

S-allele sequence diversity in natural populations of *Solanum carolinense* (Horsenettle)

ADAM D. RICHMAN*†, TEH-HUI KAO‡, STEPHEN W. SCHAEFFER§
& MARCY K. UYENOYAMA†

†Department of Zoology, Box 90325, Duke University, Durham, NC 27708–0325, ‡Department of Biochemistry and Molecular Biology and Institute of Molecular Evolutionary Genetics, The Pennsylvania State University, University Park, PA 16802, and §Department of Biology and Institute of Molecular Evolutionary Genetics, The Pennsylvania State University, University Park, PA 16802, U.S.A.

S-allele diversity in *Solanum carolinense* was surveyed in two natural populations, located in Tennessee and North Carolina, with a molecular assay to determine the genotype of individual plants. A total of 13 different *S*-alleles were identified and sequenced. There is high overlap between the two populations sampled, with 10 alleles shared in common, one allele found only in Tennessee, and two found only in North Carolina. The number of alleles in this species appears to be extremely low compared with other species with gametophytic self-incompatibility. Sequence comparisons show that most alleles are extremely different one from another in their primary sequence and a phylogenetic analysis indicates extensive trans-specific evolution of *S*-lineages. In addition, some alleles appear to be derived much more recently. The implications of these observations are discussed in the light of recent theoretical results on *S*-allele population diversity and persistence.

Keywords: RT-PCR, *S*-allele diversity, self-incompatibility, *Solanum carolinense*.

Introduction

Homomorphic self-incompatibility systems prevent selfing by inhibiting the germination or growth of pollen that express specificities held in common with the seed parent. Molecular and classical genetic studies indicate that determination of specificity in pollen and style is encoded in many species by a single locus with multiple alleles, designated the *S*-locus. In gametophytic self-incompatibility (GSI) systems, pollen specificity is determined by the haploid genotype of the gametophyte or pollen grain, while in sporophytic self-incompatibility (SSI) systems, pollen specificity is determined by the diploid pollen parent. *S*-locus polymorphism has been the subject of empirical and theoretical study for many decades, as an example of extreme diversifying selection, for which the nature of the selective forces is well understood. The conceptual basis for polymorphism in GSI systems is particularly simple because dominance in the determination of pollen specificity, typ-

ical of sporophytic systems, is absent. Single pollen donor crosses fall into one of three categories: compatible, half-compatible or incompatible, corresponding to zero, one or two *S*-alleles shared by the parents. The opportunity for mating associated with any allele is inversely related to its frequency, promoting the maintenance of many alleles within populations, as has been found in a number of empirical studies (Emerson, 1939; Ockendon, 1974; Campbell & Lawrence, 1981; Levin, 1993).

S-allele diversity in population samples has been investigated with diallel crosses to determine compatibility, as assessed by examination of pollen tube growth in the style (Emerson, 1938; Lawrence, 1975) or seed-set (Levin, 1993). Recently, the genes implicated in determining specificity have been characterized for the sporophytic system in *Brassica* (Nasrallah *et al.*, 1987) and for three different gametophytic systems: in solanaceous taxa (Anderson *et al.*, 1986), the field poppy *Papaver rhoeas* (Foote *et al.*, 1994), and the grass *Phalaris coerulea* which has a two-locus system (Li *et al.*, 1994). This advance promises to revolutionize the study of SI at the population as well as the molecular level. In this

*Correspondence and present address: Department of Biology 0116, The University of California at San Diego, La Jolla, CA 92093-0116, U.S.A.

paper we describe the direct determination of the *S*-locus genotype of individual plants sampled from natural populations through the use of the polymerase chain reaction (PCR) procedure, as has recently been used in the SSI system with cultivars of rapeseed (Brace *et al.*, 1993, 1994). This approach circumvents the laborious necessity of large crossing experiments to assess *S*-locus genotype and, further, provides sequence information. Other advantages of the direct assay include unambiguous comparisons of studies of *S*-diversity carried out at different times and in different species, and the ability to study plants not amenable to greenhouse crossing experiments, such as tree species or plants producing limited numbers of flowers.

Here we report on *S*-allele diversity in two natural populations of the Horsenettle, *Solanum carolinense* (Solanaceae), which expresses gametophytic self-incompatibility. The molecular genetics of self-incompatibility in the Solanaceae is particularly well known (reviewed by Clarke & Newbigin, 1993; Hinata *et al.*, 1993; also Lee *et al.*, 1994) and a number of *S*-allele cDNA sequences have been used to design *S*-locus specific PCR primers (Ai *et al.*, 1990; Kheyr-Pour *et al.*, 1990; Xu *et al.*, 1990; Kaufmann *et al.*, 1991). To enhance the specificity of our assay, we used reverse transcription-PCR to exploit the abundant expression of the *S*-gene message in mature styles of SI solanaceous taxa (Kirch *et al.*, 1989). PCR products from individual plants were analysed by RFLP analyses and DNA sequencing to determine their *S*-locus genotype. This approach was successful in amplifying two sequences of appropriate size from each individual examined, consistent with the expectation that all individuals are heterozygous under GSI. These sequences show homology to published *S*-allele sequences and the pattern of inheritance of a subset of these sequences was consistent with predictions for a gametophytic *S*-locus. The discovery rate of new *S*-alleles in population samples was used to estimate population *S*-allele diversities and the degree of overlap between population samples was assessed at the allelic and sequence levels. We compare our surveys of *S*-diversity in two populations of Horsenettle with studies of *S*-allelic diversity in other taxa.

Materials and methods

Collection of plant materials

Horsenettle is a weedy perennial of disturbed and open habitats in the south-east United States. Individuals die back completely above ground in winter,

resprout from rhizomes in the spring and flower in late summer. Rhizomes were collected for propagation in the greenhouse from two Horsenettle populations, 150 miles apart, in Washington County, Tennessee (TN) and Orange County, North Carolina (NC). Thirty-six rhizomes, collected from TN in April 1993, were located by the presence of flowering stalks from the previous year; the majority of these were scattered along the perimeter of a fallow pasture field of approximately 0.25 ha. An effort was made to collect across the entire area, collecting samples separated by at least several meters. A total of 36 rhizomes was collected in NC in June 1993, 30 of which were sampled at 3 m intervals on a grid centred in an area where Horsenettle plants were observed (0.1 ha), and an additional six samples were collected at random off the grid. This sampling procedure was repeated in May 1994 to replace plants lost to frost damage. Although collections of rhizomes at the NC site, made in late spring and summer, did not rely on the presence of old flowering stalks, the majority of plants sampled appeared to be derived from rhizomes, rather than from recent germinations from seed.

Reverse transcription-PCR

Styles collected 1–2 days prior to anthesis were either processed immediately or frozen at -70°C . Styler RNA was extracted using a modification of a protocol from Ai *et al.* (1990). One to several styles were ground in liquid nitrogen, then reground in 0.4 mL equal volumes 25:24:1 phenol/chloroform/isoamyl alcohol (PCI) and extraction buffer (50 mM Tris, pH 8.5, 10 mM EDTA, 2 per cent SDS). This mixture was centrifuged for 15 min at 4°C , the aqueous layer removed and extracted with PCI two more times, then precipitated by addition of 10 per cent 3 M sodium acetate and 2.5 vol. ethanol. After high speed centrifugation for 30 min at 4°C , the pellet was washed with 70 per cent ethanol and resuspended in 20 μL sterile water. An aliquot of 2–7 μL was removed for cDNA synthesis using the cDNA cycle kit (Invitrogen) and oligo dT primer. A 2 μL aliquot of the cDNA synthesis reaction was used directly in a PCR reaction without further purification, using degenerate *S*-locus-specific primers. Three primers were used: an upstream primer (PR1), targeted to the conserved region designated as C2 by Ioerger *et al.* (1991), in combination with one of two downstream primers (PR2 and PR3), corresponding to the conserved regions C3 and C4. Primer sequences were as follows: PR1: 5' GAA TTC A(TC)G GN(TC) TNT GGC CNG A 3'; PR2:

5' NNAAGC TTT CT(AG) (ACT)AC (AT)AC A(CG)G TTC CAT GCT T 3'; PR3: 5' NNG TCG AC(GA) AAA CAT ATN CCT A(TC)(TC) TCN (ATC)NT AG(TC) TC 3'. PCR was carried out in a 50 μ L volume of 50 mM KCl, 10 mM Tris HCl, pH 8.0, 1.5 mM MgCl₂, 3.0 mM DTT, 0.1 mg mL⁻¹ BSA, 0.2 mM dNTPs, 2 units AmpliTaq Taq polymerase and 100 ng of each primer. A single drop of mineral oil covered each PCR reaction. Following an initial 5 minute denaturation step at 94°C, amplification occurred over 35–40 cycles of 2 min at 45°C, 1.5 min at 72°C and 1 min at 94°C.

Cloning

RFLP analyses of PCR amplification products from individual plants indicated the presence of more than one product of closely similar size (see Results). PCR products were cloned, without purification using the Invitrogen TA cloning kit to separate these products for sequencing. Cloned sequences were amplified from individual bacterial colonies. Colonies were transferred with a sterile pipette tip to a 0.5 mL Eppendorf tube and resuspended in 15 μ L 10 mM Tris, pH 7.5, 1 mM EDTA and 50 μ g mL⁻¹ Proteinase K, incubated for 15 min at 55°C, 15 min at 80°C, spun at high speed on a desktop centrifuge for 15 min, and 8 μ L removed to a new tube for PCR. Amplification was carried out in a volume of 25 μ L, using m13 forward and reverse primers and 0.5 units Taq polymerase. PCR conditions were as above, except that the low temperature annealing step was performed at 55°C. PCR products of appropriate size were screened using restriction enzyme analysis to identify different cloned sequences.

Sequencing

A preliminary survey indicated that alleles with the same RFLP patterns from different individuals were identical in sequence whereas alleles with different RFLP patterns showed extensive sequence differences. When a new RFLP pattern was encountered, several clones were isolated and the sequence determined for both strands. Subsequent identifications of this allele in other individuals were based on either RFLP analyses or partial sequences, with the exception that the above sequencing strategy was repeated when an allele already encountered was first found in a new population sample.

Amplification products were purified for sequencing either by: (i) precipitation in 10 per cent 4 M ammonium acetate and 2 vol ethanol, spinning at

high speed at room temperature for 5 min, washing with 70 per cent EtOH, air or vacuum drying and resuspending in 20 μ L ddH₂O, or (ii) removing 20 μ L of the reaction from under mineral oil with a sterile pipette tip, washing twice in 500 μ L ddH₂O using a Microcon 30 spin centrifugation unit as per the manufacturer's recommendations, and adjusting the final volume to 20 μ L ddH₂O. Purified PCR product (2–5 μ L) was used as the sequencing template. Clones were sequenced using the BRL cycle sequencing kit and P33 end-labelled m13 primers. Sequencing primer concentration and cycling conditions conformed to the manufacturer's recommendations. Sequencing gels were immersed in 5 per cent acetic acid for 10 min, vacuum dried on 3 MM paper and exposed to X-ray film for 1–2 days.

Results

PCR and restriction fragment length polymorphism analyses

From among plants grown from rhizomes collected in the field and which subsequently flowered in the greenhouse, 30 collected from TN in 1993 and 24 from NC (eight in 1993 and 16 in 1994) were chosen for genetic analysis. RT-PCR from stylar RNA from individual plants resulted in a single band on an agarose gel of the size expected from published S-allele cDNA sequences. In most cases, restriction enzyme analysis of the products indicated the presence of more than one sequence. In some instances, it appeared that only one product was obtained for the primer combination PR1, PR3. In these instances, a second set of primers (PR1, PR2) yielded an additional product, as determined by RFLP analyses or DNA sequencing.

Sequence analysis

Plants yielding unique RFLP patterns were chosen for sequencing. In all, 13 different sequences were found (GenBank accession numbers L40539–L40551), 11 in the TN sample and 12 in NC. The sequences obtained are consistent with the observed RFLP patterns. Alleles with identical RFLP patterns common to both TN and NC (10 total) were found to be identical in sequence as well. Allele S₃ was found only in TN, alleles S₁₀ and S₁₃ only in NC. Inferred amino acid sequences for the 13 Horse-nettle sequences, 10 S-allele cDNAs and 3 homologous S-related RNases are shown in Fig. 1. Figure 2 shows the phylogeny reconstructed by the neighbour-joining algorithm (Saitou & Nei, 1987), as

RNS1	WFDYKGGTY-PENC-DASDQ-FDSST-ISDLTLBMKRWPTLAC--PQSGSAPWSEHCKICE GT
LE	WPNNDGTY-PENCDFNS-P-YDQ-SQISDLISBQQNWPILAC--PQSGSPTFWSEHCKICE GTCAMSVLTHQAY FFKALDLNQID LLSLQGGADI HPDQGSYDLVNRKAIKSAI--GTFWQCNDV-QSBSQI
X2	WFDN--VSTFLAYCSEKTKQ-YNKLQOPTIKNS-LYKRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S1NIC	WFDN--VSTFLAYCSEKTKQ-LPFDKDKKQD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S2NIC	WFDN--HTFLAYCSEKTKQ-KP-YNKLQOPTIKNS-LYKRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S3NIC	WFDN--VSTFLAYCSEKTKQ-YKLLDDKDKKQD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S2SOL	WFDN--KICLLANNC-RBYA--YNALNVRKQK-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S3SOL	WFDN--VSTFLAYCSEKTKQ-KIVK--YNKLQOPTIKNS-LYKRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S1PET	WPEI--TQFLAYCSEKTKQ-K-YRFLKDKKQD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S2PET	WPEI--KIFLAYCSEKTKQ-K-YRFLKDKKQD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S3PET	WPEI--KIFLAYCSEKTKQ-K-YRFLKDKKQD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
SF11	WFDN--VSTFLAYCSEKTKQ-KP-KPTYSYPT-GOMLAD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
SZNIC	WFDN--VSTFLAYCSEKTKQ-KP-KPTYSYPT-GOMLAD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S13C	WFDN--VSTFLAYCSEKTKQ-KP-KPTYSYPT-GOMLAD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S10SC	WFDN--VSTFLAYCSEKTKQ-KP-KPTYSYPT-GOMLAD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S4SC	WFDN--VSTFLAYCSEKTKQ-KP-KPTYSYPT-GOMLAD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S2SC	WFDN--VSTFLAYCSEKTKQ-KP-KPTYSYPT-GOMLAD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S8SC	WFDN--VSTFLAYCSEKTKQ-KP-KPTYSYPT-GOMLAD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S7SC	WFDN--VSTFLAYCSEKTKQ-KP-KPTYSYPT-GOMLAD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S9SC	WFDN--VSTFLAYCSEKTKQ-KP-KPTYSYPT-GOMLAD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S3SC	WFDN--VSTFLAYCSEKTKQ-KP-KPTYSYPT-GOMLAD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S5SC	WFDN--VSTFLAYCSEKTKQ-KP-KPTYSYPT-GOMLAD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S6SC	WFDN--VSTFLAYCSEKTKQ-KP-KPTYSYPT-GOMLAD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S11SC	WFDN--VSTFLAYCSEKTKQ-KP-KPTYSYPT-GOMLAD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL
S12SC	WFDN--VSTFLAYCSEKTKQ-KP-KPTYSYPT-GOMLAD-LDRWFDLVTBETOCLEKQNFWRKRYNKE GTC CSBRY - NQQY FFLANALDKKDFW LLTFLRHHGI I PGGN- YVQKINPTIKTIT-RGY-FMLACTS-----MEL

Fig. 1 Inferred amino acid sequences for thirteen cDNA sequences isolated from *Solanum carolinense* by RT-PCR, ten published *S*-allele cDNAs, and 3 *S*-locus related RNases. Sequences were aligned by eye. Published cDNA sequences were derived from *Nicotiana glauca* (SF11, SZNIC, S1–3NIC: Kheyr-Pour *et al.*, 1990, and references cited therein), *Solanum chacoense* (S2–S3SOL: Xu *et al.*, 1990), *Petunia inflata* (S1–3PET: Ai *et al.*, 1990; RNase: Lee *et al.*, 1992), *Lycopersicon esculentum* (LE: Jost *et al.*, 1991), and *Arabidopsis thaliana* (RNS1: Taylor & Green, 1991). The first three and last two amino acids in the *S. carolinense* sequences are based on primer sequences, and were included here for purposes of alignment but excluded from subsequent calculations of percent divergence at the amino acid and DNA sequence levels.

implemented in the Molecular Evolution Genetic Analysis package (Kumar *et al.*, 1993) from the pairwise percentage amino acid distances among these sequences (Table 1). The sequences for *S*₁₁, *S*₁₂ and the *S*-related RNase *RNS1* are omitted from Table 1 because only partial sequence information relative to the other alleles sampled was available. Their position in the phylogeny appears highly reliable, as assessed by bootstrap resampling, but the estimates of branch lengths (not shown) may be subject to error. Percentage similarity at synonymous and non-synonymous nucleotide sites is shown for all 13 *Horsenettle* sequences in Tables 2 and 3.

Single donor crosses

A number of individuals showing different RFLP patterns were crossed in the greenhouse, the progeny raised to flowering and the *S*-locus genotypes of the offspring determined by RT-PCR and RFLP analysis. One family flowered in 1994: a cross between NC43 with genotype *S*₁₁*S*₁₂ as the ovule parent and NC31 with genotype *S*₁₂*S*₁₃ the pollen

parent. As predicted for a half-compatible mating, only two genotypes were observed among the nine progeny: all offspring carried the compatible paternal allele *S*₁₃, six the maternal allele *S*₁₁ and the remaining three the maternal allele *S*₁₂. These results are consistent with Mendelian segregation at a gametophytic *S*-locus, assuming codominance in expression of self-incompatibility by paternal and maternal *S*-alleles.

Number of sequences and genotypes

Table 4 provides the sequence designations of the 26 genotypes observed (14 detected among 24 TN plants surveyed and 12 among 24 NC plants). The equality of frequency of occurrence of these sequences, as expected at equilibrium for a gametophytic *S*-locus, was tested by Mantel's (1974) statistic:

$$\chi^2_{n-1} = (n-1)(\sum C_j^2 - 4r^2/n)/(2r-4r/n),$$

where *C*_{*j*} is the number of times an allele occurs, *n* the number of alleles found and *r* the number of plants (different genotypes) sampled. The test fails

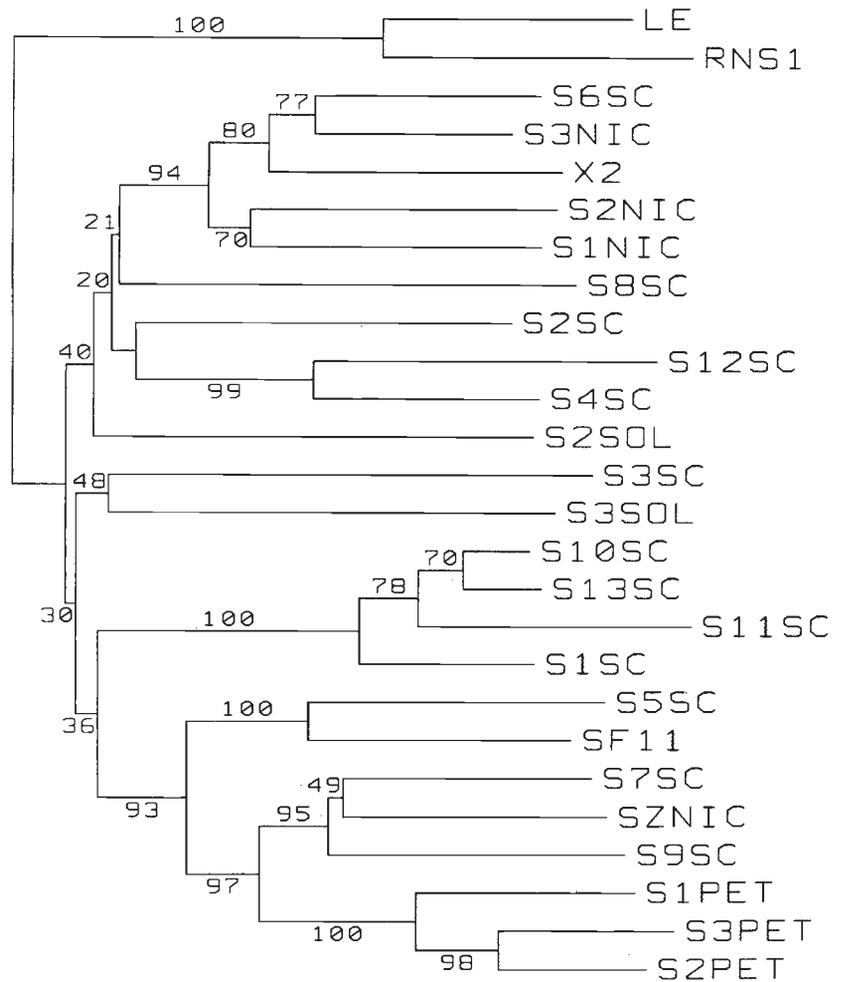


Fig. 2 Neighbour-joining phylogeny for the 23 sequences shown in Fig. 1. Bootstrap values for 1000 replicates are shown next to corresponding nodes.

to reject the hypothesis of equal frequencies for TN ($\chi^2_{13} = 10.9$, $P > 0.50$) and NC ($\chi^2_{11} = 3.3$, $P > 0.95$).

The number of sequences recovered in a finite sample may underestimate the total diversity within the population. An estimate of the total number of alleles in the population (N) is obtained by implicit solution of

$$n = N(1 - (1 - 2/N)^r),$$

where n is the number of alleles observed and r the number of plants examined (Paxman, 1963; O'Donnell & Lawrence, 1984). The estimated number of alleles in TN is 11.9 and in NC is 14.4.

For comparison to other studies, it is useful to summarize the thoroughness of sampling of S -allelic diversity using a repeatability statistic R (Campbell & Lawrence, 1981) calculated as

$$R = (m - n)/(m - 3),$$

where m is the number of alleles examined (equal to $2r$) and n is the number of different alleles found. If n equals m , no repeats have occurred and $R = 0$. R

converges on 1 as n converges on N . If only three different alleles are sampled (the minimum required within a population if it is to persist), then $R = 1$ regardless of the number of alleles examined. For TN, $R = 0.68$ and for NC, $R = 0.57$.

The degree of overlap in S -sequences found in the two population samples was assessed using the method set out in O'Donnell *et al.* (1993) using the foregoing population estimates of 12 and 14 for TN and NC, respectively. The maximum likelihood estimate of the degree of overlap is total overlap of 12 alleles.

Discussion

Specificity of the PCR assay

RT-PCR using stylar RNA as the template for cDNA synthesis resulted in amplification of a single band of expected size in each individual. Each band was subsequently found to contain two different sequences, as determined by RFLP analyses and

Table 1 Percentage similarity in amino acid identity for the sequences listed in Fig. 1, with RNS1, S11SC and S12SC omitted (see text)

	X2	S1NIC	S2NIC	S3NIC	S2SOL	S3SOL	S1PET	S2PET	S3PET	SF11	SZNIC	S1SC	S13SC	S10SC	S4SC	S2SC	S6SC	S7SC	S9SC	S3SC	S5SC	S6SC	S8SC
LE	0.70	0.68	0.66	0.68	0.68	0.67	0.73	0.74	0.75	0.66	0.72	0.66	0.67	0.68	0.69	0.73	0.69	0.69	0.65	0.68	0.79	0.81	
X2		0.43	0.45	0.33	0.55	0.58	0.61	0.62	0.63	0.58	0.58	0.59	0.55	0.55	0.49	0.58	0.58	0.62	0.62	0.34	0.61	0.55	
S1NIC			0.36	0.31	0.52	0.54	0.62	0.64	0.63	0.57	0.57	0.56	0.59	0.58	0.52	0.59	0.57	0.62	0.62	0.39	0.62	0.59	
S2NIC				0.38	0.54	0.53	0.66	0.64	0.65	0.58	0.61	0.59	0.62	0.60	0.54	0.61	0.62	0.62	0.59	0.42	0.68	0.58	
S3NIC					0.53	0.53	0.60	0.62	0.61	0.56	0.57	0.56	0.57	0.56	0.47	0.57	0.60	0.61	0.60	0.25	0.64	0.49	
S2SOL					0.61	0.61	0.62	0.60	0.63	0.58	0.60	0.55	0.54	0.53	0.50	0.59	0.58	0.65	0.59	0.53	0.73	0.60	
S3SOL						0.60	0.62	0.56	0.58	0.60	0.58	0.59	0.64	0.63	0.57	0.56	0.61	0.56	0.60	0.55	0.71	0.60	
S1PET							0.56	0.27	0.27	0.49	0.41	0.57	0.56	0.54	0.58	0.41	0.44	0.63	0.51	0.58	0.68	0.74	
S2PET								0.19	0.19	0.50	0.43	0.59	0.54	0.55	0.61	0.42	0.46	0.63	0.52	0.57	0.67	0.75	
S3PET									0.53	0.53	0.42	0.60	0.56	0.57	0.60	0.44	0.46	0.63	0.54	0.60	0.75	0.82	
SF11										0.53	0.46	0.53	0.55	0.50	0.60	0.45	0.50	0.58	0.34	0.55	0.60	0.76	
SZNIC											0.46	0.57	0.58	0.59	0.59	0.31	0.36	0.67	0.52	0.56	0.75	0.77	
S1SC												0.20	0.19	0.19	0.54	0.54	0.58	0.61	0.58	0.56	0.32	0.65	
S13SC													0.09	0.09	0.53	0.54	0.59	0.61	0.59	0.57	0.25	0.62	
S10SC															0.53	0.54	0.59	0.61	0.57	0.57	0.22	0.62	
S4SC																0.55	0.58	0.58	0.49	0.67	0.51		
S2SC																	0.33	0.61	0.48	0.55	0.73	0.81	
S8SC																		0.64	0.52	0.56	0.76	0.86	
S7SC																			0.56	0.61	0.62	0.65	
S9SC																				0.59	0.67	0.77	
S3SC																					0.62	0.65	
S5SC																						0.62	0.60

Table 2 Percentage similarity at synonymous sites for the 13 Horsenettle sequences

	S13	S10	S11	S4	S12	S2	S8	S7	S9	S3	S5	S6
S1	0.09	0.09	0.02	0.61	0.58	0.48	0.54	0.62	0.62	0.62	0.58	0.67
S13		0.12	0.03	0.58	0.59	0.53	0.58	0.63	0.65	0.60	0.56	0.65
S10			0.07	0.60	0.56	0.46	0.52	0.60	0.62	0.58	0.56	0.62
S11				0.60	0.58	0.45	0.71	0.66	0.70	0.70	0.63	0.80
S4					0.22	0.50	0.50	0.52	0.51	0.44	0.43	0.47
S12						0.60	0.55	0.51	0.58	0.50	0.51	0.50
S2							0.54	0.45	0.52	0.48	0.46	0.54
S8								0.41	0.45	0.53	0.41	0.38
S7									0.55	0.53	0.40	0.53
S9										0.58	0.51	0.53
S3											0.44	0.38
S5												0.47

Alignment was as for the amino acid sequences shown in Fig. 1.

Table 3 Percentage similarity at nonsynonymous sites for the 13 Horsenettle sequences

	S13	S10	S11	S4	S12	S2	S8	S7	S9	S3	S5	S6
S1	0.11	0.10	0.15	0.37	0.45	0.34	0.39	0.35	0.42	0.39	0.39	0.35
S13		0.08	0.12	0.37	0.45	0.32	0.38	0.34	0.40	0.38	0.38	0.36
S10			0.10	0.38	0.44	0.34	0.41	0.36	0.42	0.38	0.37	0.36
S11				0.53	0.47	0.45	0.55	0.45	0.55	0.43	0.43	0.41
S4					0.20	0.31	0.35	0.44	0.47	0.39	0.43	0.38
S12						0.36	0.36	0.55	0.60	0.41	0.47	0.41
S2							0.33	0.38	0.40	0.36	0.39	0.32
S8								0.43	0.44	0.40	0.40	0.36
S7									0.20	0.39	0.32	0.35
S9										0.43	0.34	0.37
S3											0.38	0.41
S5												0.39

Alignment was as for the amino acid sequences shown in Fig. 1.

DNA sequencing, consistent with the expectation that all individuals are heterozygous under GSI. Segregational analysis for three sequences in one cross is consistent with the prediction for alleles at a single gametophytic *S*-locus. A phylogenetic analysis indicates that the Horsenettle sequences are most closely related to published *S*-allele cDNA sequences from other solanaceous GSI taxa (Fig. 2). The RNase *X2* also falls within this group, reflecting its high homology with *S*-alleles. *X2* is expressed in the style of *Petunia inflata* (Lee *et al.*, 1992), suggesting that homologous style-specific RNases are a potential source of artefacts. However, *X2* is monomorphic in *P. inflata* (Lee *et al.*, 1992); while such a monomorphic product would be readily detected in RFLP analyses, none was observed. RFLP analyses

invariably indicated the presence of (at most) two sequences per individual whereas amplification of related regions in addition to the *S*-locus would be expected to generate at least three sequences. One reason for the apparent specificity of the assay is that stylar RNA is greatly enriched in *S*-allele message, so that sequences from other RNases were less likely to be detected.

Estimating the number of S alleles in the population

Eleven different sequences were detected in the TN population and 12 in NC from 14 and 12 individuals with distinguishable RFLP patterns, respectively. This number is considerably smaller than those

Table 4 Genotypes identified in two Horsenettle population samples

ID no.	TN	ID no.	NC
113	S_1S_2	50	S_1S_5
115	S_1S_3	7	S_1S_7
151	S_1S_8	28*	S_2S_5
125	S_1S_9	31*	S_2S_{10}
138	S_2S_4	20	S_4S_6
152	S_2S_5	16	S_4S_7
129	S_2S_6	11*	S_5S_{13}
123	S_2S_7	24*	S_6S_9
143	S_2S_{12}	29*	S_6S_{11}
109	S_3S_4	30*	S_8S_{13}
107	S_3S_9	43	$S_{12}S_{11}$
122	S_4S_6	31	$S_{12}S_{13}$
146	S_5S_9		
136	S_6S_{11}		

Sequence designations are as in Fig. 1.

*Indicates 1994 collection, otherwise all plants were collected in 1993.

reported for other species with gametophytic self-incompatibility, based on crossing experiments. For example, Levin (1993) reported finding 30 alleles in a sample of 24 individuals of *Phlox drummondii*, and in *Papaver rhoeas*, Lawrence & O'Donnell (1981) found 30 alleles in 36 plants in one population sample, 26 in 36 plants from another. While relatively few genotypes were sampled in the Horsenettle populations, the estimates for population diversity are only slightly larger than the sampled diversities. Moreover, the values of R , a measure of thoroughness of sampling of S -allelic diversity, for the population samples from TN and NC indicate that these surveys are as thorough as those carried out in the majority of other self-incompatible species (see table 1 in O'Donnell & Lawrence, 1984). Thus the disparity between censuses of Horsenettle and other SI species is not because of limited sampling in the former. Vekemans & Slatkin's (1994) study of the relationship between effective population size and the number of S -alleles maintained at equilibrium under GSI suggests that an effective population size of about 500–1000 is required to maintain 12 alleles at equilibrium, assuming mutation rates (to new specificities) between 10^{-6} and 10^{-9} per generation (see also Wright, 1960).

Tests for inequality of frequency of sequences sampled from TN and NC were not significant, a result consistent with expectation for alleles segregating at a gametophytic S -locus. In other GSI species, O'Donnell & Lawrence (1984) found that Emerson's (1939) data on S -allele frequencies in

Oenothera organensis gave a very good fit to the expectations of equal frequencies whereas S -allele frequencies in *Papaver rhoeas* deviated significantly for all population samples tested (O'Donnell & Lawrence, 1984). Levin (1993) found very good agreement with expectation in a population of *Phlox drummondii*. Elucidation of the causes of significantly unequal allele frequencies in *Papaver* is under active investigation: stochastic variation in allele frequencies alone appears insufficient (Lawrence *et al.*, 1994), and there is evidence for selection, possibly at linked loci, reflected in biased transmission of particular S -alleles and genotypes in single donor families (Lawrence & Franklin-Tong, 1994).

S-diversity and gene flow

The NC and TN sites are regarded as separate populations by reason of the wide geographical separation and the patchy distribution of this species in NC and TN, which is as expected given its weedy habit. However, gene flow is suggested by the high degree of overlap (10 of a total of 13) in the sequences found in the two populations and by the absence of sequence differences within each presumptive S -allelic class between sites. The limited sampling does not permit determination of whether the sites are at all differentiated with respect to S -locus variation. Lack of population structure was inferred in studies of *P. rhoeas* collected from British sites separated by at least 43 km, with each population sample appearing to have been drawn from the same pool of S -alleles (Campbell & Lawrence, 1981; Lawrence *et al.*, 1993). Horsenettles produce conspicuous, fleshy fruits which remain on the fruiting stalk and may be dispersed by birds, suggesting that long-distance dispersal may be a frequent event in this species. In addition, the effective migration rate for the S -locus may be higher than at other loci as immigrant S -alleles unrepresented in a population have a mating advantage over more common resident alleles.

The Horsenettle sites sampled are small in area, permitting the possibility that larger sites contain additional S -sequences. However, if the NC and TN sequences represent independent samples of the S -allelic diversity in this species, then it appears that Horsenettle S -diversity is in fact greatly impoverished with respect to other SI species. For example, *P. rhoeas* has been estimated to contain 66 different alleles across its range from samples in Spain and Britain, and more than 50 based on samples in Britain alone (Lane & Lawrence, 1993). The biochemical mechanism underlying GSI in *P. rhoeas* has been

shown to differ from that found in the Solanaceae (Franklin-Tong & Franklin, 1993; Foote *et al.*, 1994), raising the possibility that constraints on the origin of new *S*-alleles in the latter may be responsible for the marked difference in *S*-diversity in the two taxa. However, this explanation seems unlikely given that the number of different *S*-allele sequences reported in the Solanaceae, including the 13 reported here, exceeds 30. Moreover, a parallel study of *S*-diversity as assessed by RT-PCR in another solanaceous species, *Physalis crassifolia*, detected high within-population diversity similar to that found in non-solanaceous GSI taxa (A. D. Richman *et al.*, unpublished data). Perhaps low diversity in Horse-nettle is a consequence of its particular population structure (many small and probably quite short-lived colonies), which suggests small effective population size.

Phylogeny and the age of S-alleles

Despite evidence that very few base substitutions can determine different *S*-allele specificities (Royo *et al.*, 1994; Saba-El-Leil *et al.*, 1994), we observed that numerous sequence differences distinguish most specificities found in natural populations (Tables 1–3). From the phylogenetic analysis (Fig. 2), many Horsenettle *S*-sequences appear most closely related to allelic lineages found in other taxa. The pattern of trans-specific sharing of *S*-lineages among solanaceous species has been noted previously (Ioerger *et al.*, 1990). Vekemans & Slatkin's (1994) numerical simulation study suggests that the time to coalescence of 10 common allelic lines (the remaining three alleles in Horsenettle being of relatively recent origin) would be less than 10 million generations, assuming an origination rate greater than or equal to 10^{-7} per generation. In contrast, the observation of trans-specific evolution in Horsenettle suggests that the origin of the majority of alleles predates the divergence of *Nicotiana* and *Solanum* (Fig. 2), which is thought to have occurred perhaps 28 million years ago, as estimated by a molecular clock analysis of RuBCs data (Ioerger *et al.*, 1990). A low origination rate of *S*-allele specificities would be consistent with high lineage age (Wright, 1939, 1965; Vekemans & Slatkin, 1994). Yet, observations of high *S*-diversity in small populations in other GSI taxa argues against an extremely low rate of origination in those systems (Vekemans & Slatkin, 1994), and high *S*-diversity has been found within natural populations of another solanaceous species (A. D. Richman *et al.*, unpublished data). Furthermore, even in Horsenettle, several alleles appear to be of relatively

recent origin, with the average pairwise number of synonymous differences among the most closely related alleles (S_1 , S_{10} , S_{11} and S_{13}) being about a tenth the average number among the remainder (0.0564 [SE 0.0295] vs. 0.5171 [SE 0.0162]). Alternatively, very old lineages might persist even under high rates of mutation to new *S*-alleles if additional selective factors tend to prevent new specificities from becoming common or, in those cases in which new specificities do attain high frequencies, to cause the preferential replacement of the alleles from which they were derived.

Conclusions

Classical studies of the number of *S*-alleles in populations have revealed tremendous allelic diversity at the *S*-locus in homomorphic SI species and inspired a number of theoretical investigations (Wright, 1960, 1965; Nagylaki, 1975; Yokoyama & Hetherington, 1982; Uyenoyama, 1991). Molecular information on *S*-allele sequences has afforded new insights, including the observation of extreme age of *S*-lineages (Ioerger *et al.*, 1990) and an emerging picture of the molecular evolution of this locus (Clark & Kao, 1991), and a renewed development of theory and statistical analysis has followed (Clark, 1993; Vekemans & Slatkin, 1994; Uyenoyama, 1995). Approaches such as the one taken here join these previously separate aspects of research on SI, and new questions and approaches will emerge from this union. Diversification of *S*-lineages in natural populations has occurred in the context of a variety of population structures and life histories, factors that can be expected to influence both *S*-allele diversity and lineage persistence. In this study, the quite low within- and between-population diversity in Horsenettle may reflect the particular life history of this colonizing species. Surveys of *S*-allele diversity in GSI species with different life histories, coupled with more extensive theoretical treatments, should provide greater insight on the interaction of ecology and genetics in the maintenance of *S*-diversity.

Acknowledgments

We thank Dr F. Levy and Dr E. Walker of East Tennessee State University for suggesting *S. carolinense* as a study organism, locating collection sites in Tennessee, and making several collections; S. Good (Pennsylvania State University) and D. Hall (Duke University) for help with field collecting; Dr J. Kohn and Dr T. Price (UCSD) for much useful discussion and for improvement of presentation; an anonymous

reviewer for suggestions which improved the manuscript. Research was supported by Alfred P. Sloan awards in molecular evolution (A.D.R. M.K.U.), NSF grant BSR-9019917 (S.W.S. and T.-L.K.) and Public Health Service grant GM 37841 (M.K.U.).

References

- AI, Y., SINGH, A., COLEMAN, C. E., IOERGER, T. R., KHEYP-POUR, A. AND KAO, T.-H. 1990. Self-incompatibility in *Petunia inflata*: isolation and characterization of cDNAs encoding three *S*-allele-associated proteins. *Sex. Plant Reprod.*, **3**, 130–138.
- ANDERSON, M. A., CORNISH, E. C., MAU, S.-L., WILLIAMS, E. G., HOGGART, R., ATKINSON, A., BONIG, I., GREGO, B., SIMPSON, R., ROCHE, P. J., HALEY, J. D., PENSCHOW, J. D., NIALL, H. D., TREGEAR, G. W., COGLAN, J. P., CRAWFORD, R. J. AND CLARKE, A. E. 1986. Cloning of cDNA for a stylar glycoprotein associated with expression of self-incompatibility in *Nicotiana glauca*. *Nature*, **321**, 38–44.
- BRACE, J., KING G. J. AND OCKENDON, D. J. 1993. Development of a method for the identification of *S*-alleles in *Brassica oleracea* based on digestion of PCR-amplified restriction endonucleases. *Sex. Plant Reprod.*, **6**, 133–138.
- BRACE, J., KING G. J. AND OCKENDON, D. J. 1994. A molecular approach to the identification of *S*-alleles in *Brassica oleracea*. *Sex. Plant Reprod.*, **7**, 203–208.
- CAMPBELL, J. M. AND LAWRENCE, M. J. 1981. 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, A. G. 1993. Evolutionary inferences from molecular characterization of self-incompatibility alleles. In: Takahata, N. and Clark, A. G. (eds) *Mechanisms of Molecular Evolution*, pp. 79–108. Sinauer Associates, Sunderland, MA.
- CLARK, A. G. AND KAO, T.-H. 1991. Excess nonsynonymous substitution at shared polymorphic sites among self-incompatibility alleles of Solanaceae. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 9823–9827.
- CLARKE, A. E. AND NEWBIGIN, E. 1993. Molecular aspects of self-incompatibility in flowering plants. *Ann. Rev. Genet.*, **27**, 257–279.
- 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.
- FOOTE, H. C. C., RIDE, J. P., FRANKLIN-TONG, V. E., WALKER, E. A., LAWRENCE, M. J. AND FRANKLIN, F. C. 1994. Cloning and expression of a novel class of self-incompatibility (*S*) gene from *Papaver rhoeas* L. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 2265–2269.
- FRANKLIN-TONG, V. E. AND FRANKLIN, F. C. H. 1993. Gametophytic self-incompatibility: contrasting mechanisms for *Nicotiana* and *Papaver*. *Trends Cell. Biol.*, **3**, 340–345.
- HINATA, K., WATANABE, M., TORIYAMA, K. AND ISOGAI, A. 1993. A review of recent studies on homomorphic self-incompatibility. *Int. Rev. Cytol.*, **143**, 257–296.
- IOERGER, T. R., CLARK, A. G. AND KAO, T.-H. 1990. Polymorphism at the self-incompatibility locus in Solanaceae predates speciation. *Proc. Natl. Acad. Sci. U.S.A.*, **87**, 9732–9735.
- IOERGER, T. R., GOHLKE, J. R., XU, B. AND KAO, T.-H. 1991. Primary structural features of the self-incompatibility protein in Solanaceae. *Sex. Plant Reprod.*, **4**, 81–87.
- JOST, W., BAK, H., GLUND, K., TERPSTRA, P. AND BEINTEMA, J. J. 1991. Amino acid sequence of an extracellular, phosphate-starvation-induced ribonuclease from cultured tomato (*Lycopersicon esculentum*) cells. *Eur. J. Biochem.*, **198**, 1–6.
- KAUFMANN, H., SALAMINI, F. AND THOMPSON, R. D. 1991. Sequence variability and gene structure at the self-incompatibility locus of *Solanum tuberosum*. *Mol. Gen. Genet.*, **226**, 457–466.
- KHEYP-POUR, A., BINTRIM, S. B., IOERGER, T. R., REMY, R., HAMMOND, S. A. AND KAO, T.-H. 1990. Sequence diversity of pistil *S*-proteins associated with gametophytic self-incompatibility in *Nicotiana glauca*. *Sex. Plant Reprod.*, **3**, 88–97.
- KIRCH, H. H., UHRIG, H., LOTTSPEICH, F., SALAMINI, F. AND THOMPSON, R. D. 1989. Characterization of proteins associated with self-incompatibility in *Solanum tuberosum*. *Theor. Appl. Genet.*, **78**, 581–588.
- KUMAR, S., TAMURA, K. AND NEI, M. 1993. MEGA: *Molecular Evolutionary Genetics Analysis*, version 1.02. The Pennsylvania State University, University Park, PA.
- 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. 1975. The genetics of self-incompatibility in *Papaver rhoeas*. *Proc. R. Soc. Lond. B.*, **188**, 275–285.
- LAWRENCE, M. J. AND FRANKLIN-TONG, V. E. 1994. The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. IX. Evidence of an extra effect of selection acting on the *S*-locus. *Heredity*, **72**, 353–364.
- 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. 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., 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.
- LEE, H.-S., HUANG, S. AND KAO, T.-h. 1994. S proteins control rejection of incompatible pollen in *Petunia inflata*. *Nature*, **367**, 560–563.
- LEE, H.-S., SINGH, A. AND KAO, T.-h. 1992. RNase X2, a pistil-specific ribonuclease from *Petunia inflata*, shares sequence similarity with solanaceous S proteins. *Plant Mol. Biol.*, **20**, 1131–1141.
- LEVIN, D. A. 1993. S-gene polymorphism in *Phlox drummondii*. *Heredity*, **71**, 193–198.
- LI, X., NIELD, J., HAYMAN, D. AND LANGRIDGE, P. 1994. Cloning a putative self-incompatibility gene from the pollen of the grass *Phalaris coerulea*. *Plant Cell*, **6**, 1923–1932.
- MANTEL, N. 1974. Approaches to a health research occupancy problem. *Biometrics*, **30**, 355–362.
- NAGYLAKI, T. 1975. The deterministic behavior of self-incompatibility alleles. *Genetics*, **79**, 545–550.
- NASRALLAH, J. B., KAO, T.-H., CHEN, C.-H., GOLDBERG, M. L. AND NASRALLAH, M. E. 1987. Amino-acid sequence of glycoproteins encoded by three alleles of the S locus of *Brassica oleracea*. *Nature*, **326**, 617–619.
- O'DONNELL, S., LANE, M. D. AND LAWRENCE, M. J. 1993. The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. VI. Estimation of the overlap between the allelic complements of a pair of populations. *Heredity*, **71**, 591–595.
- 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.
- OCKENDON, D. J. 1974. Distribution of self-incompatibility alleles and breeding structure of open-pollinated cultivars of Brussels sprouts. *Heredity*, **33**, 159–171.
- PAXMAN, G. J. 1963. The maximum likelihood estimation of the number of self-sterility alleles in a population. *Genetics*, **48**, 1029–1032.
- ROYO, J., KOWYAMA, Y. AND CLARKE A. E. 1994. Cloning and nucleotide sequence of two S-RNases from *Lycopersicon peruvianum* (L.) Mill. *Plant Physiol.*, **105**, 751–752.
- SABA-EL-LEIL, M. K., RIVARD, S., MORSE, D. AND CAPPADOCIA, M. 1994. The S11 and S13 self-incompatibility alleles in *Solanum chacoense* Bitt. are remarkably similar. *Plant Mol. Biol.*, **24**, 571–583.
- SAITOU, N. AND NEI, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**, 406–425.
- TAYLOR, C. B. AND GREEN, P. J. 1991. Genes with homology to fungal and S-gene RNases are expressed in *Arabidopsis thaliana*. *Pl. Physiol.*, **96**, 980–984.
- UYENOYAMA, M. K. 1991. On the evolution of genetic incompatibility systems. VI. A three-locus modifier model for the origin of gametophytic self-incompatibility. *Genetics*, **128**, 453–469.
- UYENOYAMA, M. K. 1995. A generalized least squares estimate for the origin of sporophytic self-incompatibility. *Genetics*, **139**, 975–992.
- VEKEMANS, X. AND SLATKIN, M. 1994. Gene and allelic genealogies at a gametophytic self-incompatibility locus. *Genetics*, **137**, 1157–1165.
- 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. 1965. The distribution of self-incompatibility alleles in populations. *Evolution*, **18**, 609–619.
- XU, B., MU, J., NEVINS, D. L., GRUN, P. AND KAO, T.-H. 1990. Cloning and sequencing of cDNAs encoding two self-incompatibility associated proteins in *Solanum chacoense*. *Mol. Gen. Genet.*, **224**, 341–346.
- YOKOYAMA, S. AND HETHERINGTON, L. E. 1982. The expected number of self-incompatibility alleles in finite plant populations. *Heredity*, **48**, 299–303.