

# Number of incompatibility alleles in clover and other species

M. J. LAWRENCE

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

Atwood's (1942, 1944) data on *Trifolium repens*, those of Williams & Williams (1947) on *T. pratense* and those of Williams (1951) on *T. hybridum* have been re-analysed to provide maximum likelihood estimates of the number of incompatibility alleles in the populations or breeders' stocks from which the samples investigated were obtained. These new estimates suggest that populations of *T. repens* contain about 100 alleles and those of *T. pratense* contain up to twice this number. The single estimate from *T. hybridum*, however, suggests that the species in North America possesses only 17 *S*-alleles. The estimates from clover populations are compared with those from the nine most thoroughly investigated species of other self-incompatible flowering plants and confirm the long-held belief that populations of *T. repens* and *T. pratense* contain more *S*-alleles than those of the latter. It is argued that the most likely explanation of the large number of *S*-alleles that natural populations of these species appear to contain is that they are substructured into a large number of semi-isolated neighbourhoods.

**Keywords:** number of incompatibility alleles, self-incompatibility, substructure, *Trifolium* spp.

## Introduction

It has long been apparent that populations of self-incompatible species of clover, in which self-incompatibility is controlled by a single, multiallelic gene expressed gametophytically in the pollen (Williams & Silow, 1933; Atwood, 1940; Williams, 1951), contain a large number of *S*-alleles and more recently, that this number is larger than the number of *S*-alleles in populations of other self-incompatible species. This conclusion is based on the number,  $n$ , of different alleles found when  $m$  alleles have been examined, because, with one exception, these data have not been used to estimate the number of alleles,  $N$ , in the populations from which the samples were taken. It would be convenient if such estimates were available to compare them with those from other self-incompatible species. This is the chief purpose of this paper.

## Data and analytical procedure

Estimates of  $N$  have been obtained from the data of Atwood (1942, 1944) on *Trifolium repens*, those of Williams & Williams (1947) on *T. pratense* and from

those of Williams (1951) on *T. hybridum*. All three of these investigators used the same analytical procedure whereby each plant of a sample was crossed for seed with a tester that was homozygous for the *S*-gene, these testers having been produced prior to the investigations by the forced self-pollination of individuals that displayed a very small amount of pseudo-self-compatibility. Each cross was, therefore, of the type  $S_i S_j \times S_1 S_1$ , where  $S_i$  and  $S_j$  were alleles from the population in question and, hence, was expected to produce two equally frequent classes of offspring,  $S_1 S_i$  and  $S_1 S_j$ , respectively. All pollinations between the offspring of different parents, as well as those of the same parent are, in these circumstances, expected to be either half-compatible or incompatible, indicating either that  $S_i \neq S_j$ , or  $S_i = S_j$ , respectively. The number of *S*-alleles in a sample was determined by making all possible crosses, diallel fashion, among these offspring, the compatibility of a cross being determined by seed-set.

If each parent is represented in a diallel by a half-compatible pair of offspring, both of its alleles can be assayed and the conventional maximum likelihood equation

$$\varepsilon n = N[1 - (1 - 2/N)^n], \quad (1)$$

\*Correspondence.

where  $r$  is the number of plants in the sample ( $m = 2r$ ), can be used to estimate the number of alleles in the population, from which the sample has been drawn, by setting  $en$  equal to  $n$  and solving for  $N$  by iteration (Paxman, 1963; O'Donnell & Lawrence, 1984). If, on the other hand, only one offspring from each parent is included in the diallel, the appropriate maximum likelihood equation is

$$en = N[1 - (1 - 1/N)^r] \quad (2)$$

where  $r$  now equals  $m$ .

Atwood (1942) examined the progeny of two sets of plants which were taken from collections of potential breeding material of *T. repens* of predominantly North American origin. In the first, he found that 25 *S*-alleles of 26 tested were different and in the second that 34 *S*-alleles of 41 were different. Each of these sets contained seven homozygotes, however, and, in the first, one of the 10 parents was represented in the diallel by only one genotype; the results from these plants have been excluded from analysis in order to obtain balanced sets of data, giving  $n = 17$  and  $r = 9$  for the first set and  $n = 28$  and  $r = 17$  for the second. Estimates of  $N$  can then be obtained from each set by the use of eqn 1.

Atwood (1944) analysed the progeny of 49 plants taken from two long-established pastures in central Pennsylvania, where white clover occurs naturally, and found that 36 *S*-alleles of the 49 tested in the first were different and that 39 *S*-alleles of the 49 tested in the second sample were different. These data can be used as they stand. Each of the parents of these samples was represented by only one plant in these diallels, so eqn 2 has been used to estimate  $N$ .

The data of Williams & Williams (1947) from their first sample of 24 plants taken from a population known as English late-flowering red clover, in which they found that 41 *S*-alleles of 48 examined were different, can also be used as they stand. It is necessary to exclude four plants from the data of their second sample, taken from a population known as English broad red clover, in which they found that 37 *S*-alleles of 40 examined were different, because these were single representatives of their parents, giving  $n = 33$  and  $r = 18$ . Equation 1 can then be used to estimate the number of alleles in each of these populations.

Williams (1951) investigated the number of *S*-alleles in a sample of 11 *T. hybridum* plants raised from commercial seed of North American origin and found that 13 *S*-alleles of 21 examined were different. One of these plants, however, inadver-

tently contributed only a single genotype to the diallel. If the data from this plant are excluded,  $n = 12$  and  $r = 10$ , and eqn 1 can be used to estimate  $N$ .

The derivation of the maximum likelihood eqns 1 and 2 assumes that the incompatibility genotypes in the populations from which the samples have been drawn at random are equal (Paxman, 1963; O'Donnell & Lawrence, 1984). Although tests of equal allele frequency (Campbell & Lawrence 1981b) have not been carried out on any of these data sets, because none contains sufficient information to make this worthwhile, there is no indication in any of them that the frequencies of the alleles are other than equal. The use of these estimators, therefore, seems to be justified.

## Results

The estimates of the number of *S*-alleles in these clover populations, shown in Table 1, confirm the belief that they contain a large number of these alleles in the case of *T. repens* and *T. pratense*, but not in the case of *T. hybridum*. The estimate of 101 alleles from Atwood's (1944) second sample may be higher than that from the first because he stated that whereas the first population contained 'a few thousand' plants, the second contained 'a few hundred thousand' plants. Estimates from breeders' stocks (Atwood, 1942) concern the number of alleles in the species, because they were composed of material collected from a number of different populations; these estimates are expected to be larger than those from single populations. It is surprising, therefore, that these estimates are not much larger than those from Atwood's 1944 data. It is possible that the first of these species estimates is larger than the second because it included material collected from New Zealand as well as from North America, whereas the second sample contained material from the latter source only.

The number of *S*-alleles in each of the two non-pedigree populations of *T. pratense* (Williams & Williams, 1947), is even higher than that in *T. repens* populations. The present estimates are smaller than those of Paxman (1963) and O'Donnell & Lawrence (1984), which were based on Bateman's (1947) rather questionable interpretation of Williams & Williams' data.

The estimate of only 17 *S*-alleles in *T. hybridum*, derived from two North American commercial varieties, is unexpectedly low and may be a consequence of these varieties being founded on a very small

**Table 1** Maximum likelihood estimates of the number of *S*-alleles, *N*, in populations of three species of clover

Species	Material	<i>n</i>	<i>m</i>	<i>R</i>	<i>N</i>
<i>T. repens</i> <sup>1</sup>	Breeders stock (1)	17	18	0.07	<b>139</b>
	Breeders stock (2)	28	34	0.19	<b>80</b>
<i>T. repens</i> <sup>2</sup>	Natural population (1)	36	49	0.28	<b>74</b>
	Natural population (2)	39	49	0.22	<b>101</b>
<i>T. pratense</i> <sup>3</sup>	Commercial variety (1)	41	48	0.16	<b>143</b>
	Commercial variety (2)	33	36	0.09	<b>193</b>
<i>T. hybridum</i> <sup>4</sup>	Commercial varieties	12	20	0.47	<b>17</b>

*n* is the number of alleles found, *m* the number of alleles examined and *R* is the repeatability of the assay (Campbell & Lawrence, 1981a). Superscripts on species names indicate source of data: 1, Atwood (1942); 2, Atwood, (1944); 3, Williams & Williams (1947); 4, Williams (1951).

number of plants, rather than this species possessing fewer alleles than the other clovers.

Estimates of the number of incompatibility alleles in populations of the nine most thoroughly investigated species of other flowering plants (Table 2) confirm that these are lower than those of the clovers, despite the fact that the former differ in respect of the genetics of their incompatibility system, their life history and their ecology. On this evidence, therefore, the self-incompatible species of flowering plants appear to fall into two distinct groups: one containing the clovers and the other, the rest.

## Discussion

The obvious question raised by these results is why populations of clovers appear to contain more alleles than those of other species. Three possible reasons can be advanced in an attempt to answer this question. First, as Emerson (1938) showed, the classification of pollinations by seed-set is less reliable than by the direct observation of the behaviour of pollen in or on the pistil. The probability of misclassification is, thus, higher with the former than with the latter. Whenever undetected misclassification occurs, the conclusion drawn from the data is always that the number of alleles in a sample is greater, or that the incompatibility system of the species in question is more complex than is actually the case (Lawrence *et al.*, 1983). Because there appears to be no distinct zone in which the growth of incompatible pollen is regularly inhibited in their pistils (our own observations and, independently, those also of T. Michaelson-Yates, personal communication), it is not possible to classify pollinations in either *T. repens* or *T. pratense* by observing

the behaviour of pollen in the style; all of the data from these species, therefore, is unavoidably based on seed-set. All of the investigators involved, however, were aware of the risk of misclassification and took considerable care to ensure that all plants were of the intended genotype. All initially doubtful pollinations were also repeated until a satisfactory outcome was obtained. Thus, although the repeatabilities of the clover assays, because of the large numbers of alleles identified, are generally lower than those of the species listed in Table 2, which ranged from 0.40 for *Phlox drummondii* to 0.72 for the third sample of *Papaver rhoeas*, this work in all other respects was at least as thorough as that carried out on other species and, indeed, was more thorough than most. The possibility that the estimates of the number of alleles in these clover populations has been inflated by undetected misclassification can, therefore, be virtually ruled out.

The second possibility is that the molecular basis of self-incompatibility in clovers is different from that of other species with the same genetical system, such that their allelic diversity at the *S*-locus is much higher. It has been shown that the molecular basis of self-incompatibility in several solanaceous species differs from that in *Papaver rhoeas* (McClure *et al.*, 1989; Clark *et al.*, 1990; Xu *et al.*, 1990; Singh *et al.*, 1991; for the former and Franklin-Tong *et al.*, 1991 for *P. rhoeas*), and so there is no difficulty in supposing that the clovers may constitute a third group in this respect. The ultimate limit to allelic diversity at the *S*-locus in populations of self-incompatible species appears, however, to be dynamic rather than molecular, because of the attenuation of the strength of frequency-dependent selection as the number of alleles in the population rises. Thus, as

**Table 2** Number of alleles, *N*, known or estimated to be present in populations and species of eight flowering plants

System	Species	In populations		In species	
		Known	Estimated	Known	Estimated
A	<i>Oenothera organensis</i> <sup>1</sup>	45	30	45	30
	<i>Papaver rhoeas</i> <sup>2</sup>	27	35+		
		25	32+	45	66
		31	38+		
	<i>Phlox drummondii</i> <sup>3</sup>	30	45	—	—
	<i>Solanum carolinense</i> <sup>4</sup>	11	12	13	—
		12	14		
	<i>Physalis crassifolia</i> <sup>5</sup>	28	44	—	—
B	<i>Lolium perenne</i> <sup>6</sup>	17S	31+S	—	—
		17Z	31+Z	—	—
C	<i>Brassica oleracea</i> <sup>7</sup>	—	—	49	—
	<i>Sinapis arvensis</i> <sup>8</sup>	35	—	—	—
	<i>Ipomoea trifida</i> <sup>9</sup>	5	—		
		15	—		
		19	—	49	—
		16	—		
		21	—		
	6	—			

The entries in the first column of the table show the genetical control of their self-incompatibility: A the one-locus, gametophytic system; B the two-locus, gametophytic system; and C the one-locus, sporophytic system. Superscripts on the species names indicate the sources of data and estimates which are as follows: 1, Emerson (1939, 1940) and O'Donnell & Lawrence (1984); 2, Campbell & Lawrence (1981b); Lawrence *et al.* (1993) and Lane & Lawrence (1993); 3, Levin (1993); 4, Richman *et al.* (1995); 5, Richman *et al.* (1996); 6, Fearon *et al.*, (1994); 7, Ockendon, (1985); 8, Stevens & Kay, (1989); 9, Koyama *et al.* (1994).

we have shown elsewhere (Lawrence *et al.*, 1994), the probability that a pollen grain carrying a particular *S*-allele alights at random on a stigma carrying the same allele in a population containing 200 equally frequent alleles is only 1 per cent. It is difficult to see how selection this weak would be sufficient to counteract the very considerable effect of drift revealed in the simulations of Brooks *et al.* (1996), unless the effective size of the population was exceedingly large. Hence, the coding capacity of the *S*-locus is by itself insufficient to account for the large number of alleles in clover populations. It is, of course, possible that the mutation rate at the *S*-locus in clovers is much higher than in other species; Wright's (1939) calculations show, for example, that a population whose effective size is no larger than 5000 could maintain 200 alleles at equilibrium if the mutation rate from one functional allele to another was  $1 \times 10^{-3}$ . The difficulty with this possibility is that there is no evidence that the mutation rate at the *S*-locus in clovers is any higher

than that in other species, in which it appears to be very low (Lewis, 1948, 1949, 1951; Hayman & Richter, 1992).

The third possibility, which arises from discussion of the second, is that clover populations are highly subdivided, rather than panmictic. Thus, as long-lived, vegetatively reproducing and insect-pollinated species, it might be expected that populations of clovers consist of a large number of semi-isolated neighbourhoods, between which gene flow is limited. But whereas Wright (1939) showed that a population with this structure could maintain a larger number of alleles at equilibrium than a panmictic one, the degree of isolation required for this to occur had to be considerable, a conclusion later confirmed by Crosby (1965) by computer simulation of the polymorphism. Brooks *et al.* (1996) showed, however, that although limited seed and pollen dispersal had a negligible effect on the distribution of *S*-alleles in their simulation of the polymorphism, the addition of variation in plant size caused a

marked spatial clustering of alleles. In these circumstances, the strength of frequency-dependent selection is stronger locally than in the population at large, so that a population with this structure should be able to maintain more alleles at equilibrium than a panmictic one. The vegetative spread of individual clover plants could cause extreme variation in plant size in long-established populations. It may not be a coincidence that a long-established population of another species with the capacity to reproduce vegetatively, *Lolium perenne*, appears able to maintain an even larger number of pollen incompatibility phenotypes than those of the clovers, as for a species with a two-locus system, this is the product of the number of alleles at the *S*-locus times the number at the *Z*-locus, which from the data given in Table 2, appears to be at least 961 for the population investigated by Fearon *et al.* (1994). In this sense, *L. perenne* should be grouped with the clovers, rather than with the other species of flowering plants.

Gliddon & Saleem (1985) found some evidence of substructuring with respect to isozyme polymorphism in one of the two natural populations of *T. repens* that they investigated, an effect which may have been reinforced by specific interactions between genotypes of this species and those of the other dominant species in the sward, *L. perenne* (Gliddon & Trathan, 1985). Hence, although there is much less evidence on this question than is desirable, the possibility that the large number of *S*-alleles in clover populations is the result of their subdivision appears to be the most likely of the three possibilities considered. It would, of course, be highly desirable for more recent data on the number, frequency and distribution of *S*-alleles in populations of these clover species to be available, which could be used to test this possibility.

### Acknowledgements

It is a pleasure to acknowledge the useful discussions I have had with my colleague, Dr D. F. Marshall, on some aspects of this paper. I am also grateful to Terry Michaelson-Yates of IGER, Aberystwyth, for his consent to cite his unpublished observations on the absence of a distinct zone of inhibition of incompatible pollen tubes in the styles of clover species.

### References

ATWOOD, S. S. 1940. Genetics of cross-incompatibility among self-incompatible plants of *Trifolium repens*. *J. Am. Soc. Agron.*, **32**, 955–968.

- 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.
- BATEMAN, A. J. 1947. Number of *S*-alleles in a population. *Nature*, **160**, 337.
- BROOKS, R. J., TOBIAS, A. M. AND LAWRENCE, M. J. 1996. The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. XI. The effects of limited pollen and seed dispersal, overlapping generations and variation in plant size on the variance of *S*-allele frequencies in populations at equilibrium. *Heredity*, **76**, 367–376.
- 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 localities. *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*-alleles in a natural population (R106). *Heredity*, **46**, 81–90.
- CLARK, K. R., OKULEY, J. J., COLLINS, P. D. AND SIMS, T. L. 1990. Sequence variability and developmental expression of *S*-alleles in self-incompatible and pseudo-self-compatible *Petunia*. *Pl. Cell*, **2**, 815–826.
- CROSBY, J. L. 1965. Self-incompatibility alleles in the population of *Oenothera organensis*. *Evolution*, **20**, 567–579.
- 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. Gaz.*, **101**, 890–911.
- FEARON, C. H., CORNISH, M. A., HAYWARD, M. D. AND LAWRENCE, M. J. 1994. Self-incompatibility in ryegrass. X. Number and frequency of alleles in a natural population of *Lolium perenne* L. *Heredity*, **73**, 254–261.
- FRANKLIN-TONG, V. E., ATWAL, K. K., HOWELL, E. C., LAWRENCE, M. J. AND FRANKLIN, F. C. H. 1991. Self-incompatibility in *Papaver rhoeas* L.: there is no evidence for the involvement of stigmatic ribonuclease activity. *Pl. Cell Env.*, **14**, 423–429.
- GLIDDON, C. AND SALEEM, M. 1985. Gene-flow in *Trifolium repens* – an expanding genetic neighbourhood. In: Jacquard, P., Heim, G. and Antonovics, J. (eds), *Genetic Differentiation and Dispersal in Plants*, pp. 293–309. NATO ASI Series G. Ecological Sciences. Springer-Verlag, Berlin.
- GLIDDON, C. AND TRATHAN, P. 1985. Interactions between white clover and perennial ryegrass in old permanent pasture. In: Haeck, J. and Woldendorp, J. W. (eds), *Structure and Functioning of Plant Populations 2*, pp. 161–169. North-Holland, Amsterdam.
- HAYMAN, D. L. AND RICHTER, J. 1992. Mutations affecting self-incompatibility in *Phalaris coerulea* Desf. (Poaceae). *Heredity*, **68**, 495–503.

- KOWYAMA, Y., TAKAHASHI, H., MURAOKA, K., TANI, T., HARA, K. AND SHIOTANI, I. 1994. Number, frequency & dominance relationships of S-alleles in diploid *Ipomoea trifida*. *Heredity*, **73**, 275–283.
- LANE, M. D. AND LAWRENCE, M. J. 1993. The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. VII. The number of S-alleles in the species. *Heredity*, **71**, 596–602.
- LAWRENCE, M. J., FEARON, C. H., CORNISH, M. A. AND HAYWARD, M. D. 1983. The genetical control of self-incompatibility in ryegrass. *Heredity*, **51**, 461–466.
- 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. Cross-classification of the S-alleles of samples from three natural populations. *Heredity*, **71**, 581–590.
- LAWRENCE, M. J., O'DONNELL, S., LANE, M. D. AND MARSHALL, D. F. 1994. The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. VIII. Sampling effects as a possible cause of unequal allele frequencies. *Heredity*, **72**, 345–352.
- LEVIN, D. A. 1993. S-gene polymorphism in *Phlox drummondii*. *Heredity*, **71**, 193–198.
- LEWIS, D. 1948. Structure of the incompatibility gene. I. Spontaneous mutation rate. *Heredity*, **2**, 219–236.
- LEWIS, D. 1949. Structure of the incompatibility gene. II. Induced mutation rate. *Heredity*, **3**, 339–355.
- LEWIS, D. 1951. Structure of the incompatibility gene. III. Types of spontaneous and induced mutation. *Heredity*, **5**, 399–414.
- MCLURE, B. A., HARING, V., EBERT, P. R., ANDERSON, M. A., SIMPSON, R. J., SAKIYAMA, F. AND CLARKE, A. E. 1989. Style self-incompatibility products of *Nicotiana glauca* are ribonucleases. *Nature*, **342**, 955–957.
- OCKENDON, D. J. 1985. Genetics and physiology of self-incompatibility in *Brassica*. In: Sussex, I. (ed.) *Current Communications in Molecular Biology*, pp. 1–6. Cold Spring Harbor, New York.
- 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.
- PAXMAN, G. J. 1963. The maximum likelihood estimation of the number of self-sterility alleles in a population. *Genetics*, **48**, 1029–1032.
- RICHMAN, A. D., KAO, T.-H., SCHAEFFER, S. W. AND UYENOYAMA, M. K. 1995. S-allele sequence diversity in natural populations of *Solanum carolinense* (Horsenettle). *Heredity*, **75**, 405–415.
- RICHMAN, A. D., UYENOYAMA, M. K. AND KOHN, J. R. 1996. S-allele diversity in a natural population of *Physalis crassifolia* (Solanaceae) (ground cherry) assessed by RT-PCR. *Heredity*, **76**, 497–505.
- SINGH, A., AI, Y. AND KAO, T.-H. 1991. Characterization of ribonuclease activity of three S-allele-associated proteins of *Petunia inflata*. *Pl. Physiol.*, **96**, 61–68.
- STEVENS, J. P. AND KAY, O. O. N. 1989. The number, dominance relationships and frequencies of self-incompatibility alleles in a natural population of *Sinapis arvensis* L. in South Wales. *Heredity*, **62**, 199–205.
- WILLIAMS, R. D. AND SILOW, R. A. 1933. Genetics of red clover (*Trifolium pratense* L.) compatibility. I. *J. Genet.*, **27**, 341–361.
- 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.
- XU, B., MU, J., NEVINS, D., GRUN, P. AND KAO, T.-H. 1990. Cloning and sequencing of cDNAs encoding two self-incompatibility associated proteins in *Solanum chacoense*. *Sex. Plant Reprod.*, **3**, 54–60.