17 Koerselman, W. and Meuleman, A.F.M. **The vegetation N:P ratio: a new tool to detect the nature of nutrient limitation,** *J. Appl. Ecol.* (in press)

18 De Baar, H.J.W. (1994) **von Liebig's law of the minimum and plankton ecology (1899–1991),** *Prog. Oceanog.* 33, 347–386

19 Gorham, E. (1991) **Biogeochemistry: its origins and development,** *Biogeochemistry* 13, 199–239

20 Vitousek, P.M. and Reiners, W.A. (1991) **Ecological succession and nutrient budgets,** *Curr. Contents* 42, 18

21 Chapin, F.S. *et al.* (1994) **Mechanisms of primary succession following deglaciation at Glacier Bay, Alaska,** *Ecol. Monogr.* 64, 149–175

22 Vitousek, P.M. *et al.* (1993) **Nutrient limitations to plant growth during primary succession in Hawaii Volcanoes National Park,** *Biogeochemistry* 23, 197–215

23 Berendse, F. (1990) **Organic matter accumulation and nitrogen mineralization during secondary succession in heathland ecosystems,** *J. Ecol.* 78, 413–427

24 Koerselman, W., Bakker, S.A. and Blom, M. (1990) **Nitrogen, phosphorus and potassium mass balances for two small fens surrounded by pastures,** *J. Ecol.* 78, 428–442

25 Verhoeven, J.T.A., Kemmers, R.H. and Koerselman, W. (1993) **Nutrient enrichment of freshwater wetlands,** in *Landscape Ecology of a Stressed Environment* (Vos, C.C. and Opdam, P., eds), pp. 33–59, Chapman & Hall

26 Erisman, J.W. (1991) **Acid deposition in the Netherlands,** *RIVM Rep.* (No. 723001002), 1–72

27 Aerts, R., Wallen, B. and Malmer, N. (1992) **Growth-limiting nutrients in Sphagnum-dominated bogs subject to low and high atmospheric nitrogen supply,** *J. Ecol.* 80, 131–140

28 Wassen, M.J., Olde Venterink, H.G.M. and De Swart, E.O.A.M. (1994) **Nutrient concentrations in mire vegetation as a measure of nutrient limitation in mire ecosystems,** *J. Veg. Sci.* 6, 5–16

29 Boyer, M.L.H. and Wheeler, B.D. (1989) **Vegetation patterns in spring-fed calcareous fens: calcite precipitation and constraints on fertility,** *J. Ecol.* 77, 597–609

30 Koerselman, W. (1992) **The nature of nutrient limitation in Dutch dune slacks,** in *Coastal Dunes: Geomorphology, Ecology and Management for Conservation* (Carter, R.W.G., Curtis, T.G.F. and Sheehy-Skeffington, M.J., eds), pp. 189–199, Balkema

Learning from rejection: the evolutionary biology of single-locus incompatibility

Adam D. Richman and Joshua R. Kohn

The self-incompatibility (S-) locus of flowering plants is among the most polymorphic known. PCR methods can now be used to estimate both the number of alleles in natural populations and their sequence diversity. The number of alleles provides an estimate of recent effective population size, thus the S-locus provides a tool for examining how species characteristics affect population size. Sequence relationships among alleles provide another estimate of population size extending millions of years into the past. Relationships between S-alleles and related genes provide a means of dating the age of origin of incompatibility systems and determining which, if any, angiosperm families share incompatibility by homology.

Adam Richman and Joshua Kohn are at the Dept of Biology, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0116, USA.

Species in a wide array of flowering plant families prevent self-fertilization by physiological rejection of self pollen. The genetic basis of rejection of self pollen is best understood in single-locus homomorphic gametophytic and sporophytic self-incompatibility systems. Under gametophytic self-incompatibility (GSI), if the S-allele carried by the pollen grain matches either allele of the ovule parent, the pollen is rejected. Thus, matings fall into three categories: (1) compatible (no alleles shared between maternal and paternal parents); (2) semi-compatible (one allele shared, half of the pollen rejected); and (3) incompatible (two alleles shared, all pollen rejected). In sporophytic self-incompatibility (SSI), the genotype of the pollen parent (rather than that of the haploid pollen grain) determines whether or not pollen will be rejected, and dominance relationships among alleles play a role in determining mating type. Alleles at the self-incompatibility (S-) locus are less frequently rejected when rare, leading to diversifying selection and some of the highest levels of allelic polymorphism known for any locus. Many other self-incompatibility (SI) mechanisms are found in flowering plants, including heteromorphic incompatibility (distyly and tristyly) and homomorphic systems controlled by two or more loci. However, the conceptual simplicity of single-locus SI, coupled with recent advances in understanding its molecular basis, make the S-locus particularly amenable to population and evolutionary genetic study.

Polymorphism at the S-locus

Two empirical findings have been pivotal in generating theoretical inquiry into the dynamics of the S-locus. The first was the discovery that natural populations contain very large numbers of S-alleles. Emerson[1] found 45 different S-alleles in GSI *Oenothera organensis*, an endemic species that he estimated to comprise only 500 individuals, raising the question of how this remarkable diversity is maintained in a small, isolated population. Wright[2] modeled frequency dependent selection (FDS) at a gametophytic S-locus and determined the number of S-alleles maintained in a finite population under the balance of drift, mutation and selection. For *Oe. organensis*, Wright and later workers[3] were forced to conclude that population numbers had recently declined from sizes between 2000 and 10 000. More recent population estimates of *Oe. organensis* place the number of individuals near 5000 (Ref. 4), a remarkable validation of theoretical prediction. Subsequent studies in several taxa have recovered similarly large estimates of the

  PII: S0169-5347(96)10051-3

Box 1. The number of S-alleles and effective population size

Given the number of gametophytic self-incompatibility (GSI) alleles, n, in a sample of r diploid individuals, the number of alleles in the population, N, can be found by iteratively solving Paxman's[35] estimator:

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

This estimator assumes that genotype frequencies are equal, as expected under GSI. The assumption that the frequency distribution of alleles in the sample is drawn from a uniform distribution can be tested using Mantel's[36] test:

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

where C_j is the number of times an allele occurs, n is the number of alleles found, and r is the number of genotypes sampled. Alternative estimators for the number of alleles in a population are available if allele frequencies are not uniform. Substantial deviations have been detected only in *Papaver rhoeas*[37].

The population size required to maintain a specified number of S-alleles is a solution of:

$$n_a = \int_{1/2N_e}^{1} \phi(x)dx$$

where $\phi(x)dx$ is the number of alleles whose frequencies fall in interval from x to $x + dx$, and where:

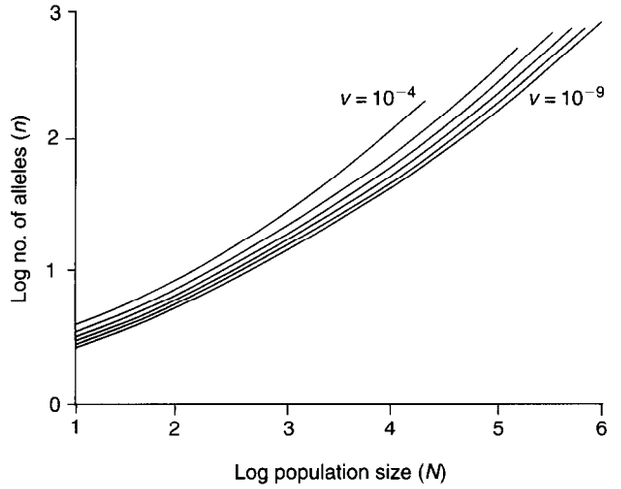
$$\phi(x) = 4N_e v e^{2N_e a x}(1-2x)^{2N_e b-1}x^{-1},\ 0 < x < 0.5,\ a = 1/[(1-J)(1-2J)],\ b = 1/[2(1-J)] + v$$

and J is the effective homozygosity obtained as a solution of:

$$\sqrt{8\pi N_e}\exp\{2N_e J/[(1-J)(1-2J)]\} = (1-J)^{-1/2}(1-2J)^{-N_e[1/(-J)+2v]}$$

where N_e is effective population size, and v is the origination rate.

The relationship between the number of alleles maintained in a population (N) and the effective population size (N_e), for values of the origination rate (v) of new alleles from 10^{-4} to 10^{-9} gene^{-1} generation^{-1}, is shown below. Within the expected range of origination rates ($v = 10^{-7}$–10^{-9}; see text) estimates of N_e are relatively insensitive to the value of v.



number of S-alleles in natural populations of GSI and SSI plants (Box 1, Table 1), and resulting estimates of population size are generally several thousand individuals. Data on S-allele numbers provide the most common genetic estimates of effective population size (N_e) for plants but have gone largely ignored in recent discussions of plant population size[5–7].

Characterization of the stylar gene products of the S-locus in several families led to the second major evolutionary finding. Despite the discovery that the molecular basis of SI systems differs widely[8], all these systems exhibit allelic variation that can be extremely old, often pre-dating the origin of species in which the alleles currently reside. In the Solanaceae, the stylar gene product of the S-locus is an

Table 1. Numbers of S-alleles and corresponding estimates of effective population size (N_e) in natural populations of gametophytic self-incompatible (GSI) and sporophytic self-incompatible (SSI) plants

Species	Number of S-alleles	N_e	Refs
GSI			
Papaver rhoeas	34, 38, 42	>2000	38
Phlox drummondii	45	≥5000	39
Trifolium repens	>36		40
T. pratense	>35		41
Oenothera organensis	45	≥5000	1
Physalis crassifolia	44	6000–10 000	23
Solanum carolinense	14	500–1000	22
SSI[a]			
Brassica campestris	20–30		42
Iberis amara	≥22		43
Raphanus raphanistrum	≥22		44
Sinapis arvensis	52		45

[a]For SSI plants, allele numbers represent minimum estimates, and estimates of N_e are not possible unless dominance relationships among alleles are ignored.

RNase[8], and gene genealogies of S-alleles cloned from *Solanum*, *Petunia* and *Nicotiana* show that alleles from different species and genera frequently group together[9], a pattern called trans-specific evolution (Box 2). These genera are thought to have diverged approximately 30 million years (My) ago (Ref. 9). The age of S-alleles is also reflected in the extremely high levels of sequence divergence (44% amino acid divergence, on average), which is thought to have accumulated in the absence of recombination at the S-locus in the Solanaceae[10,11]. Similar evidence for trans-specific evolution and extensive sequence divergence of S-alleles is found in the SSI family Brassicaceae[12].

A similar pattern of extensive trans-specific evolution was previously observed for the MHC loci of primates and other vertebrates[13,14], motivating studies of the role of balancing selection in maintaining allelic variation. Takahata[15–18] pioneered the application of coalescent theory for the inference of historical population size to loci under balancing selection. In particular, he derived the remarkable result that the genealogy of selectively different lineages is mathematically analogous to that of a neutral gene genealogy, differing only by a constant scaling factor (see Box 3), thus allowing the application of theory developed for the treatment of selectively neutral gene-genealogies. Vekemans and Slatkin[19] extended this approach by deriving the scaling factor specific to FDS at the gametophytic S-locus.

The discovery that S-lineages persist over millions of years places an additional constraint on models of S-allele diversity, which must simultaneously explain large numbers of alleles in populations and long persistence times of allelic lineages. While Wright[2] showed that hypermutability increases the equilibrium number of alleles, high origination rates also increase allelic turnover, reducing the time to coalescence (T_c). Vekemans and Slatkin[19] modeled evolution at a gametophytic S-locus in order to identify combinations of population size and origination rate consistent with both long persistence and large numbers of extant alleles. This model, in combination with empirical data, allows broad bounds to be placed on the origination rate of new alleles.

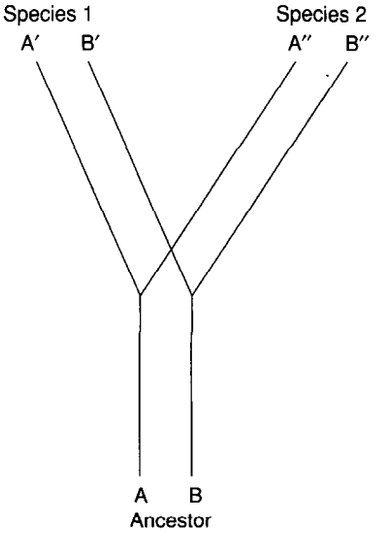
In the Solanaceae, assays of large pollen populations estimated the rate of spontaneous mutations to be $0.2–0.4\times10^{-6}$ gene^{-1} generation^{-1} (Ref. 20). However, all such mutations caused self-compatibility rather than conversion to a new mating type, so the rate of origination of new specificities is expected to be lower than this value. In order for coalescence times of alleles to reach 40 My as implied by the allele genealogy, an origination rate as high as 10^{-6} would require a long-term N_e of 4×10^5 (Ref. 19). At this population size, 375 alleles are expected to be maintained, a value in excess of all observed allele numbers, even for species-wide estimates[21]. At an origination rate of 10^{-7}, a long-term population size of 2×10^4 is required for T_c to equal 40 My and 66 alleles are maintained at equilibrium, a value more in keeping with observed allele numbers. Origination rates $\leq10^{-9}$ gene^{-1} generation^{-1} increase T_c but make it difficult to maintain observed allele numbers in relatively small, isolated populations[19].

Molecular tools for assaying S-diversity

Until recently, data on allele numbers and sequence relationships came from different sources. Labor-intensive diallel crossing studies were used to estimate the number of alleles in populations, their frequencies and the degree of overlap in the alleles found among populations of a species. Sequence data, on the other hand, were limited to relatively few alleles cloned from various species without reference to natural population structure. PCR (polymerase chain reaction)-based methods (Box 4) can now be used to amplify S-alleles from individuals, thus providing rapid population surveys of S-allele numbers as well as sequence information for gene-genealogical studies. Application of models to data on allele numbers and sequence diversity from natural populations provides a tool for examining recent and long-term population dynamics.

Box 2. Trans-specific evolution

Trans-specific evolution occurs when a polymorphism present in an ancestral taxon is transmitted to descendants. In the figure below, alleles A and B were transmitted to extant Species 1 and 2 from their ancestor. Both alleles have accumulated changes through time but a gene-genealogy will group A′ with A″ and B′ with B″, so that sister alleles will occur in different species. Trans-specific evolution is a common feature of loci under balancing selection, such as the S-locus in plants (Fig. 1) and the MHC loci of vertebrates[13,14,46]. In contrast, neutral polymorphism usually coalesces within extant species.



For instance, a PCR-based approach was used to survey S-allele numbers and sequence diversity in natural populations of two solanaceous species[22,23]. In *Solanum*

Box 3. Estimating long-term effective population size

Gene-genealogies from loci under balancing selection can be used to estimate long-term effective population size (N_e). One method uses the number of trans-specific lineages in a sample assuming a stochastic model for lineage turnover. Specifically, we calculate the probability $g_{nk}(t)$ that given a sample of n extant alleles, they coalesce into k lineages t generations ago. It is convenient to replace t by t', defined as:

$$t' = \frac{t}{2N_e f_s} \quad (1)$$

where f_s is a scaling factor specific to frequency dependent selection at the S-locus:

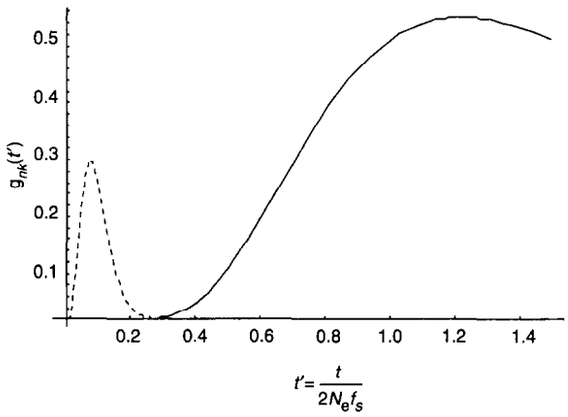
$$f_s = \frac{\sqrt{2}}{16N_e^2 va(J - v/a)^2} \quad (2)$$

where v is the origination rate, and J and a are defined in Box 1.
The probability, $g_{nk}(t')$, is then given by:

$$g_{nk}(t') = \sum_{m=k}^{n} \frac{(2m-1)(-1)^{m-k} k_{(m-1)} n_{[m]}}{k!(m-k)!n_{(m)}} \exp\left\{-\frac{m(m-1)t'}{2}\right\} \quad (3)$$

for $2 \leq k \leq n$, where $n_{[m]} = n(n-1)(n-2)\ldots(n-m+1)$, $n_{(m)} = n(n+1)(n+2)\ldots(n+m-1)$, and m is a variable changing from k to n (Refs 17,19). For a given v, we then calculate the probability $g_{nk}(t')$ over the range of t' to find the value of t' that maximizes the probability of observing n and k.

The phylogeny in Fig. 1 was used to estimate k for *Solanum carolinense* and *Physalis crassifolia*. An S-lineage was considered to be as old or older than the divergence of *Nicotiana* from the clade containing both *Solanum* and *Physalis*, if it inserted into the phylogeny at a node ancestral to an allele from *Nicotiana* or *Petunia*[26]. Nine such ancient lineages are found in *S. carolinense*, while only two occur in the sample of alleles from *P. crassifolia* (see Fig. 1). Likelihood curves for both species (*S. carolinense*, dotted line, $k = 9$, $n = 13$; *P. crassifolia*, solid line, $k = 2$, $n = 28$) are shown in the figure below.



The value of N_e associated with the maximum-likelihood estimate of t' is obtained from eqn (1) above.

Assuming the time since divergence of *Nicotiana* from the clade containing both *Physalis* and *Solanum* is 30 My (Ref. 9), and assuming an average generation time of 2 years, the N_e for origination rates (v) from 10^{-7} to 10^{-9} are presented below. For all origination rates, the long-term N_e for *P. crassifolia* is substantially smaller than for *S. carolinense* in contrast to estimates of current N_e based on the number of alleles (Box 1).

	Effective population size (N_e)	
Origination rate (v)	*S. carolinense*	*P. crassifolia*
10^{-7}	6.5×10^5	0.6×10^4
10^{-8}	2.2×10^4	0.9×10^2
10^{-9}	3.0×10^2	0.6×10^2

carolinense, a weedy rhizomatous plant of the southeastern USA and northern Mexico, sampling of a total of 24 individuals from two populations separated by approximately 250 km revealed a total of (only) 13 alleles, 11 of which were present in both populations providing a species-wide estimate of 13–14 S-alleles. By contrast, allelic diversity in *Physalis crassifolia*, a perennial subshrub of southwestern deserts, was similar to that found in previous studies of GSI plants. Twenty-eight alleles were found in the first 22 individuals sampled, providing a single population estimate of 44 alleles. The striking difference in the numbers of alleles found in these species may be attributable to differences in ecology. The rhizomotous *S. carolinense* occurs in small, short-lived patches on disturbed sites suggesting small N_e. By contrast, *P. crassifolia* lacks the weedy habit occurring in undisturbed habitats of the Mojave and Colorado deserts and does not reproduce vegetatively. It is interesting to note that species of SSI Asteraceae thought to suffer reduced reproduction owing to low mating-type diversity reproduce clonally, a trait shared in common with *S. carolinense*[24,25].

Sequences of S-alleles from *S. carolinense* and *P. crassifolia* indicate surprising differences in the tempo and mode of diversification. Alleles from *S. carolinense* show extensive trans-specific evolution (Fig. 1). The 13 alleles represent nine lineages that pre-date the divergence of *Solanum* and *Nicotiana* (Fig. 1, Box 3). In *P. crassifolia*, on the other hand, there is extensive, relatively recent allelic diversification, with evidence for only two old (transgeneric) lineages (Fig. 1, Box 3). Recent diversification is indicated also by the relatively low level of sequence divergence among alleles within each of the old lineages, resulting in significant differences in average allele age between the samples from the two species[26]. The clustering of *Physalis* alleles is even more striking given molecular phylogenetic information that places *Physalis* and *Solanum* as sister genera relative to the more basal *Nicotiana* and *Petunia*[27,28]. The sequence diversity of alleles found in *Physalis* implies that most extant alleles were derived relatively recently, as, for example, following a population bottleneck, rather than as the result of continuous turnover in the absence of fluctuation in population size (Box 3). Following a bottleneck, the rate of successful invasion by mutant alleles with new specificities substantially increases, leading to rapid diversification within the extant lineages[17]. The striking differences in both allele number and age among the first species surveyed indicates that the S-locus may provide a wealth of information pertaining to current and to historical population processes.

Origins of the S-locus

Molecular analysis of S-gene products in GSI Solanaceae and Papaveraceae, as well as SSI Brassicaceae, have revealed three different molecular mechanisms of rejection and, by extension, independent origins of SI in these families or their ancestors[8]. In contrast, the stylar S-gene products in GSI Rosaceae and Scrophulariaceae have recently been cloned and found to be RNases as in the Solanaceae[29–31]. Whether this represents homology of the SI systems of these dicot families is unclear. Some phylogenetically intervening families (e.g. Brassicaceae) have evolved independent SI mechanisms but the possibility remains that gametophytic SI could be homologous in a number of higher dicot families. Equally intriguing is the possibility that the use of RNases as incompatibility genes in different plant families represents convergence. It has been suggested that RNases functioning in defense against pathogen invasion through floral organs may have been converted to function in self-pollen rejection[32]. Identification of RNases involved in defense may lead to discovery of the antecedents of the S-locus. Whether or not S-RNases in different angiosperm families are homologous, the PCR methods used to survey S-alleles in the Solanaceae will probably be useful in any plant family with an RNase-based incompatibility system.

The fact that related non-S-genes are present in various plant families suggests an approach to dating the origin of SI by estimating the time taken since the change of the non-S-genes from their ancestral function to their use as an SI mechanism. What is needed is a measure of the rate of substitution along branches of the gene-family tree, and a divergence time from the fossil record, or another source, with which to calibrate the clock. Uyenoyama[33] used sequence data on the *SLG/SRK/SLR* multigene family in the Brassicaceae to estimate the time since the origin of SI. This estimate was 4–5 times as long as the time since divergence of the *Brassica oleracea* and *B. campestris*, an event assumed to have occurred 10 My ago.

In addition to estimating the age of the SI system found in the Brassicaceae, Uyenoyama's method estimated rates of substitution in S-alleles relative to those of related genes not subject to frequency-dependent selection. Rates of substitution at the S-locus of Brassicaceae were found to be significantly slower than those of related genes. This result was important since hypermutability had been suggested as a factor in the maintenance of extreme polymorphism[34]. Instead, the large numbers of alleles and their high levels of sequence divergence appeared to be brought about through the long persistence of allelic lineages. In fact, it appeared that most S-lineages appeared soon after the origin of the system, and subsequently, the origination rate of allelic lineages slowed down. To account for this

Box 4. Molecular methods to assay S-allele variation

Brace *et al.*[48] were the first to describe a PCR (polymerase chain reaction) method for amplification of S-alleles using sporophytic self-incompatible (SSI) *Brassica oleracea*. PCR primers were constructed using alignments of known sequences, and amplification products were digested with a battery of restriction enzymes to identify different alleles. Working with a reference population of 48 lines that were homozygous for different S-alleles, Brace *et al.*[48] found only two pairs of lines in which restriction fragment patterns of amplification products were identical. Lines with identical patterns were found to be incompatible, representing duplications in the reference collection, an unexpected confirmation of the reliability of the method.

Taking advantage of the abundance of S-locus mRNA in stylar tissue, S-alleles were amplified from two species of gametophytic self-incompatible (GSI) Solanaceae, *Solanum carolinense* and *Physalis crassifolia*, using reverse-transcriptase (RT-) PCR[22,23]. Primer pairs, designed to complement the conserved regions of published S-allele sequences from solanaceous plants, amplify a fragment of expected size from each individual of both species. Restriction digests reveal the presence of two alleles in each amplification product, which is consistent with the expectation that all individuals are heterozygous under GSI. Two S-allele sequences can be cloned and sequenced from each amplification product. In crosses predicted to be semi-compatible, based upon the RT-PCR assay of parental individuals, only the compatible paternal allele is transmitted to all offspring.

PCR approaches circumvent the necessity of large crossing experiments to determine the number of mating type alleles in a population, while providing sequence information for evolutionary study at and above the species level. Other advantages of molecular assays include unambiguous comparisons of alleles found in studies of S-diversity carried out at different times and in different species, and the ability to study plants that are not amenable to greenhouse crossing experiments.

pattern, Uyenoyama speculates that descendant alleles may preferentially replace their parents. Alternatively, the slowdown in the appearance of new S-alleles might reflect constraints on the nature of changes that can generate new specificities. Given the recent data from *Physalis*, it would appear that some lineages, at least, can produce more than 20 different specificities. Further analysis of the limits of within-lineage diversification will require the results of molecular studies of additional species and populations.

Conclusion

Empirical data from the S-locus have provided impetus for the development of population genetic theory several times during the 20th century. Emerson's[1] discovery of extreme polymorphism prompted Wright[2] and others[3] to model allele numbers and sampling distributions under frequency dependent selection. The discovery of extensive trans-specific evolution at the MHC loci of vertebrates and the S-locus of plants has helped to promote extensions of coalescence theory to loci under overdominant and frequency dependent selection[15,19]. Identification of the genes responsible for rejection specificity permits the application of molecular methods to the study of natural populations. We have only begun to accumulate samples from nature. The striking differences in both allele number and age among samples from *S. carolinense* and *P. crassifolia*[26] suggest that a wealth of information concerning ecology and history may be stored at the S-locus. In particular, the S-locus will provide a useful tool for investigation of the relationship of species ecology to effective population size. The prevalence of historical bottlenecks at the S-locus provides a means for evaluating the frequency of founder-event speciation.

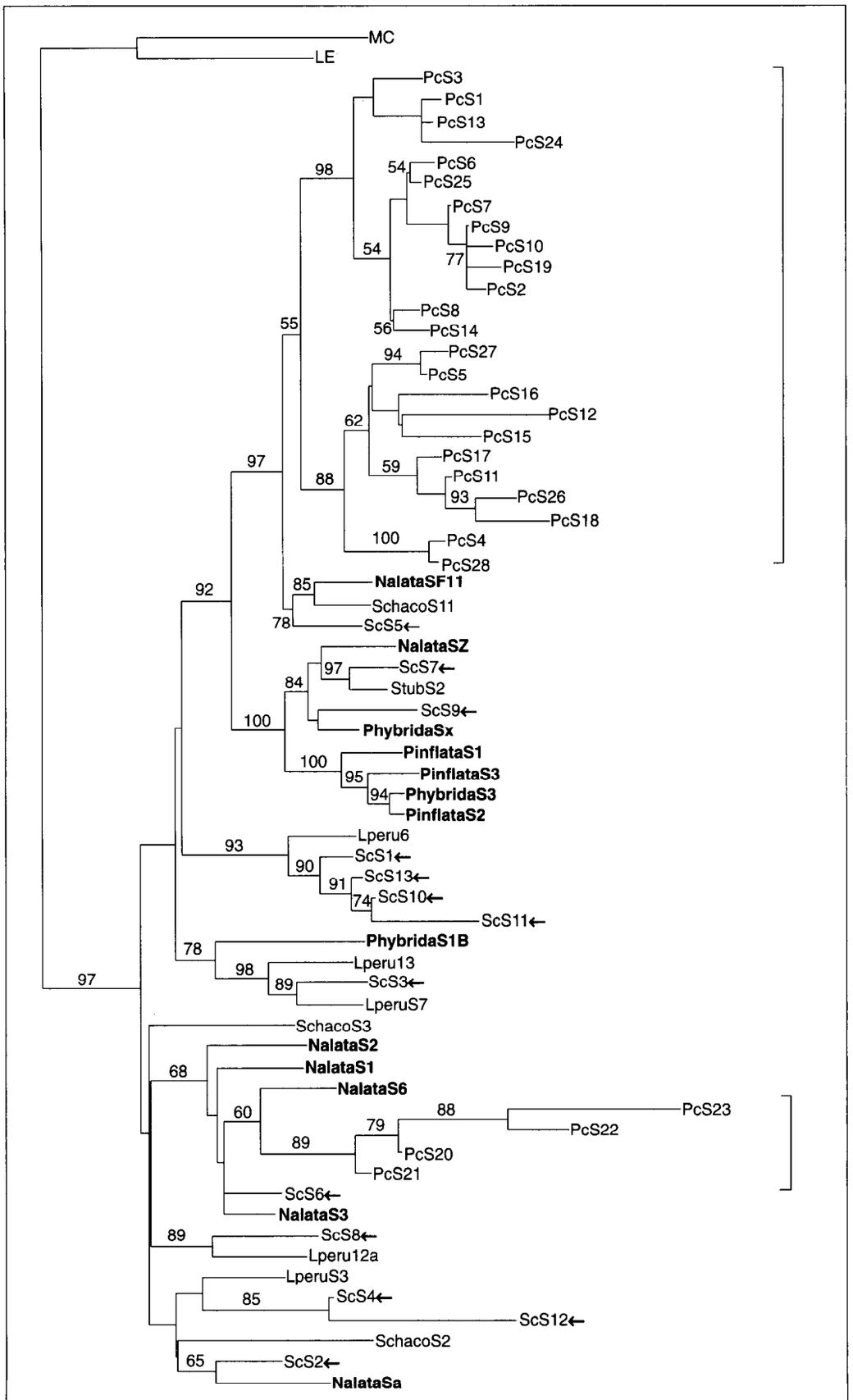


Fig. 1. The phylogeny of 63 (partial) S-allele sequences from the Solanaceae. The neighbor-joining tree was derived from amino acid distances using the PAM 001 Dayhoff similarity matrix[47]. Numbers indicate consensus bootstrap values that exceed 50%. Brackets indicate alleles from *Physalis crassifolia* (Pc)[23], and arrows indicate alleles from *Solanum carolinense* (Sc)[22]. Other solanaceous taxa are *S. chacoense* (Schaco), *S. tuberosa* (Stub), *Lycopersicon peruvianum* (Lperu), *Petunia hybrida* (**Phybrida**), *P. inflata* (**Pinflata**) and *Nicotiana alata* (**Nalata**). The tree is rooted with non-S RNases from *Lycopersicon esculentum* (LE) and *Momordica charantia* (MC; see Ref. 26 for sequence sources). Taxa in bold represent basal genera

Acknowledgements

The support of NSF to both authors made this work possible. We are grateful to M.K. Uyenoyama and two anonymous reviewers for comments on the manuscript.

References

1 Emerson, S. (1939) **A preliminary survey of the *Oenothera organensis* population,** *Evolution* 24, 524–537
2 Wright, S. (1960) **On the number of self-incompatibility alleles maintained in equilibrium by a given mutation rate in the population of a given size: A re-examination,** *Biometrics* 16, 61–85
3 Yokoyama, S. and Hetherington, L.E. (1982) **The expected number of self-incompatibility alleles in finite plant populations,** *Heredity* 48, 299–303
4 Levin, D.A., Ritter, K. and Ellstrand, N.C. (1979) **Protein polymorphism in the narrow endemic *Oenothera organensis*,** *Evolution* 33, 534–542
5 Husband, B.C. and Barrett, S.C.H. (1992) **Effective population size and genetic drift in tristylous *Eichhornia paniculata* (Pontederiaceae),** *Evolution* 46, 1875–1990
6 Frankham, R. (1995) **Effective population size/adult population size ratios in wildlife: a review,** *Genet. Res.* 66, 95–107
7 Schoen, D.J. and Brown, A.H.D. (1991) **Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants,** *Proc. Natl. Acad. Sci. U. S. A.* 88, 4494–4497
8 Matton, D.P. *et al.* (1994) **Self-incompatibility: How plants avoid illegitimate offspring,** *Proc. Natl. Acad. Sci. U. S. A.* 91, 1927–1997
9 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
10 Kheyr-Pour, A. *et al.* (1990) **Sequence diversity of pistil S-proteins associated with gametophytic self-incompatibility in *Nicotiana alata*,** *Sex. Plant Reprod.* 3, 88–97
11 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
12 Dwyer, K.G. *et al.* (1991) **DNA sequences of self-incompatibility genes from *Brassica campestris* and *B. oleracea*: Polymorphism predating speciation,** *Plant Mol. Biol.* 16, 481–486
13 Klein, D. *et al.* (1993) **Extensive MHC variability in cichlid fishes of Lake Malawi,** *Nature* 364, 330–334
14 Klein, J. *et al.* (1993) **Trans-specific *Mhc* polymorphism and the origin of species in primates,** *J. Med. Primatol.* 22, 57–64
15 Takahata, N. (1990) **A simple genealogical structure of strongly balanced allelic lines and trans-species evolution of polymorphism,** *Proc. Natl. Acad. Sci. U. S. A.* 87, 2419–2423
16 Takahata, N. and Nei, M. (1990) **Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci,** *Genetics* 124, 967–978
17 Takahata, N. (1993) **Evolutionary genetics of human paleo-populations,** in *Mechanisms of Molecular Evolution* (Takahata, N. and Clark, A.G., eds), pp. 1–21, Sinauer
18 Takahata, N. (1993) **Allelic geneology and human evolution,** *Mol. Biol. Evol.* 10, 2–22
19 Vekemans, X. and Slatkin, M. (1994) **Gene and allelic genealogies at a gametophytic self-incompatibility locus,** *Genetics* 137, 1157–1165
20 Lewis, D. (1948) **Mutation of the incompatibility gene. I. Spontaneous mutation rate,** *Heredity* 2, 219–236
21 Lane, M.D. and Lawrence, M.J. (1994) **The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. VII. The number of S-alleles in the species,** *Heredity* 71, 596–602
22 Richman, A.D. *et al.* (1995) **S-allele sequence diversity in natural populations of *Solanum carolinense* (Horsenettle),** *Heredity* 75, 405–415
23 Richman, A.D., Uyenoyama, M.K. and Kohn, J. (1996) **S-allele diversity in a natural population of ground cherry *Physalis crassifolia* (Solanaceae) (ground cherry) assessed by RT-PCR,** *Heredity* 76, 497–505
24 Reinartz, J.A. and Les, D.H. (1994) **Bottleneck-induced dissolution of self-incompatibility and breeding system consequences in *Aster furcatus* (Asteraceae),** *Am. J. Bot.* 81, 446–455
25 DeMauro, M.M. (1993) **Relationship of breeding system to rarity in the lakeside daisy (*Hymenoxys acaulis* var. *glabra*),** *Conserv. Biol.* 7, 542–550
26 Richman, A.D., Uyenoyama, M.K. and Kohn, J.R. (1996) **Contrasting patterns of allelic diversity and gene genealogy at the self-incompatibility locus in two species of Solanaceae,** *Science* 273, 1212–1216
27 Olmstead, R.G. and Palmer, J.D. (1992) **A chloroplast DNA phylogeny of the Solanaceae: subfamilial relationships and character evolution,** *Ann. MO Bot. Gard.* 79, 346–360
28 Olmstead, R.G. and Sweere, J.A. (1994) **Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in the Solanaceae,** *Syst. Biol.* 43, 467–481
29 Broothaerts, W. *et al.* (1995) **cDNA cloning and molecular analysis of two self-incompatibility alleles from apple,** *Plant Mol. Biol.* 27, 499–511
30 Janssens, G.A. *et al.* (1995) **A molecular method for S-allele identification in apple based on allele-specific PCR,** *Theoret. Appl. Genet.* 91, 691–698
31 Xue, Y. *et al.* (1996) **Origin of allelic diversity in *Antirrhinum* S locus RNases,** *Plant Cell* 8, 805–814
32 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
33 Uyenoyama, M.K. (1995) **A generalized least squares estimate for the origin of sporophytic self-incompatibility,** *Genetics* 139, 975–992
34 Trick, M. and Heizmann, P. (1992) **Sporophytic self-incompatibility systems: Brassica S gene family,** *Int. Rev. Cytol.* 140, 485–524
35 Paxman, G.J. (1963) **The maximum likelihood estimation of the number of self-sterility alleles in a population,** *Genetics* 48, 1029–1032
36 Mantel, N. (1974) **Approaches to a health research occupancy problem,** *Biometrics* 30, 355–362
37 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
38 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
39 Levin, D.A. (1993) ***S*-gene polymorphism in *Phlox drummondii*,** *Heredity* 71, 193–198
40 Atwood, S.S. (1944) **Oppositional alleles in natural populations of *Trifolium repens*,** *Genetics* 29, 428–435
41 Williams, R.D. and Williams, W. (1947) **Genetics of red clover (*Trifolium pratense* L.) compatibility. III. The frequency of incompatibility *S* alleles in two non-pedigree populations of red clover,** *J. Genet.* 48, 69–79
42 Nou, I.S. *et al.* (1991) **Variation of *S*-alleles and *S*-glycoproteins in a naturalized population of self-incompatible *Brassica campestris* L.,** *Jpn J. Genet.* 66, 227–239
43 Bateman, A.J. (1954) **Self-incompatibility systems in angiosperms. II. *Iberis amara*,** *Heredity* 8, 305–332
44 Karron, J.D., Marshall, D.L. and Oliveras, D.M. (1990) **Numbers of sporophytic self-incompatibility alleles in populations of wild radish,** *Theoret. Appl. Genet.* 79, 457–460
45 Stevens, J.P. and Kay, Q.O.N. (1989) **The number, dominance relationships and frequencies of self-incompatibility alleles in a natural population of *Sinapis arvenis* L. in South Wales,** *Heredity* 62, 199–205
46 Ayala, F.J. (1995) **The myth of Eve: molecular biology and human origins,** *Science* 270, 1930–1936
47 Felsenstein, J. (1993) *PHYLIP (Phylogeny Inference Package) (Version 3.5c)*, J. Felsenstein, University of Washington
48 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.* 7, 169–176