles is at the north-western periphery of the geographical range of the species or that, despite theoretical expectation to the contrary, the number of S-alleles in the species as a whole is not very large, so that its constituent populations, provided that they are long-established and large, come to contain by mutation most, if not quite all, of the alleles of the species. The obvious way of testing

these hypotheses is to find out whether a population that is geographically remote from British populations contains a different subset of alleles, as predicted on the first hypothesis, or substantially the same set of alleles as the British populations, as predicted on the second hypothesis. In this paper, we report the results obtained from the cross-classification of the alleles of a sample taken from a Spanish population against those of one of the British populations we have examined.

Materials and methods

The 30 Spanish plants used in this cross-classification experiment were raised from seed collected from a natural population found in the province of Cordoba in Southern Spain by P. Gibbs of the Department of Biology and Pre-Clinical Medicine, University of St. Andrews, to whom we are indebted for this material. The seed was collected by taking one or more ripe capsules from each of 24 well-spaced plants chosen at random, the seed from each plant being packeted separately from that of every other, following the procedure employed when collecting seed from the British natural populations. The original intention was to raise one plant from each packet of seed taken from this Spanish population. Unfortunately, the package containing this seed became damaged in transit, so that seed from different plants became mixed. The 30

*Correspondence.

Spanish plants raised in this experiment, therefore, unavoidably originated from a bulk sample of seed from this population. The disadvantage of plants raised from seed in bulk for present purposes is that when two or more are found to contain the same S-allele, one cannot be sure whether this is due to the relatively high frequency of this allele in the population from which the sample has been drawn, an inference which is valid when the seed of each wild plant is kept separate from that of every other, or because these plants were raised from seed that came from the same wild plant, in which case they are, of course, expected to contain one or other of the maternal S-alleles (and, indeed, may also share one or more paternal alleles if the bulk of the pollen giving rise to this seed came from a single male parent). In this latter case, the relatively high frequency of an allele in a small sample may indicate no more than common ancestry rather than genuine differences in allele frequency in the population.

The dormancy of the Spanish seed was checked in a small-scale germination experiment in which seed was incubated at 25°C on moist filter pads. Although the germination percentage of this seed was, at 36 per cent, much higher than that obtained with untreated seed from British populations (2–5 per cent), the seed was nevertheless given a short acid treatment of 30s to break its residual dormancy, which gave an acceptably high germination percentage.

Although the taxonomic identity of this material was not in serious doubt, some species of the genus *Papaver* cannot easily be distinguished on sight in the wild. Selfs and crosses were made, therefore, the latter involving both Spanish plants and those of British origin. No seed was obtained from the self-pollination of any of the Spanish plants, indicating that they were as self-incompatible as British plants. The amount of seed yielded by the inter-population (Spanish × British) crosses was similar to that obtained from intra-population crosses, both Spanish and British. Furthermore, plants raised in the following season from the seed of these inter-population crosses were as fertile as those raised from the seed of intra-population crosses. There is no reason to doubt, therefore, that the original identification of the plants of the Spanish population as *P. rhoeas* was correct.

Each of the 30 Spanish plants was crossed to each of ten pairs of fully compatible testers from the R104 population whose incompatibility genotype had been ascertained previously, because they had been used in the final large experiment involving the cross-classification of the alleles of the R102 sample against those of the R104 sample (Lawrence *et al.*, 1993). Between them, the ten pairs of testers contained all of the 25 S-alleles that had been identified in the R104 sample.

Though the cross-classification of the alleles of the Spanish sample against those of the R104 testers was very nearly complete, this experiment differed from previous cross-classification experiments in one important respect. Thus, while the genotypes of the R104 testers could be ascertained because each pair was a member of a full-sib family produced by crossing parents whose genotype had been determined in a previous experiment (Lawrence & O'Donnell, 1981), the incompatibility genotypes of the Spanish plants were not known prior to their cross-classification against the R104 testers. Hence, the number of S-alleles contained in the Spanish sample was neither known in advance nor could be determined in the present experiment.

Results

This experiment was carried out towards the end of a winter season when the plants involved had been in flower for nearly five months, so that some of them were past their best. Although attempts were made to overcome this problem by repeating pollinations when the outcome of the first was not certain and by using stored pollen that had been gathered from some of these plants earlier in the season, the classification of some of them could not be completed or remained in doubt at the end of the experiment. Thus, 28 (4.7 per cent) of the pair-wise comparisons between the Spanish plants and the R104 testers gave pollinations which were unclassifiable, 5 (0.8 per cent) were scored as probably half-compatible and 36 (6 per cent) as probably fully compatible. While problems with the classification of pollinations were occasionally encountered in previous cross-classification experiments, they were resolved by repeating questionable comparisons in later experiments, an opportunity not available in the present experiment because of the status of the Spanish material. Nevertheless, 531 (88.5 per cent) of the pollinations were classified, 38 of these being clearly half-compatible and 493 fully compatible. Thus, while these results are perhaps less well-founded than those from previous experiments, they are not much less so.

A summary of these results is shown in Tables 1 and 2. Sixteen of the Spanish plants turned out to possess one of the R104 S-alleles, four plants contained two and the remaining ten possessed alleles that did not occur in the R104 sample. Thus, in all, 24 of the 60 S-alleles contained in the Spanish sample occurred in both samples. The identity of the R104 alleles that occurred in nine of the Spanish plants could not be completely resolved because the alleles in question occurred only once among the 20 R104 testers. This problem did not arise in previous experiments because

one or other of a pair of alleles involved in an identity always occurred more than once in the R104 sample, so that it also occurred more than once in the 19 full-sib families produced by crossing the plants whose incompatibility genotype had been determined in the R104 diallel experiment (Lawrence & O'Donnell, 1981). Thus, of the three alternatives pairs shown in Tables 1 and 2, only S_2 and S_{27} occurred once in the R104 sample, S_4 occurring three times, S_8 five times and S_{17} and S_{18} twice. Hence, each of these alternatives could have been decided one way or another if the full set of 19 R104 families had been raised in this experiment rather than the minimum set of ten containing all 25 of the R104 alelles. To have done this, however,

would have meant virtually doubling the size of the experiment which was regarded as too high a price to pay for a rather small gain in information. The ambiguities concerning two of these identities, namely S_2 or S_4 and S_{17} or S_{18}, raise no problems because, as each occurs only once, only one or other of these pairs of alleles can be involved in the identity. The ambiguity concerning the third of these alternative pairings cannot be resolved, however, because some of the seven identities may have involved S_8 and others S_{27}. If this were the case then eight, rather than seven R104 alleles also occurred in the Spanish sample. Nevertheless, in what follows, we have generally prefered the more conservative assumption that only seven alleles occur in both samples.

The final point worth making about these results is to note that the frequencies with which the R104 alleles occur in the Spanish sample are very unequal (Table 2). Thus, while four of these alleles, S_2 (or S_4), S_5, S_7 and S_{17} (or S_{18}) occur only once and S_{14} twice, S_8 (or S_{27}) occurs seven and S_{26} eleven times in this small sample. Since the Spanish plants were raised from a bulk sample of seed, however, some of these plants may be related as half-sibs. While this must be regarded as a distinct possibility for the seven identities involving S_8 or S_{27}, it appears less likely for the 11 identities concerning the R104 allele S_{26} which may, therefore, occur at a higher frequency in the Spanish population than the other alleles, although on present evidence, this can be no more than a surmise.

Lastly, there appears to be no obvious relationship between the frequency of the alleles in the Spanish sample and those of the same alleles in either the R104

Table 1 Partial or complete genotype of the 20 Spanish plants that contained R104 alleles. The ten plants not shown here were fully compatible with each of the 20 R104 testers and, hence, contained none of the 25 S-alleles of the R104 sample

Plant no.	Genotype	Plant no.	Genotype
1	S_7	14	S_2 or S_4
2	S_8 or S_{27}	15	S_8 or S_{27}
3	S_{26}	16	S_{17} or S_{18}
5	S_{26}	17	S_5S_8 or S_5S_{27}
7	S_{26}	18	S_8 or S_{27}
8	S_8 or S_{27}	21	S_{26}
9	S_8S_{26} or $S_{26}S_{27}$	23	S_{26}
10	S_{14}	25	S_{26}
12	S_{26}	26	S_{26}
13	S_8S_{26} or $S_{26}S_{27}$	28	$S_{14}S_{26}$

Table 2 Frequency of occurrence of the R104 alleles that also occur in the Spanish sample and the frequency of their occurrence in the British samples, R102, R104 and R106

R104 alleles occurring in Spanish sample	Frequency of occurrence		Frequency of occurrence in British samples		
			R102	R104	R106
S_2 or S_4	1	S_2	0	1	0
		S_4	3	3	2
S_5	1		3	2	11
S_7	1		3	9	7
S_8 or S_{27}	7	S_8	0	5	0
		S_{27}	5	1	2
S_{14}	2		1	4	4 or 2
S_{17} or S_{18}	1	S_{17}	3	2	5
		S_{18}	3	2	0
S_{26}	11		1	4	1
Total no. of alleles examined	60		72	71	102

or other British samples which, although similar to the conclusions drawn from the latter (Lawrence *et al.*, 1993), is vitiated by our uncertainty about the cause of the unequal allele frequencies in the former.

Discussion

As the results obtained from the cross-classification of the alleles of the Spanish sample against those of the R104 sample show that no less than 24 out of a total of 60 alleles of the former also occur in the latter, it is clear that there must be a considerable overlap between the complements of alleles of the natural populations from which these samples have been drawn. In order to estimate this overlap, using the procedure given in the previous paper (O'Donnell *et al.*, 1993), we need to know the number of alleles in each sample (n_1 and n_2), the number of alleles that occur in both (y) and the estimated number of alleles in each population ($\hat{N}_1$ and $\hat{N}_2$). The number of alleles in the R104 sample was $n_1 = 25$, the estimate of the number of alleles in the R104 population was $\hat{N}_1 = 32$ (Lawrence *et al.*, 1993) and the number of alleles identified as occurring in both was, from the previous section, $y = 7$. The total number of alleles occurring among the Spanish plants is not known because the classification of their compatibility genotype was incomplete. If we are to assume, however, that the seven R104 alleles identified in the Spanish plants can be regarded as a randomly drawn subset of the alleles in this sample, an approximate estimate of the number of alleles contained in the 30 plants may be obtained as $(7 \times 60)/24 = 17.5$, which gives an estimate of $\hat{n}_2 = 18$ by rounding off to the nearest integer.

An estimate of the number of alleles in the Spanish population is less easily obtained. Thus, in the absence of a knowledge of the frequency of occurrence of the alleles in the Spanish sample because of the incomplete classification of these plants, it is not possible to use the E_2 estimator of O'Donnell and Lawrence (1984); nor is it reasonable to assume, in view of the results given in Table 2, that the genotype frequencies are equal in the population from which this sample has been drawn, which is an assumption made with the maximum likelihood procedure (Paxman, 1963). We shall assume, therefore, that the number of alleles in the Spanish population is similar to the number of alleles in the British populations examined, which averaging over the estimates obtained from each, gives $N_2 = 35$. Then with $n_1 = 25$, $\hat{n}_2 = 18$, $\hat{N}_1 = 32$, $\hat{N}_2 = 35$ and $y = 7$, the number of alleles that the Spanish and R104 populations have in comon is $\hat{x} = 17$, giving an estimated overlap between their complements of alleles of 53 per cent.

This estimate is, of course, less reliable than those obtained from the British populations because we have had to estimate, rather than determine, the number of alleles in the Spanish sample, n_2, and have had to assume, in the absence of any obvious alternative, that the number of alleles in the Spanish population, N_2, is similar to that in the British populations. We have, therefore, examined the effect of varying both y and N_2 on this estimate of overlap. It turns out that even if y were reduced to six, so as to exclude the questionable identity involving the presence of the R104 allele S_7 in the Spanish sample and $\hat{N}_2$ is reduced to the R104 value of 32, which was the lowest of the estimates from British populations, the estimated overlap between the allelic complements of these populations, although reduced to 44 per cent, still indicates that this is substantial. Conversely, if y is held at 7 and $\hat{N}_2$ increased to 38, the highest estimate from a British population (R106), the overlap only rises to 59 per cent (see Lane, 1990 for details). Again, if we assume that both of the R104 alleles, S_7 and S_{28}, rather than just one or other of them, occurred in the Spanish sample, y rises to 8 so that with $\hat{N}_2 = 35$, $\hat{x} = 18$, giving an overlap of 56 per cent. Lastly, it could be argued that because *P. rhoeas* is generally regarded as having a Mediterranean origin (McNaughton & Harper, 1964), populations in that region might contain more S-alleles than those on the periphery of its range, so that in taking the average of the British populations, we have underestimated the number of alleles in the Spanish population. Suppose, then, that $\hat{N}_2 = 50$. Then with $y = 7$, $\hat{x} = 25$ and our estimate of the overlap between the allelic complements of the R104 and Spanish populations rises to 78 per cent.

Taken at face value, these calculations suggest that whereas our estimate of overlap depends, as expected, on the values we choose for y and N_2, we are unlikely to be seriously misled by errors in estimation of these parameters, because all estimates of the percentage overlap indicate that this is substantial. The least satisfactory aspect of the application of O'Donnell's procedure to the present data is that y and n_2 are not independent because the latter is estimated as $(y \times 60)/n_i$, where n_i is the number of alleles in the Spanish sample which also occur in the R104 sample (i.e. 24). In consequence, n_2 rises and falls with y. Though y cannot be greater than either n_1 or n_2, whichever is the smaller, in general n_2 will be independent of y when the former is found from data obtained from a fully classified sample of plants, as was the case with n_1. Not knowing the actual value of n_2 for the Spanish sample, we have no alternative but to estimate this statistic if we are to attempt to estimate

the overlap between the allelic complements of the Spanish and British populations.

The chief purpose of this experiment was to attempt to distinguish between two hypotheses concerning the virtually complete overlap between the allelic complements of the three British populations we have examined (O'Donnell *et al.*, 1993). The first of these hypotheses proposed that the number of S-alleles in these populations is limited because the British Isles are on the north-western periphery of the geographical range of the *P. rhoeas* and, hence, contain only a subset of the alleles of the species. The second hypothesis, however, proposed that despite theoretical considerations to the contrary, the number of S-alleles in the species is not very large and that British populations contain most, if not quite all, of these alleles.

The results we have obtained from the present experiment are not wholly consistent with either of these hypotheses. On the one hand, they can be regarded as supporting the first hypothesis because 47 per cent of the alleles of the Spanish population do not appear to occur in the British populations, which suggests that the latter do indeed possess only a subset of the alleles of the species. On the other hand, that a population 1700 km distant from the British populations appears to contain a set of alleles 53 per cent of which also occur in the latter, suggests that the number of alleles in the species is not all that large. These results, therefore, can be regarded as at least partly consistent with the second hypothesis. Indeed, as these hypotheses are by no means mutually exclusive, it is, of course, possible that both hold with these data, although further evidence is clearly required before we could be confident that this is so.

An approximate estimate of the number of alleles in the species can be obtained from these data from Fisher's (1961) missing plot procedure. Thus, with $\hat{N}_1 = 32$, $\hat{N}_2 = 35$ and $\hat{x} = 17$, the number of alleles occurring in the R104 population but not in the Spanish one, can be found as $\hat{N}_1 - \hat{x} = 15$; the number of alleles occurring only in the Spanish population as $\hat{N}_2 = \hat{x} = 18$ and, as we have seen, the estimated number of alleles occurring in both is $\hat{x} = 17$. The total number of alleles that appear to be present in these populations is, therefore 50. From Fisher's procedure, we find that the number of alleles in the species that occur in neither population is 15.88. Hence, the estimated number of alleles in the species is found as $50 + 15.88 = 65.88$ or 66 to the nearest whole number.

It is, of course, questionable whether the use of Fisher's method is valid with numbers that are estimates, rather than those observed in samples taken from two or more populations. An alternative estimate of the number of alleles in the species can be obtained

by recalling that the actual number of alleles identified in the samples taken from the British populations was 45 (Lawrence *et al.*, 1993). Then because 47 per cent of the alleles of the Spanish sample appeared not to occur in the R104 sample, an estimate of the number of alleles in the species can be found from $45 + (0.47 \times 45) = 66.15$. While this estimate is not wholly independent from that obtained by Fisher's method, the fact that it is very similar to the latter is encouraging.

There are two reasons for believing that these estimates may be biased. First, as pointed out earlier, if both of the R104 alleles, S_8 and S_{27}, occurred in the Spanish sample, the estimated overlap rises to 56 per cent. Fisher's method then yields an estimate of 62 alleles and the second method, an estimate of 65. Again, if the number of alleles in the Spanish population is appreciably larger than in the British populations (e.g. $\hat{N}_2 = 50$), our estimate the overlap between the allelic complements of the Spanish and British populations rises to 78 per cent, which gives estimates of 64 and 59 from Fisher's and our method of estimation, respectively. Second, as the cross-classification of the alleles of the R106 sample against those of the R102 and R104 samples was less complete than that of the latter pair, there are grounds for suspecting that some of the 12 alleles that apparently occurred only in the former may, in fact, also occur in one or other of the other samples (see Table 6 of Lawrence *et al.*, 1993). If this were the case, the total number of alleles in the British samples would be less than 45 so that our second estimate of the number of alleles in the species would be less than 66. Hence, for either or both of these reasons, it is possible that our rather approximate estimate of 66 alleles is biased upwards, although, as our calculations show, the magnitude of this bias is likely to be small.

Now, the number of alleles in the species must, of course, be related to the number of alleles in its constituent populations. If the gene frequency-dependent selection that maintains the polymorphism had an unconditional effect on increasing the number of alleles in the population, we would expect both the number in the latter and the species as a whole to be large. As our estimate of 66 alleles for the species is clearly not consistent with this expectation, the effect of this selection cannot be unconditional. In the previous paper (O'Donnell *et al.*, 1993), we drew attention to the fact that the strength of this selection becomes progressively attenuated as the number of alleles in the population rises; that is, the selective advantage of new allele that enters the population by mutation or migration is negatively related to the number of alleles already present. For example, whereas cross-incompatibility,

the force driving the frequency dependent selection, is no less than 33.33 per cent in a population containing only three equally frequent alleles, it is only 0.53 per cent in one containing 20 equally frequent alleles. Why, then, do we find populations that appear to contain substantially more than this number of alleles?

There are at least two reasons why this argument is too simple. First, the percentage incompatibility in a population containing a given number of unequally frequent alleles could be greater than in one containing the same number of equally frequent alleles. As the allele frequencies appear to be significantly unequal in each of the three British populations we have examined (Campbell & Lawrence, 1981; Lawrence & O'Donnell, 1981) this point is of obvious relevance to *P. rhoeas* and suggests that it would be worth investigating the reduction in the rate of attenuation of the frequency dependent effect when allele frequencies are allowed to become unequal by computer simulation. Secondly, the argument assumes that the population is panmictic. It is possible, however, that a population which is substructured into a number of groups of related individuals, because both pollen and seed dispersal is limited, might be capable of carrying more alleles at equilibrium that one in which mating is random. If true, this might explain why populations of long-lived perennial species that reproduce vegetatively, like *Trifolium repens* (Atwood, 1944) and *T. pratense* (Williams & Williams, 1947) appear to contain many more alleles than those of either *Oenothera organensis* (Emerson, 1939, 1940) or *P. rhoeas*. In view of Wright's (1939) investigation of the amount of isolation required between subpopulations of *Oe. organesis* to maintain the observed number of alleles in a population, considered at that time to consist of only 500 individuals, the conditions under which a substructured population could maintain a greater number of alleles than a panmictic one of the same size could turn out to be quite restrictive. Nevertheless, this possibility could be investigated by computer simulation of the polymorphism. The chief point that emerges from this discussion, however, is that the number of alleles in a population could well be subject to what it is convenient to refer to as a dynamic restraint, due to the attenuation of the strength of frequency dependent selection as the number of alleles in the population rises, even in a population that is very large.

The other major factor determining the number of alleles of a species is the number of functional allelic forms in which is it possible for the S-gene to exist. While it is convenient to assume that there are an infinite number of alleles, each newly arising mutation being unique, when investigating the mathematical properties of the polymorphism, this assumption is unlikely to hold in practice. Nevertheless, if the number of functional allelic forms that can be coded for by the S-gene is very large and the mutation rate from one such form to another is not very small, the number of alleles in the species would be expected to be much larger than the number in any of its constituent populations because of the effect of drift in the latter (Wright, 1939). Conversely, if the number of alleles in the species is relatively small, large and long-established populations might be expected to contain most of them, so that each such population would be expected to contain essentially the same complement of alleles. In these circumstances, the number of alleles in the species and, hence, in its constituent populations, would be subject to what can be termed a molecular restraint. Furthermore, as the genetical evidence suggests that self-incompatibility has originated independently a number of times during the evolution of the flowering plants, it is possible that the coding capacity of the S-gene may be greater in some species than others. For example, the coding capacity of the gene of *T. repens* may be greater than that of *P. rhoeas*. Given that they exist, differences in coding capacity might be due to an 'accident' of evolution, in which case the S-genes of the self-incompatible species of a family of flowering plants would be expected to have similar coding capacities. Alternatively, such differences might be the result of different adapative responses to different reproductive biologies of the kind mentioned earlier; if so, we would not expect coding capacity to be approximately constant in families whose species had different reproductive strategies. In the absence of a knowledge of the molecular basis of allelic diversity we have no way at present of distinguishing between these alternatives. Now that the stigmatic version of the S-gene in *P. rhoeas* has been cloned (H. C. Foote *et al.,* unpublished data), however, we should in the not too distant future be able to find out whether the relatively small number of S-alleles in this species is due to a molecular restraint.

References

ATWOOD, S. S. 1944. Oppositional alleles in natural populations of *Trifolium repens*. *Genetics*, **29**, 428–435.

CAMPBELL, J. M. AND LAWRENCE, M. J. 1981. The population of the self-incompatibility polymorphism in *Papaver rhoeas*. II. The number and frequency of S-alleles in a natural population (R106). *Heredity*, **46**, 81–90.

EMERSON, S. 1939. A preliminary survey of the *Oenothera organenesis* population. *Genetics*, **24**, 524–537.

EMERSON, S. 1940. Growth of incompatible pollen in *Oenothera organensis*. *Bot. Gaz.*, **101**, 890–911.

FISHER, R. A. 1961. Possible differentiation in the wild population of *Oenothera organensis*. *Aust. J. Biol. Sci.*, **14**, 76–78.

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. 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.

LAWRENCE, M. J., LANE, M. D., O'DONNELL, S. AND FRANKLIN-TONG, V. E. 1993. The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. V. The cross-classification of the S-alleles of samples from three populations. *Heredity*, **71**, 581–590.

McNAUGHTON, I. M. AND HARPER, J. L. 1964. Biological Flora of the British Isles. No. 99. Papaver L. *J. Ecol.*, **52**, 767–793.

O'DONNELL, S. AND LAWRENCE, M. J. 1989. 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.

O'DONNELL, S., LANE, M. D. AND LAWRENCE, M. J. 1993. The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. VI. The estimation of the overlap between the allelic complements of a pair of populations. *Heredity*, **71**, 596–602.

PAXMAN, G. J. 1963. The maximum likelihood estimation of the number of self-sterility alleles in a population. *Genetics*, **48**, 1029–1032.

WILLIAMS, W. AND WILLIAMS, R. D. 1947. Genetics of red clover (*Trifolium pratense* L.) compatibility. III. The frequency of S-alleles in two non-pedigree populations of red clover. *J. Genet.*, **48**, 69–72.

WRIGHT, S. 1939. The distribution of self-sterility alleles in populations. *Genetics*, **24**, 538–552.