

# Estimating the number, frequency, and dominance of *S*-alleles in a natural population of *Arabidopsis lyrata* (Brassicaceae) with sporophytic control of self-incompatibility

BK Mable<sup>1</sup>, MH Schierup<sup>2</sup> and D Charlesworth<sup>1</sup>

<sup>1</sup>Institute of Cell, Animal and Population Biology, University of Edinburgh, Ashworth Laboratories, King's Buildings, Edinburgh EH9 3JT, UK; <sup>2</sup>Department of Ecology and Genetics, University of Aarhus, Ny Munkegade, Bygn. 540, DK-8000 Aarhus C, Denmark

In homomorphic plant self-incompatibility (SI) systems, large numbers of alleles may be maintained at a single Mendelian locus. Most estimators of the number of alleles present in natural populations are designed for gametophytic self-incompatibility systems (GSI) in which the recognition phenotype of the pollen is determined by its own haploid genotype. In sporophytic systems (SSI), the recognition phenotype of the pollen is determined by the diploid genotype of its parent, and dominance differs among alleles. We describe research aimed at estimates of *S*-allele numbers in a natural population of *Arabidopsis lyrata* (Brassicaceae), whose SSI system has recently been described. Using a combination of pollination studies and PCR-based identification of alleles at a locus equivalent to the *Brassica SRK* gene, we identified and sequenced 11

putative alleles in a sample of 20 individuals from different maternal seed sets. The pollination results indicate that we have not amplified all alleles that must be present. Extensive partial incompatibility, nonreciprocal compatibility differences, and evidence of weakened expression of SI in some genotypes, prevent us from determining the exact number of missing alleles based only on cross-pollination data. Although we show that none of the theoretical models currently proposed is completely appropriate for estimating the number of alleles in this system, we estimate that there are between 13 and 16 different *S*-alleles in our sample, probably between 16 and 25 alleles in the population, and discuss the relative frequency of alleles in relation to dominance.

*Heredity* (2003) 90, 422–431. doi:10.1038/sj.hdy.6800261

**Keywords:** sporophytic self-incompatibility; Brassicaceae; *Arabidopsis lyrata*; *S*-alleles; partial incompatibility

## Introduction

The self-incompatibility (SI) recognition sequences of plants maintain extremely large number of alleles for long periods of time because of frequency-dependent selection. Opportunities for mating of each *S*-allele are inversely related to the alleles' frequencies, which maintains many alleles within populations (Wright, 1939). *S*-allele polymorphism was first described in detail for tobacco by East and Yarnell (1929), but it was the seminal study of Emerson (1938, 1939) on *Oenothera organensis* that captured the attention of Wright (1939), followed by a series of other theoretical (eg Fisher, 1947, 1958; Wright, 1960, 1964; Yokoyama and Nei, 1979), and more empirical (Emerson, 1938, 1939; Bateman, 1947) population geneticists (see review by Lawrence, 2000). The early theoretical studies analyzed the effects of mutation rate, effective population size, and migration rates in determining the equilibrium number of alleles and later work used these results to develop approaches for estimating allele numbers in populations from those found in samples.

In gametophytic SI (GSI; as found in *O. organensis*), the recognition phenotype of the pollen is determined by expression of its own haploid genotype so that, in diploids, each pollen grain expresses only a single allele, and all individuals are heterozygous at the *S*-locus. In sporophytic SI (SSI), the phenotype of the pollen is determined by the diploid genotype of the male parent plant (Kusaba *et al.*, 2002; Shiba *et al.*, 2002).

For GSI, estimation of the number of alleles is possible assuming that (1) essentially all individuals are heterozygotes at the *S*-locus, (2) all alleles are at equal frequencies in the population, and (3) the alleles are associated randomly so that the probability of any pair is  $1/n(n-1)$ , where  $n$  is the number of alleles. However, with large numbers of alleles ( $n > 30$ ), there are more than 870 possible genotypes and it is unlikely to observe any one genotype in a limited sample. For GSI, most empirical studies (reviewed by Lawrence, 2000) have estimated the number of alleles via cross-pollination studies, but recently *S*-alleles have been identified by RT-PCR, assuming that different sequences represent functionally different allelic specificities (Richman *et al.*, 1995, 1996).

In SSI, the problem of estimating allele numbers from cross-pollination data is more difficult. Diploid expression of recognition proteins in the sporophytic anther

tissue gives the potential for dominance interactions among alleles. In the Brassicaceae (the most intensively studied family with SSI), dominance relationships are often nonlinear, and may differ in the pollen and stigma (Bateman, 1954; Sampson, 1964; Thompson and Taylor, 1966; Hatakayama *et al*, 1998). This invalidates the simple assumptions that can be made for GSI estimators. Since recessive alleles can be effectively 'hidden', they may be more frequent in populations than dominant alleles, as predicted by theoretical studies (Bateman, 1952; Sampson, 1964) and supported by empirical studies (Thompson and Taylor, 1966; Ockendon, 1974; Stevens and Kay, 1989; Kowiyama *et al*, 1994). Homozygotes for recessive alleles are also possible. Simulation studies have found that the numbers of alleles maintained for SSI systems depend both on the dominance relations and on the frequencies of different dominance classes in populations (Charlesworth, 1988; Schierup *et al*, 1997; Schierup, 1998; Schierup *et al*, 1998; Vekemans *et al*, 1998). However, the general prediction that recessive alleles should have higher frequencies than dominant ones is valid only for undivided populations (Schierup *et al*, 1997, 2000).

In cultivated species of the genus *Brassica*, more than 30 S-alleles have now been identified in *B. campestris* (and more than 50 in *B. oleracea*) by a combination of cross-pollination studies (eg, Wallace, 1979) and molecular characterization of S-alleles using PCR-based screens (Nou *et al*, 1993a; Brace *et al*, 1994; Kusaba *et al*, 1997; Nishio and Kusaba, 2000) and large-scale cloning (Boyes and Nasrallah, 1993; Nasrallah *et al*, 2000). In these species, S-alleles can be divided into two classes based on their dominance (class I) or recessivity (class II) in pollen (Ockendon, 1974; Hatakayama *et al*, 1998; Kusaba *et al*, 2001; Shiba *et al*, 2002). It is not, however, clear how cultivation history may have affected the frequencies of S-alleles of different dominance, and information from natural populations is required. A few studies have attempted to estimate numbers of alleles and dominance relationships in naturally occurring species with SSI (*Raphanus raphanistrum*: Sampson, 1967; *Sinapis arvensis*: Ford and Kay, 1985; *Raphanus sativus*: Karron *et al*, 1990; *Ipomoea trifida*: Kowiyama *et al*, 1994), but the genes controlling SI have not yet been sequenced for most of these species, which are not very closely related to *Brassica* (although all except *Ipomoea* are in the Brassicaceae), so that the results cannot be coordinated with sequence data.

Molecular characterization of the number of alleles is also complicated in SSI, because of high polymorphism and because the loci belong to extensive gene families (Schierup *et al*, 2001). In *Brassica*, the S-locus region contains at least three genes that have been implicated in SI reactions (see review by Nasrallah, 2000). The two stigma-expressed genes, S-locus glycoprotein (SLG) and S-related kinase (SRK), share an extracellular domain (the S-domain) but SRK is found to be sufficient for determination of specificity (Takasaki *et al*, 2000). The pollen-expressed gene S-locus cysteine rich (SCR or SP11) is located in the same region, but is a member of a distinct gene family (Schopfer *et al*, 1999; Takayama *et al*, 2000). The extremely high polymorphism (eg, Stein *et al*, 1996; Nishio and Kusaba, 2000; Schierup *et al*, 2001) means that PCR primers will not amplify all alleles, making allele discovery and identification difficult. The presence of paralogous loci in the genome adds the

potential of misclassifying sequences as S-alleles when PCR-based methods are used, so that establishing linkage to the locus controlling SI phenotypes (the S-locus) is essential (Schierup *et al*, 2001).

We have studied SSI in natural populations of *Arabidopsis lyrata* (Brassicaceae) from North America (formerly *Arabis lyrata*), Iceland, and Scotland (formerly *Arabis petraea*). Combining classical pollination and PCR-based approaches (described in Charlesworth *et al*, 2000) we have identified a locus, *Aly13*, with homology to *Brassica* SRK alleles that is linked to SI. Kusaba *et al* (2001) independently cloned a stigma-expressed cDNA (which they called *AIISK*) from a plant from one of the same North-American populations used by us, and identified two alleles (*Sa* and *Sb*), both of which were identified independently in our study (our *Aly13-13* and *Aly13-20*, respectively). Kusaba *et al* (2002) recently showed that these alleles were codominant in the stigma (with weakening of *Sa*), whereas *Sa* is recessive to *Sb* in the pollen. This difference appears to be related to the expression of the *SCR* gene in the pollen: the pollen-dominant allele (*Sb*) is expressed both sporophytically (ie in the tapetal cells) and gametophytically (ie in the pollen), whereas expression of the pollen-recessive allele is exclusively sporophytic, and at a lower level than the dominant allele. In a previous study (Schierup *et al*, 2001), we identified 22 putative SRK alleles from North America and Iceland and, using several independent families, we demonstrated the linkage of 11 of these to the S-locus. Four of these 11 alleles were codominant with one another (including *Aly13-20*), two appeared to be recessive to the others (in both pollen and stigma), and one appeared to have intermediate dominance (*Aly13-13*).

Using the same approach used to identify this locus, we set out to characterize the allele number and dominance relationships in a large natural population from Iceland. Here we describe the *Aly13* alleles identified in this population, comparing the results with those from other populations, and discuss our current knowledge about dominance relationships among these alleles.

## Methods

### Plants and S-phenotype determinations

Plants were raised from seeds collected from Esja mountain about 20 km north of Reykjavik, on a steep southfacing scree slope, at an altitude of approximately 100 m. Seeds were planted in Fisons F2 compost with the addition of 30% 2–3 mm grit and raised in a greenhouse under a minimum of 12 h light (with artificial light when necessary). Up to seven seedlings were raised from each of 50 maternal seed groups. As plants began to flower, they were placed under bridal veil covers for a minimum of 3 days prior to pollination.

Pollinations were performed by rubbing anthers on stigmas of intact flowers. Each combination of parents was pollinated reciprocally, using three replicate flowers. Compatibility was scored as fruit set 7–10 days after pollination. Rather than scoring pollinations by fluorescent staining of pollen tube growth, fruit set was monitored in order to produce second-generation progeny to test dominance relationships. Where the results

were unclear, additional pollinations were done. Each individual used was also self-pollinated to determine the strength of SI. A total of 30 plants were used in crosses, but not all flowered sufficiently to be crossed to a substantial number of other plants. In total, plants from 20 maternal parents were reciprocally crossed to one another (Table 1).

For each pollination, seed set was scored as negative, positive, or 'small'. 'Negative' fruits contained no seeds and positive fruits contained at least three seeds. Statistics on all fruits collected (fruit length, number of seeds, number of seeds per fruit length) indicated the existence of an additional discrete class, which we term 'small'. These fruits were smaller than average when mature (fruit : seed length ratio of 0.2 or less) and had one to three seeds (complete details of these analyses can be obtained from the first author). Pollinations were conducted over a 5-month period (February–July) because these plants produce few flowers at any time. Fruits were collected when mature (4–7 weeks).

### S-genotype determinations

Leaves were collected from each of the plants used in the crosses. Three to four leaves were frozen at  $-80^{\circ}\text{C}$  and DNA extracted using a CTAB method (Junghans and Metzloff, 1990). Genotypes at the *Aly13* locus (the candidate *SRK* gene for *A. lyrata*) were scored as in Schierup *et al* (2001). In brief, primers designed in conserved regions of the *S*-domain and known to amplify a number of the *Aly13* alleles (13F1+SLGR) were first used to amplify DNA extracted from each individual (Table 2). Strong amplification products of the expected size (approximately 1 kb) were cloned into TA cloning vectors (Invitrogen PCR TOPO 4.1) and three variants of each type detected by restriction digestions were sequenced using the universal primers M13F and M13R. Sequences were aligned to previously characterized 'alleles' to identify allelic types. To identify further sequence types not amplified by the general primers, we used the general reverse primer SLGR (which amplifies all alleles that we have so far found) with allele-specific forward primers designed to amplify each known *Aly13* sequence type (see Table 2 and Schierup *et al*, 2001). For each product identified in each individual in this screen, direct sequences were obtained, using as sequencing primers the specific primer with SLGR. Expected fragment sizes ranged from 474 to 1000 bp (Table 2) depending on the location of the forward primer in the *S*-domain. For each set of sequencing reactions, chromatographs were aligned and base-calling errors corrected using SeqEd, and then pasted into a manual alignment program (Se-al 1.0, Rambaut, 1996) for comparison with previously identified *Aly13* alleles (Schierup *et al*, 2001). Alignments were further corrected by translating to amino-acid sequences (using MacClade 4.0, Maddison and Maddison, 2000). For cloned products, consensus sequences were obtained by correcting PCR mutations to the nucleotide found in the majority of clones for a given individual. Putative alleles whose sequences did not match any of the previously identified alleles were given new allele numbers. Synonymous and nonsynonymous divergence between alleles (ie the number of synonymous changes per synonymous site,  $K_s$ , and of nonsynonymous changes per nonsynonymous site,  $K_a$ )

were computed according to Nei and Gojobori (1986) using DNAsp 3.53 (Rojas and Rojas, 1999).

## Results

### S-phenotype determinations

Table 1 summarizes the results of the crossing experiments for the 20 plants from individual maternal seed groups that flowered sufficiently to be crossed to the majority of the other plants used. The individuals in Table 1 are grouped primarily according to phenotypic assessment of compatibility relationships but the molecular genotyping results are also used. We first consider the phenotypic data. Eight individuals produced some fruits when self-pollinated (99R 18/1, 19/2, 25/1, 37/4, 36/3, 32/1, 98I 36/1). Individual 99R 25/1 produced three full-sized fruits from six self-pollinations, and overall 9/13 (69%) of the fruits produced were classified as 'small', suggesting weak expression of SI (see below). Out of the 380 cross-pollination combinations possible, 355 were performed (10/190 possible reciprocal crosses could not be performed; an additional five crosses were completed only in one direction). Of the 175 complete reciprocal crosses performed, 17% showed clear non reciprocal incompatibility. Out of the 355 crosses, 40 appeared to be fully incompatible and produced no fruits, 45 appeared to be predominantly incompatible but some fruits were produced (37/60 or 62% of them classified as 'small'), 156 were fully compatible, yielding only full-sized fruits, 97 appeared to be predominantly compatible but produced some small fruits, and finally, 17 could not be reliably categorized.

### S-genotype determinations

The *Aly13* sequence types determined for each individual from the Icelandic population are shown in Table 1, along with the crossing data. In all, 11 sequence types were found and sequenced for all individuals in which they were found. For 10 of them (*Aly13*-1, *Aly13*-6, *Aly13*-9, *Aly13*-11, *Aly13*-12, *Aly13*-14, *Aly13*-15, *Aly13*-16, *Aly13*-22, and *Aly13*-25), we have evidence of linkage to the *S*-locus, the first four reported in Schierup *et al* (2001), and the remainder based on families from other Icelandic populations (J Bechsgaard, LH Nielsen and MH Schierup, unpublished results; BK Mable and J Beland, unpublished results). One sequence (*Aly13*-7) was not closely linked to the other *Aly13* alleles, based on segregation analysis in three independent families studied. In Table 1, individuals with *Aly13*-7 are indicated with asterisks, but these sequences are not considered further. If most individuals are heterozygous at the *S*-locus, it is clear from Table 1 that we have not identified all the alleles present. Individuals 99R 10/2, 18/1, 2/2, 11/3, and 13/3 show reciprocal incompatibility despite the fact that their one identified allele is different in each plant. These plants must therefore share an as yet unidentified allele (indicated by 'X'). Individuals 99R 1/2 and 98I 36/1 share *Aly13*-11 and show nonreciprocal incompatibility to one another, but may not share their other allele, as they differ in their compatibility reactions with other individuals. Individuals 99R 5/4 and 2/2 also have unidentified alleles, and may or may not share alleles with other individuals. Based on this information, we can estimate a minimum

**Table 1** Results of cross-pollinations among 20 individual plants from a population in Iceland (recipient plants are listed in columns against donor plants in rows) indicating individual number (all plant names are prefixed with 99R except where indicated), genotype determinations at the *Aly13* locus, overall pollination results (*shaded cells* indicate predominantly incompatible cross-pollinations, *cells on the diagonal* indicate self-pollinations, and open cells indicate predominantly compatible cross-pollinations) and a detailed ranking of the 'strength' of compatibility

<i>Aly13</i> Allele	Plant	13-1/X	13-1/X	X/?	13-1/X*	13-6/X*	13-1/6	13-1/9	13-1/9*	13-1/?	13-9/22	13-1/22*	13-12/?*	13-12/16	13-1/14*	13-1/25	13-1/25*	13-1/25	13-11/?	13-11/?	13-1/15*
		10/2	18/1	2/2	13/3	11/3	19/2	25/1	37/4	5/4	35/5	36/3	44/4	14/1	9/5	32/1	28/1	43/1	1/2	981 36/1	20/2
13-1/X	10/2	1	2	2	2	2	5	5	5	6	4	6	5	6	5	5	4	5	5	5	6
13-1/X	18/1	1	2	1	1	2	4	6	6	6	6	5	6	6	6	6	6	–	6	2	6
X/?	2/2	2	1	1	1	1	5	5	5	5	5	3	6	5	6	5	1	5	5	5	5
13-1/X*	13/3	2	1	1	1	1	5	6	6	6	–	–	1	6	6	2	6	–	6	2	–
13-6/X*	11/3	2	1	2	1	1	6	6	5	6	3	4	5	6	2	5	5	6	6	6	6
13-1/6	19/2	6	6	4	4	6	2	5	3	6	6	2	6	5	3	6	4	6	6	5	6
13-1/9	25/1	6	5	5	5	6	5	3	6	6	2	1	6	5	6	6	6	–	6	6	2
13-1/9*	37/4	5	5	6	–	6	5	1	2	1	2	6	5	6	5	5	6	–	5	6	–
13-1/?	5/4	6	4	1	2	6	6	6	1	1	6	6	5	5	3	6	3	1	6	5	1
13-9/22	35/5	4	5	6	–	4	5	6	6	6	1	1	5	6	6	5	6	–	6	6	6
13-1/22*	36/3	6	2	6	–	6	2	6	6	6	–	2	–	6	6	6	6	6	6	5	–
13-12/?*	44/4	5	2	6	1	6	5	6	2	6	6	–	1	1	5	5	6	6	5	6	1
13-12/16	14/1	6	6	6	2	6	5	2	5	5	5	4	5	1	6	5	6	6	5	3	6
13-1/14*	9/5	5	5	6	6	3	5	6	5	5	6	1	6	5	1	2	4	6	6	6	6
13-1/25	32/1	6	5	5	6	3	5	6	4	6	6	–	5	3	6	2	1	1	6	6	1
13-1/25*	28/1	5	6	4	4	6	6	4	5	5	5	1	5	5	5	1	1	1	6	6	3
13-1/25	43/1	6	–	6	6	2	5	–	–	2	–	6	2	3	3	1	1	1	6	6	–
13-11/?	1/2	6	6	6	6	6	6	6	6	6	6	5	1	5	5	6	1	6	1	2	1
13-11/?	981 36/1	1	5	6	1	2	3	5	6	5	5	6	5	6	5	5	2	6	6	2	5
13-1/15*	20/2	4	6	6	–	6	5	6	5	1	5	–	5	6	6	5	5	–	1	5	1

\*Plant also has an unlinked allele, *Aly13-7*. 1=completely incompatible; 2=predominantly incompatible but with some small fruits; 3=predominantly incompatible but with some full-sized fruits; 4=predominantly compatible but with some ambiguity; 5=predominantly compatible but with some small fruits; 6=completely compatible.

**Table 2** List of allele-specific primers used to screen S-domain genotypes at the *Aly13* locus used in combination with the general reverse primer, SLGR, which is complementary to near the end of the S-domain

Primer name	Sequence (5'–3')	Priming site <sup>a</sup>	Product <sup>b</sup>
13–1F1	ttcaacaacaactcacaagaa	446	13–1
13–2F1	agactcctcaaaaccaacgggagt	444	13–2
13–3F2	gtataaactcgaaatgctag	593	13–3
13–4F2	gatgaaactgggtgggatgagc	521	13–4
13–5F1	ctccaaaactaatgacttagacggg	446	13–5
13–6f1	cttaccgagttgtaactctac	955	13–6
13–7f1	gtacttccaagaaaatttg	810	13–7
13–8f1	cagagatgcaaagatggaaac	715	13–8
13–9f1	gaggtcacttacacgttctg	710	13–9
13–10f1	gtataattttacggagaacggc	774	13–10
13–11F1	atgccagaggcggcacaatcaa	710	13–11
13–12F1	ccagatatgacgacaatggacc	765	13–12
13–13F1	tggagatgcaaaaatgggacg	732	13–13, 13–16, 13–25
13–14F1	cagagatgcaaagatggaaag	720	13–14
13–15F1	caaagagggaagtgccttcat	800	13–15
13–16F1	ttggtttgttaccatcttg	670	13–16
13–17F1	ccacagcttactctagactc	710	13–17
13–18F1	aaagaggagatcgcttacc	710	13–18
13–19F1	gagttttacctgagttttctac	540	13–19
13–20F1	aacaatgggaatacatggtctc	540	13–20
13–22F1	acacgttcttaatgaccaacaag	825	13–22
13–23F1	aacaacagagactcgccaactc	671	13–23
13–24F1	cagagatacaagatctccaatg	379	13–15
13–25F1	agcgtagggaattagagcgc	856	13–25
SLGR <sup>c</sup>	atctgacataaagatcttgacc	1329	general

<sup>a</sup>Priming sites are given in bp from the 5' end of the S-domain based on an alignment of the sequence types found in the Icelandic population with the complete S-domain sequences of *Aly13*–13 and *Aly13*–20 (referred to as *AlskA* and *AlskB* in Kusaba *et al*, 2001).

<sup>b</sup>Some allele-specific primers amplified more than one product; only alleles amplified in the Icelandic population are shown.

<sup>c</sup>All forward primers were used with the general reverse primer SLGR, which begins 60 bp from the 3' end of the S-domain.

**Table 3** Allele frequencies and genotype combinations of *Aly13* sequence types found in an Icelandic population

Frequency	Allele	13–1	13–6	13–9	13–11	13–12	13–14	13–15	13–16	13–22	13–25	X <sup>a</sup>	? <sup>b</sup>
0.33	13–1	—	1	2	0	0	1	1	0	1	3	3	1
0.048	13–6	1	—	0	0	0	0	0	0	0	0	1	0
0.071	13–9	2	0	—	0	0	0	0	0	1	0	0	0
0.048	13–11	0	0	0	—	0	0	0	0	0	0	0	2
0.048	13–12	0	0	0	0	—	0	0	1	0	0	0	1
0.024	13–14	1	0	0	0	0	—	0	0	0	0	0	0
0.024	13–15	1	0	0	0	0	0	—	0	0	0	0	0
0.024	13–16	0	0	0	0	1	0	0	—	0	0	0	0
0.048	13–22	1	0	1	0	0	0	0	0	—	0	0	0
0.071	13–25	3	0	0	0	0	0	0	0	0	—	0	0
0.119	X <sup>a</sup>	3	1	0	0	0	0	0	0	0	0	—	1
	? <sup>b</sup>	1	0	0	2	1	0	0	0	0	0	1	—

<sup>a</sup>Unidentified allele predicted to be shared by a number of individuals.

<sup>b</sup>Unidentified alleles for which it is less certain which are shared; the two unknown alleles with *Aly13*–11 are predicted to be different.

of 13 and a maximum of 16 alleles in the sample of 20 individuals.

#### Dominance determinations

The unidentified allele (13-X) shared by individuals 99R 10/2, 18/1, 2/2, 13/3, and 11/3 is expressed (depending on the individual) along with 13–1, 13–6, and another unidentified allele, and appears to be dominant to 13–6 in both pollen and stigma (based on crosses between 99R 11/3 and 19/2). Allele 13–1 is

found in many individuals, but does not appear to be completely dominant to any of the alleles with which it is found, although it may be partially expressed in stigmas in individuals with alleles 13-X, 13–25 (see the cross between 99R 13/3 and 32/1), 13–22 (based on crosses involving 99R 36/3), and 13–15 (based on crosses involving 99R 20/2) in some cases. Alleles 13–9 and 13–22 appear to be codominant with one another in stigmas, but 13–9 appears to be recessive to 13–22 in pollen (see the crosses between

**Table 4** Summary of *Aly13* alleles found in populations of *A. lyrata* from Iceland, North America (NA), Scotland, Japan, and Austria

Population	Iceland <sup>a</sup>	NA	Scotland	Japan	Austria
13-1 <sup>b</sup>	Yes	Yes	Yes	Yes	Yes
13-3	No	Yes	No	No	No
13-4	Yes	No	Yes	No	No
13-5	Yes	No	Yes	No	No
13-6	Yes	No	No	No	No
13-7 <sup>c</sup>	Yes	No	No	Yes	Yes
13-8	No	No	Yes	No	No
13-9	Yes	No	No	No	No
13-10	No	No	Yes	No	No
13-11	Yes	No	No	No	No
13-12	Yes	No	No	No	No
13-13	No	Yes	No	No	No
13-14	Yes	No	No	No	No
13-15	Yes	No	No	No	No
13-16	Yes	No	No	No	No
13-17	No	No	Yes	No	No
13-18	Yes	No	No	No	No
13-19	No	Yes	No	No	No
13-20	No	Yes	No	No	No
13-22	Yes	No	No	No	Yes
13-25	Yes	No	No	No	No
13-27	Yes	No	No	No	No
13-28	No	No	No	Yes	No
13-29	No	No	No	Yes	No
13-30	No	No	No	Yes	No

<sup>a</sup>Populations from Iceland were the only ones surveyed systematically; presence or absence of alleles from other populations is based on a more limited number of independent individuals sampled opportunistically while trying to identify the *S*-locus in *A. lyrata* (roughly 10 each from the North American and Scottish populations and four each from the Japanese and Austrian populations).

<sup>b</sup>Note that *Aly13*-1 has been found in all regions sampled.

<sup>c</sup>*Aly13*-7 is not linked to self-incompatibility, but shows high sequence similarity to the linked alleles.

**Table 5** Pairwise divergence values showing synonymous (Ks: above diagonal) and nonsynonymous (Ka: below diagonal) variation

	<i>Sa</i> <sup>a</sup>	<i>Sb</i>	13-1	13-6	13-7	13-9	13-11	13-12	13-14	13-15	13-16	13-22	13-25
<i>Sa</i>	—	1.078	0.865	1.116	0.881	0.859	0.904	1.033	1.142	0.952	0.389	0.816	0.557
<i>Sb</i>	0.315	—	0.962	1.392	0.937	0.701	0.814	0.890	1.757	0.592	1.042	0.842	0.953
13-1	0.300	0.254	—	1.869	0.898	0.501	0.565	1.124	1.585	0.624	0.900	0.819	0.821
13-6	0.331	0.361	0.399	—	0.603	1.214	0.686	1.215	0.415	1.559	1.051	0.730	1.356
13-7	0.371	0.386	0.348	0.313	—	0.805	0.619	0.761	0.659	0.816	0.897	0.668	0.946
13-9	0.334	0.291	0.285	0.368	0.392	—	0.700	0.936	1.043	0.491	0.704	0.673	0.707
13-11	0.340	0.291	0.323	0.338	0.399	0.248	—	1.029	0.684	0.630	0.832	0.505	0.885
13-12	0.289	0.280	0.271	0.287	0.356	0.372	0.320	—	1.461	1.370	1.055	0.669	1.277
13-14	0.348	0.335	0.387	0.166	0.328	0.361	0.349	0.308	—	1.030	1.582	0.789	1.206
13-15	0.315	0.250	0.204	0.318	0.344	0.217	0.262	0.298	0.314	—	0.720	0.647	0.768
13-16	0.281	0.310	0.332	0.360	0.323	0.328	0.362	0.299	0.351	0.322	—	0.737	0.268
13-22	0.321	0.307	0.239	0.380	0.389	0.248	0.242	0.324	0.380	0.232	0.334	—	0.775
13-25	0.280	0.315	0.328	0.356	0.360	0.333	0.309	0.286	0.366	0.286	0.163	0.361	—

<sup>a</sup>Two alleles from North-American plants that were also used in studies by Kusaba *et al* (2000, 2002) are included for comparison. *AlskA* and *AlskB* (Kusaba *et al*, 2000) or *Sa* and *Sb* (Kusaba *et al*, 2002) correspond with our *Aly13*-13, and *Aly13*-20, respectively.

99R 35/5 and 36/3). Similarly, allele 13-12 appears to be expressed with 13-16 (it could be codominant or dominant) in stigmas but shows weakened expression in pollen (see the cross between 99R 44/4 and 14/1) which could be due to interactions with 13-16 or to the unidentified allele in 99R 44/4. The three individuals with the genotype 13-1/13-25 show correlated but not identical compatibility patterns. Environmental or genotype specific effects may be responsible for these discrepancies (ie context may change expression patterns).

#### Relative frequency of *Aly13* alleles

Table 3 summarizes the allele frequencies and allelic combinations found in the 20 individuals tested, assuming that the *S*-domain sequence types identified are alleles (except for *Aly13*-7, which we know to be unlinked, see above). Allele frequencies were far from equal, with more than twice as many of *Aly13*-1 as the next most common allele. The unlinked sequence *Aly13*-7 was also found often, but not in every individual and shows no discernable effect on the compatibility patterns or the strength of the incompatibility response.

## Divergence among alleles

Of the 11 sequences found in this population, 10 have been described previously (Schierup *et al*, 2001) and one (*Aly13-25*) is a new sequence type (Table 4). Comparison of the sequences of alleles from all individuals indicated that, except for *Aly13-1* (for which there were three polymorphic sites), sequences of the same *Aly13* allele were identical between different individuals, in contrast with high divergence between alleles (shown in Table 5).

## Discussion

### Estimating the number of S-alleles

The high degree of partial and nonreciprocal incompatibility found in crosses between the 20 individuals studied here makes it difficult to classify individuals into clear incompatibility groups to estimate allele numbers based on pollination data alone. This contrasts with the within-family comparisons we have used previously, which showed some partial incompatibility but few nonreciprocal differences (Schierup *et al*, 2001). Most of the methods that have been proposed for estimating the number of S-alleles in sporophytic systems rely on testing relationships using bud pollinations to obtain homozygous test strains (eg Ford and Kay, 1985; Stevens and Kay, 1989) but inbreeding depression makes it very difficult to produce viable homozygotes in *A. lyrata* for the majority of alleles (MH Schierup and J Bechsgaard, unpublished results). As a possible alternative, Tajima *et al* (1994) suggested statistical methods for estimating effective numbers of alleles and expected heterozygote frequencies (for both gametophytic and sporophytic SI systems) from the proportion of incompatible pollen estimated by cross-pollinations, based on a simple model. However, the approach relies on detailed knowledge of the frequency of different dominance classes, which can be very complex, and thus difficult to determine accurately in practice. For example, using bud pollinations in *Sinapis arvensis*, Ford and Kay (1985) and Stevens and Kay (1989) found 28 alleles in a sample of 35 plants and observed codominance in both pollen and stigma in 43% of heterozygotes, dominance of one allele in the pollen with codominance in the stigma in 50%, dominance of one allele in the stigma with codominance in pollen in 5%, and dominance of one allele in both pollen and stigma in only 2%. Somewhat similarly, in wild populations of *Brassica campestris*, Hatakeyama *et al* (1998) found that codominance was common in the stigma but rarer in pollen, and that many alleles differed in dominance in pollen and stigmas. We do not yet have a complete picture of dominance relationships among alleles in *A. lyrata*, but information from crosses using North American populations indicates that several alleles are dominant in both pollen and stigma in the five families studied (Schierup *et al*, 2001), suggesting that dominance classes cannot be assumed to be the same as in other systems. Thus, we are unable to use the currently available methods to estimate allele numbers based only on pollination data. Nevertheless, our estimate of sample allele numbers based on combined pollination and genotype data (ie 13 to 16 alleles in a sample of 20 individuals) is similar to those found in other systems. For example, Karron *et al* (1990) used a similar approach to ours and estimated 32 alleles in a sample of 50 individuals from two populations.

The next step is to use the number of alleles in the sample to predict the total number of alleles in the population. For our data, we do not know if any of the individuals are homozygotes (we have repeatedly attempted to test progeny to assess this, but survivorship has been too low), but clearly both allele and genotype frequencies are highly unequal (see Table 3). Thus, none of the estimators for GSI (Bateman, 1947; Fisher, 1947; Paxman, 1963; O'Donnell and Lawrence, 1984) is satisfactory, especially as the theoretical predictions for numbers and dynamics of alleles in sporophytic systems depend strongly on the dominance model assumed (Charlesworth, 1988; Schierup *et al*, 1997; Schierup, 1998; Vekemans *et al*, 1998). For example, if  $n$  is the number of different alleles found in a sample of  $m$  copies of the S-gene examined, using our sample of  $m=40$  alleles and our minimum ( $n=13$ ) and maximum estimates ( $n=16$ ) of the number of alleles, Paxman's (1963) maximum likelihood estimator gives estimates of population allele numbers as  $N=13.0$  and  $16.1$ , respectively. This would suggest that we have sampled all of the alleles in the population, which is very unlikely. O'Donnell and Lawrence's (1984) estimators designed to allow for uneven allele frequencies in GSI systems give slightly higher estimates (E1 yields  $N=15.5$  and  $20.5$ , and E2 yields  $N=16.8$  and  $24.3$ ), but it is unclear how dominance in SSI systems will affect these estimators. As an approximate measure of how many more alleles might be expected in the population (ie the thoroughness of the analysis), Stevens and Kay (1989) modified the repeatability index ( $R$ ) of Campbell and Lawrence (1981) to deal with sporophytic systems, by correcting for the difference in the minimum numbers of alleles necessary in the two kinds of systems, with  $R$  estimated as  $1-[(n-2)/(m-2)]$ . Our minimum and maximum estimates of  $n$  yield estimated repeatability indices of  $R=0.71$  or  $0.63$ , respectively, so that the total number of alleles in the population should be between 18 and 25. We thus conclude that there are probably between 16 and 25 alleles in the population. We have currently identified 28 alleles from populations sampled in North America, Scotland, Austria, and Iceland (Table 4) and have predicted an additional three to six new alleles here. This is similar to population allele numbers found in other SSI systems.

### Partial compatibility

Partial self-compatibility (indicated by the production of both small and full-sized fruits) was evident in self-pollinations of several individuals, particularly those with the genotype 13-1/13-9. Partial self-compatibility was also found in individuals with the genotypes 13-1/13-X, 13-1/13-6, 13-1/13-25, and 13-1/13-22, and in one of the two individuals with allele 13-11, which may suggest weakening of SI in these alleles or allele combinations. Stevens and Kay (1989) reported weakening of S-allele expression in some crosses but attributed it to recessive alleles segregating at a modifier locus. 'Leaky' expression of self-incompatibility has also been reported in wild *Brassica campestris* populations (Nou *et al*, 1993b; Hatakeyama *et al*, 1998). In cultivated *Brassica*, individuals with two recessive alleles tend to show at least partial self-compatibility, whereas plants with one recessive and one dominant allele are usually

highly SI but effectively cross-compatible with any plant not carrying the dominant allele (Ockendon, 1974). Partial compatibility has implications both for selection of dominant alleles that do not allow 'leakage' (Ockendon, 1974) and for allowing small populations to survive in the face of limited mate availability. We are currently investigating whether the noted weakening of SI occurs predominantly among recessive alleles in *A. lyrata*.

#### Dominance determinations

Although partial incompatibility makes interpretation difficult in some cases, we can make preliminary inferences about dominance based on our intrapopulation pollinations. Alleles 13-X and 13-22 appear to be dominant (Group I), 13-12 and 13-9 (Group II) are recessive in pollen but codominant in stigmas with Group I, and 13-6 and 13-1 (Group III) are recessive to Group I in both pollen and stigmas (although 13-1 may show partial expression with some other alleles in pollen in some cases). We previously suggested (Schierup *et al*, 2001) three dominance classes in *A. lyrata*: dominant alleles (including *Aly13-20* and *Aly13-22*), an intermediate class (represented by *Aly13-13*, which is not found in the Icelandic population), and recessive alleles (including *Aly13-1*). More recent evidence suggests at least four dominance classes, 13-1 being the only allele found (so far) in the most recessive class (J Bechsgaard, LH Nielsen, and MH Schierup, unpublished results). We are currently testing these dominance predictions using intrafamily crosses from progeny raised from the plants studied here.

#### Relative frequencies of alleles

Bateman (1952) predicted that recessive alleles should in general be at higher frequencies than dominant ones, but recent models suggest that this relationship may be more complex. Sampson (1974) argued that alleles with the highest frequency will belong to the dominance level with the fewest number of alleles, regardless of the dominance order. Using simulations of an evolving hierarchy of dominance levels, Schierup *et al* (1997, 1998, 2000) found that incorporating more complicated demographic parameters (mate availability, fecundity selection, and population subdivision) resulted in varying predictions. For example, if dominance occurs only in pistils, turnover of alleles favors increasing dominance over evolutionary time rather than recessive alleles. The recessive alleles would then be predicted to be a mixture of old alleles that were previously dominant and new alleles having arisen as recessive, both with a short expected lifetime in the population. Uyenoyama (2000) suggested that at equilibrium in a finite undivided population all alleles have the same expected transmission rate and that, with just two dominance classes (as in *Brassica*), there should be multiple dominant alleles at low frequencies, and a single high-frequency recessive allele. This latter prediction may be supported by the data presented here if *Aly13-1* is indeed the most recessive allele and is consistent with the findings of Kowiyama *et al* (1994), who showed that the most recessive allele in a six-level dominance hierarchy ( $S_3$ ) in *Ipomoea trifida* was the only one present in all six populations studied.

All these theoretical models assume that alleles may mutate to other alleles with different specificities and any dominance level. This is unlikely to be true in *Brassica*, where recent evidence shows that dominance is controlled during mRNA expression, and that the sequences of both *SRK* and *SCR* genes of recessive alleles are much more closely related than those of dominant alleles (Watanabe *et al*, 2000; Shiba *et al*, 2002), suggesting that mutations between recessive and dominant alleles can rarely occur. In *A. lyrata* there appears to be at least three dominance classes and no clear correlation between dominance level and sequence divergence in the S-domain (Table 5), suggesting shifts between dominance classes may occur more often than in *Brassica*.

To test predictions of the various models, we need to know dominance relationships of alleles in natural populations. It is difficult to obtain this information, but some data exist. Among *Brassica* cultivars, recessive alleles generally occur at higher frequencies than dominant alleles (Thompson and Taylor, 1966; Ockendon, 1974) and the same is true in natural populations of *R. raphanistrum* (Sampson, 1964), *I. trifida* (Kowiyama *et al*, 1994) and *S. arvensis* (Stevens and Kay, 1989). In our Icelandic study population, the *Aly13-1* allele is not only found at a much higher frequency than other alleles (Table 3) but it is also the most widespread. Of the 28 alleles that we have identified so far, *Aly13-1* is the only one found in all regions surveyed (Table 4). Most other alleles known to be linked to the S-locus are restricted to a single geographic region (although only the Icelandic population has been sampled extensively and systematically). In contrast, the other allele that seems to be recessive (13-6) is restricted to the Icelandic populations, in which it is found as often as one of the dominant alleles (13-22). We are currently conducting more detailed investigations of the dominance relationships among the alleles described in this study, as well as more extensive surveys of relative frequency of alleles in relation to dominance with the goal of providing more explicit tests of the models described above.

#### Acknowledgements

We thank Jane Charlesworth and Jaclyn Beland for technical assistance and Thor-Ellen Thorhallsdottir for seed collection. The thoughtful advice of two anonymous reviewers significantly improved the manuscript. This work was supported by grants from the Biotechnology and Biological Sciences Research Council of the United Kingdom (to DC but including support for BKM), the Natural Sciences and Engineering Council of Canada (to BKM), and the Danish Natural Sciences Research Council (Grant no. 1262 to MHS).

#### References

- Bateman AJ (1947). Number of S alleles in a population. *Nature* **160**: 337.
- Bateman AJ (1952). Self-incompatibility systems in angiosperms. I. Theory. *Heredity* **6**: 285-310.
- Bateman AJ (1954). Self-incompatibility systems in angiosperms. II. *Iberis amara*. *Heredity* **8**: 305-332.
- Boyes DC, Nasrallah JB (1993). Physical linkage of the SLG and SRK genes at the self-incompatibility locus of *Brassica oleracea*. *Mol Gen Genet* **236**: 369-373.



- Brace J, King GJ, Ockendon DJ (1994). A molecular approach to the identification of S-alleles in *Brassica oleracea*. *Sex Plant Reprod* 7: 203–208.
- Campbell JM, Lawrence MJ (1981). The population genetics of the self-incompatibility polymorphism in *Papaver rhoas*. II. The number and frequency of S-alleles in a natural population (R106). *Heredity* 46: 181–190.
- Charlesworth D (1988). Evolution of homomorphic sporophytic self-incompatibility. *Heredity* 60: 445–453.
- Charlesworth D, Awadalla P, Mable BK, Schierup MH (2000). Population-level studies of multiallelic self-incompatibility loci, with particular reference to Brassicaceae. *Ann Bot* 85(Suppl. A): 227–239.
- East EH, Yarnell SM (1929). Studies on self-sterility. VIII. Self-sterility allelomorphs. *Genetics* 14: 455–487.
- Emerson S (1938). The genetics of self-incompatibility in *Oenothera organensis*. *Genetics* 23: 190–202.
- Emerson S (1939). A preliminary survey of the *Oenothera organensis* population. *Genetics* 24: 524–537.
- Fisher RA (1947). Number of self-sterility alleles. *Nature* 160: 797–798.
- Fisher RA (1958). *The Genetical Theory of Natural Selection*, 2nd edn. Dover Publications: New York.
- Ford MA, Kay QON (1985). The genetics of incompatibility in *Sinapis arvensis* L. *Heredity* 54: 99–102.
- Hatekayama K, Watanabe M, Takasaki T, Ojima K, Hinata K (1998). Dominance relationships between S-alleles in self-incompatible *Brassica campestris* L. *Heredity* 80: 241–247.
- Junghans SH, Metzloff M (1990). A simple and rapid method for the preparation of total plant DNA. *Biotechniques* 8: 176.
- Karron JD, Marshall DL, Oliveras DM (1990). Numbers of sporophytic self-incompatibility alleles in populations of wild radish. *Theor Appl Genet* 79: 457–460.
- Kowyama Y, Takahasi H, Muraoka K, Tani T, Hara K, Shiotani L (1994). Number, frequency and dominance relationships of S-alleles in diploid *Ipomoea trifida*. *Heredity* 73: 275–283.
- Kusaba M, Dwyer KG, Hendershot J, Vrebalov J, Nasrallah JB, Nasrallah ME (2001). Self-incompatibility in the genus *Arabidopsis*: characterization of the S locus in the outcrossing *A. lyrata* and its autogamous relative *A. thaliana*. *Plant Cell* 13: 627–643.
- Kusaba M, Nishio T, Satta Y, Hinata K, Ockendon D (1997). Striking sequence similarity in inter- and intra-specific comparisons of class I *SLG* alleles from *Brassica oleracea* and *Brassica campestris*: implications for the evolution and recognition mechanism. *Proc Natl Acad Sci USA* 94: 7673–7678.
- Kusaba M, Tung C-W, Nasrallah ME, Nasrallah JB (2002). Monoallelic expression and dominance interactions in anthers of self-incompatible *Arabidopsis lyrata*. *Plant Phys* 128: 17–20.
- Lawrence MJ (2000). Population genetics of the homomorphic self-incompatibility polymorphisms in flowering plants. *Ann Bot* 85(Suppl. A): 221–226.
- Maddison DR, Maddison WP (2000). *MacClade 4: Analysis of Phylogeny and Character Evolution. Version 4.0*. Sinauer Associates: Sunderland, MA.
- Nasrallah JB (2000). Cell-cell signalling in the self-incompatibility response. *Curr Opin Plant Biol* 3: 368–373.
- Nasrallah ME, Kandasamy M, Chang M-C, Stadler Z, Lim S, Nasrallah JB (2000). Identifying genes for pollen-stigma recognition in crucifers. *Ann Bot* 85(Suppl. A): 125–132.
- Nei M, Gojobori T (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 3: 418–426.
- Nishio T, Kusaba M (2000). Sequence diversity of *SLG* and *SRK* in *Brassica oleracea* L. *Ann Bot* 85(Suppl. A): 141–146.
- Nou IS, Watanabe M, Isuzugawa K, Isogai A, Hinata K (1993a). Comparison of S-alleles and S-glycoproteins in a naturalized population of self-incompatible *Brassica campestris*. L. *Jpn J Genet* 66: 227–239.
- Nou IS, Watanabe M, Isuzugawa K, Isogai A, Hinata K (1993b). Comparison of S-alleles from a wild population of *Brassica campestris* L. at Balcesme, Turkey and their characterisation by S-glycoproteins. *Sex Plant Reprod* 6: 71–78.
- Ockendon DJ (1974). Distribution of self-incompatibility alleles and breeding structure of open-pollinated cultivars of brussels sprouts. *Heredity* 33: 159–171.
- O'Donnell S, Lawrence MJ (1984). The population genetics of the self-incompatibility polymorphism in *Papaver rhoas*. IV. The estimation of the number of alleles in a population. *Heredity* 53: 495–507.
- Paxman GJ (1963). Maximum likelihood estimation of the number of self-sterility alleles in a population. *Genetics* 48: 1029–1032.
- Rambaut A (1996). Se-AL: sequence alignment editor, version 1.0 alpha1. Distributed over the world wide web (<http://evolve.zps.ox.ac.uk>).
- Richman AD, Kao T-H, Schaeffer SW, Uyenoyama MK (1995). S-allele sequence diversity in natural populations of *Solanum carolinense* (Horsenettle). *Heredity* 75: 405–415.
- Richman AD, Uyenoyama MK, Kohn JR (1996). S-allele diversity in a natural population of *Physalis crassifolia* (Solanaceae) (ground cherry) assessed by RT-PCR. *Ann Bot* 85: 241–245.
- Rojas J, Rojas R (1999). DNAsp version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* 15: 174–175.
- Sampson DR (1964). A one-locus self-incompatibility system in *Raphanus raphanistrum*. *Can J Genet Cytol* 6: 435–445.
- Sampson DR (1967). Frequency and distribution of self-incompatibility alleles in *Raphanus raphanistrum*. *Genetics* 56: 241–251.
- Sampson DR (1974). Equilibrium frequencies of sporophytic self-incompatibility alleles. *Can J Genet Cytol* 16: 611–618.
- Schierup MH (1998). The number of self-incompatibility alleles in a finite, subdivided population. *Genetics* 149: 1153–1162.
- Schierup MH, Mable BK, Awadalla P, Charlesworth D (2001). Identification and characterization of a polymorphic receptor kinase gene linked to the self-incompatibility locus of *Arabidopsis lyrata*. *Genetics* 158: 387–399.
- Schierup MH, Vekemans X, Charlesworth D (2000). The effect of subdivision on variation at multi-allelic loci under balancing selection. *Genet Res* 76: 51–62.
- Schierup MH, Vekemans X, Christiansen FB (1997). Evolutionary dynamics of sporophytic self-incompatibility alleles in plants. *Genetics* 147: 835–846.
- Schierup MH, Vekemans X, Christiansen FB (1998). Allelic geneologies in sporophytic self-incompatibility systems in plants. *Genetics* 150: 1187–1198.
- Schopfer CR, Nasrallah ME, Nasrallah JB (1999). The male determinant of self-incompatibility in *Brassica*. *Science* 286: 1697–1700.
- Shiba H, Iwano M, Entani T, Ishimoto K, Che F-S, Satta Y, Ito A, Takada Y, Watanabe M, Isogai A, Takayama S (2002). The dominance of alleles controlling self-incompatibility in *Brassica* pollen is regulated at the RNA level. *Plant Cell* 14: 491–504.
- Stein JC, Dixit R, Nasrallah ME, Nasrallah JB (1996). SRK, the stigma-specific S locus receptor kinase of Brassica, is targeted to the plasma membrane in transgenic tobacco. *Plant Cell* 8: 429–445.
- Stevens JP, Kay QON (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.
- Tajima F, Tokunaga T, Miyashita NT (1994). Statistical methods for estimating the effective number of alleles, expected heterozygosity and genetic distance in self-incompatibility locus. *Jpn J Genet* 69: 287–295.
- Takasaki T, Hatakeyama K, Suzuki G, Watanabe M, Isogai A, Hinata K (2000). The S receptor kinase determines self-incompatibility in *Brassica* stigma. *Nature* 403: 913–916.

- Takayama S, Shiba H, Iwano M, Shimosato H, Che F-S (2000). The pollen determinant of self-incompatibility in *Brassica campestris*. *Proc Natl Acad Sci* **97**: 1920–1925.
- Thompson KF, Taylor JP (1966). Non-linear dominance relationships between S alleles. *Heredity* **21**: 345–362.
- Ueyonama MK (2000). Evolutionary dynamics of self-incompatibility alleles in Brassica. *Genetics* **156**: 351–359.
- Vekemans X, Schierup MH, Christianssen FB (1998). Mate availability and fecundity selection in multiallelic self-incompatibility systems in plants. *Evolution* **52**: 19–29.
- Wallace DH (1979). Procedures for identifying S-allele genotypes in Brassica. *Theor Appl Genet* **54**: 249–265.
- Watanabe M, Ito A, Takada Y, Ninomiya C, Kakizaki T, Takahata Y *et al* (2000). Highly divergent sequences of the pollen self-incompatibility (S) gene in class-I haplotypes of *Brassica campestris* (syn. *rapa*) L. *FEBS Lett* **473**: 139–144.
- Wright S (1939). The distribution of self-sterility alleles in populations. *Genetics* **24**: 538–552.
- Wright S (1960). On the number of self-incompatibility alleles maintained in equilibrium by a given mutation rate in a population of a given size: a re-examination. *Biometrics* **16**: 61–85.
- Wright S (1964). The distribution of self-incompatibility alleles in populations. *Evolution* **18**: 609–618.
- Yokoyama S, Nei M (1979). Population dynamics of sex-determining alleles in honey bees and self-incompatibility alleles in plants. *Genetics* **91**: 609–626.