Heredity (2006) 96, 434–444

www.nature.com/hdy

Gametophytic self-incompatibility in *Lycium parishii* (Solanaceae): allelic diversity, genealogical structure, and patterns of molecular evolution at the *S-RNase* locus

AE Savage[1,2] and JS Miller[1]

[1]*Department of Biology, Amherst College, McGuire Life Sciences Building, Amherst, MA 01002, USA;* [2]*Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853, USA*

We characterized allelic diversity at the locus controlling self-incompatibility (SI) for a population of *Lycium parishii* (Solanaceae) from Organ Pipe National Monument, Arizona. Twenty-four partial sequences of S-RNase alleles were recovered from 25 individuals. Estimates of allelic diversity range from 23 to 27 alleles and, consistent with expectations for SI, individuals are heterozygous. We compare S-RNase diversity, patterns of molecular evolution, and the genealogical structure of alleles from *L. parishii* to a previously studied population of its congener *L. andersonii*. Gametophytic SI is well characterized for Solanaceae and although balancing selection is hypothesized to be responsible for high levels of allelic divergence, the pattern of selection varies depending on the portion of the gene considered. Site-specific models investigating patterns of selection for *L. parishii* and *L. andersonii* indicate that positive selection occurs in those regions of the *S-RNase* gene hypothesized as important to the recognition response, whereas positive selection was not detected for any position within regions previously characterized as conserved. A 10-species genealogy including S-RNases from a pair of congeners from each of five genera in Solanaceae reveals extensive transgeneric evolution of *L. parishii* S-RNases. Further, within *Lycium*, the D_n/D_s ratios for pairs of closely related alleles for intraspecific versus interspecific comparisons were not significantly different, suggesting that the S-RNase diversity recovered in these two species was present prior to the speciation event separating them. Despite this, two S-RNases from *L. parishii* are identical to two previously reported alleles for *L. andersonii*, suggesting gene flow between these species. *Heredity* (2006) **96,** 434–444. doi:10.1038/sj.hdy.6800818; published online 19 April 2006

Keywords: balancing selection; gametophytic self-incompatibility; *Lycium parishii*; Solanaceae; S-RNase; transgeneric evolution

Introduction

Genetically controlled self-incompatibility (SI) systems are widespread among angiosperm species and have evolved repeatedly, presumably as a mechanism to avoid self-fertilization and enforce outcrossing. As plants are sessile, often cosexual (ie, possessing both male and female gender function), and typically produce multiple flowers within individuals, SI systems have the potential to govern both individual and population patterns of mating. In addition, SI systems can affect the evolution of other traits associated with the mating system (eg, inbreeding depression: Jain, 1976; Charlesworth and Charlesworth, 1987; gender expression: Charlesworth, 1985; Miller and Venable, 2000).

Among homomorphic SI systems, gametophytic self-incompatibility (GSI) is widespread among angiosperm lineages (Kao and McCubbin, 1996). Two genes control the incompatibility reaction in Solanaceae, one expressed in the pistil and another in the pollen. Molecular studies have shown that the pistil S-gene product is a glycoprotein (Kehyr-Pour and Pernes, 1985; Anderson *et al*, 1986) with ribonuclease (RNase) activity (McClure *et al*, 1990). More recently, the pollen determinant (*PiSLF*) was confirmed in *Petunia inflata* and, like the pistil S-RNase, *PiSLF* shows high levels of amino acid (AA) diversity for the alleles examined to date (Sijacic *et al*, 2004). When the haploid S-genotype of the pollen grain matches that of either of the two S-RNases expressed in the pistil of the maternal plant, pollen tube growth is terminated. At the level of individuals, such genetic control enforces heterozygosity. At the level of populations, strong negative frequency dependence shelters alleles from extinction and creates strong selection for novel alleles; both these factors result in the maintenance of large numbers of S-alleles within populations (Wright, 1939).

GSI is well studied in Solanaceae and has been characterized for natural populations of species from several genera (eg, *Lycium*, Richman, 2000; *Petunia*, Wang *et al*, 2001; *Physalis*, Richman *et al*, 1996, Richman and Kohn, 1999; *Solanum*, Richman *et al*, 1995; *Witheringia*, Richman and Kohn, 2000, Stone and Pierce, 2005). Data from these species generally meet expectations; individuals are heterozygous and populations are highly

Correspondence: JS Miller, Department of Biology, Amherst College, McGuire Life Sciences Building, Amherst, MA 01002, USA.
E-mail: jsmiller@amherst.edu
Received 15 June 2005; accepted 28 February 2006; published online 19 April 2006

polymorphic at the *S-RNase* locus (see reviews by Richman and Kohn, 1996; Lawrence, 2000; Castric and Vekemans, 2004). In addition, S-RNases are often highly divergent from one another (Richman and Kohn, 1996), and genealogies indicate that S-RNases are ancient, often grouping more closely with alleles from other species and genera (Ioerger *et al*, 1990; Richman and Kohn, 2000). This latter phenomenon is known as transpecific (TS) or transgeneric (TG) evolution, where polymorphisms are passed down from a common ancestor to multiple descendant species and even genera (Ioerger *et al*, 1990).

TG evolution is extensive among S-RNases sampled across several genera in Solanaceae (Ioerger *et al*, 1990), and patterns of TG evolution have been used to infer historical events in the evolutionary history of various Solanaceae. Whereas alleles from *Lycium andersonii* and *Solanum carolinense* showed extensive TG evolution (12 and eight TG lineages, respectively), alleles from two species of *Physalis* and *Witheringia maculata* were confined to three TG lineages (Richman and Kohn, 2000). The authors suggest that a genetic bottleneck in the common ancestor of *Physalis* and *Witheringia* resulted in the loss of entire S-RNase lineages. Such a bottleneck is consistent with the observed pattern of both the shared identity and reduced number of TG lineages in *Physalis* and *Witheringia*, as well as current estimates of species phylogenies in Solanaceae (Olmstead *et al*, 1999).

In addition to variation in the number of TS or TG lineages across species and genera, S-RNase allele number also varies. Despite equal numbers of TS lineages between two species of *Witheringia*, estimates of S-RNase diversity were over twice as high for *W. solanacea* compared to *W. maculata* (Stone and Pierce, 2005). Similarly, Richman and Kohn's (1999) comparison of *Physalis cinerascens* and *Physalis crassifolia*, showed variation in allele number but not allelic age, as measured by the number of TG lineages. Richman and Kohn (2000) point out that variation in TG lineages is expected to evolve more slowly (compared to allele number) since every member-allele must be lost before the extinction of a TG lineage. In contrast, the number of alleles in a population may be more sensitive to recent changes in species ecology. These authors suggest that a major determinant of allele number in populations is a result of differences among species in life-history characteristics that affect both population size and longevity. Several studies have considered the relationship between S-RNase diversity and life-history characteristics (Richman and Kohn, 1999, 2000; Lu, 2001; Stone and Pierce, 2005); however, no clear association has emerged in Solanaceae. Further sampling of species closely related to those previously studied can strengthen overall patterns of TG evolution in Solanaceae and highlight differences between congeners.

Patterns of molecular evolution (ie, D_n/D_s, the ratio of nonsynonymous substitutions per nonsynonymous site to synonymous substitutions per synonymous site) are often used to measure positive selection at the *S-RNase* gene in species with GSI. Substitution rates can be calculated over the entire region; however, given the conserved and hypervariable regions described by Ioerger *et al* (1991), substitution patterns may vary between regions. Structural analysis of a S-RNase from *Nicotiana alata* (Ida *et al*, 2001) indicates that the hypervariable regions are located on the surface of the protein, thus corroborating previous suggestions that these regions are involved in the recognition mechanism (Ioerger *et al*, 1991). Genealogical relationships among alleles can also complicate patterns of molecular evolution. Specifically, distantly related alleles are expected to accumulate neutral mutations over time, and these substitutions may obscure patterns of positive selection (Richman *et al*, 1996). To take this into account, Richman *et al* (1996) characterized D_n/D_s ratios for closely related alleles in *Physalis crassifola* and *S. carolinense*; in both species, D_n/D_s was greater than 1.0 when comparing alleles that were relatively similar, but <1.0 for comparisons of more divergent alleles (see also Lu, 2002).

TG and TS lineages are useful for examining deep genealogical relationships among alleles (eg, ascertaining bottleneck events, Richman and Kohn, 1999, 2000), but may confound more recent evolutionary patterns such as the timing of diversification of alleles. The D_n/D_s ratio can also be used to compare pre- versus post-speciation diversification of S-alleles for closely related species. If S-RNase diversification largely precedes speciation, the D_n/D_s ratio should be similar for allelic comparisons within species and between closely related species. In contrast, if diversification is predominately post-speciation, D_n/D_s is expected to be higher in intraspecific comparisons (compared to interspecific allele pairs), since replacement substitutions among alleles within a single species would be contributing to allelic diversity. More extensive sampling from closely related species within genera are needed to detect species-specific patterns of S-RNase diversification.

Miller and Venable (2002) demonstrated that fruit production, seed set, and pollen tube number in a natural population of hermaphroditic *L. parishii* was significantly higher following outcross pollination compared to self-pollination. Based on these crossing studies, these authors suggested that *L. parishii* was self-incompatible. Miller and Venable (2000, 2002) also argued that, for some species of *Lycium*, GSI had been rendered nonfunctional as a result of polyploidy, eventually leading to the evolution of gender dimorphism (ie, separate females and hermaphrodites). An important assumption of this hypothesis is that GSI is ancestral in *Lycium*; thus, confirming GSI in the hermaphroditic close relatives (eg, *L. parishii*; Levin and Miller, 2005) of the dimorphic species is an important first step in evaluating this hypothesis.

Here, we analyze partial S-RNase sequences from *L. parishii* and investigate SI in this species by genotyping 25 individuals from a population in southern Arizona. Specifically, we: (1) investigate the molecular signature of SI (ie, heterozygosity of individuals and allelic diversity within populations), (2) explore molecular evolution of the *S-RNase* gene using D_n/D_s ratios calculated for different regions of the gene and for site-specific models of substitution patterns, (3) compare S-RNase diversity and patterns of allele diversification in *L. parishii* to previous work on its congener *L. andersonii* (Richman, 2000), and (4) explore the genealogy (including the number of TS and TG lineages) of Solanaceae S-RNases using species pairs for five genera in the family.

Methods

Study species and population sampling

Lycium L. (Solanaceae) is a genus of approximately 80 species distributed worldwide, but particularly concen-

trated in South America, southwestern North America and southern Africa (Hitchcock, 1932; Chiang-Cabrera, 1981; Bernardello, 1986; Venter, 2000; Miller, 2002; Levin and Miller, 2005). *Lycium* species are long-lived perennial shrubs and many inhabit arid to semiarid environments, though some are found in coastal saline habitats. The majority of species are hermaphroditic, have perfect flowers, and produce red, fleshy berries. The study population of *L. parishii* was from two adjacent sampling sites in Organ Pipe National Monument (OPNM) located along the US-Mexican border along South Puerto Blanco Drive (31°54′00″ N, 112°51′00″ W, Pima County, AZ, USA). *L. parishii* is common throughout OPNM, and plants in this population were fairly evenly spaced, although plants were larger and found at higher densities along washes. Ten to twenty styles from mature floral buds were dissected from flowers on each plant and frozen immediately on dry ice. Styles were collected from nine plants in March 2001 and an additional 16 plants in March 2004. Individual plants selected for sampling were a minimum of 5 m apart, although it is possible that what appeared to be a single individual may have been more than one genetic individual, especially where plants were large.

RT-PCR, cloning, RFLP analysis, and S-RNase sequencing

Three to 12 styles from each individual were ground in liquid nitrogen and stylar mRNA was obtained using Qiagen's RNeasy Plant Mini Kit (Valencia, CA, USA). We synthesized cDNA using the Novagen First Strand cDNA synthesis kit (Darmstadt, Germany) and amplified a portion of the *S-RNase* gene between conserved regions C2 and C5 (Ioerger *et al*, 1991) using degenerate primers PR1 (5′-GAA TTC A(TC)G GN(TC) TNT GGC CNG A-3′) and PR3 (5′-NNG TCG AC(GA) AAA CAT ATN CCT A(TC)(TC) TCN (ATC)NT AG(TC) TC-3′) from Richman *et al* (1995). We ran 40 cycles (94°C, 15 s; 45°C, 60 s; 72°C, 60 s) on a GeneAmp PCR system 2700 (Applied Biosystems, Foster City, CA, USA). Amplification products were cloned into the pT7Blue vector using the Novagen Perfectly Blunt Cloning kit (Darmstadt, Germany) to separate allele copies. Individual colonies were isolated with sterile pipette tips and amplified using the PCR primers and conditions described above. Inserts can be incorporated into the vector in either direction (5′–3′ or vice versa), thus primers PR1 and PR3 were used for this amplification to avoid incorrect RFLP interpretations. The colony PCR products were screened for polymorphisms using RFLP analysis with a battery of restriction enzymes including *Taq*I, *Hha*I, *Hae*III, and *Dde*I. Approximately 20 colonies were selected from each individual for RFLP analysis. All colonies identified as having unique RFLP banding patterns were amplified with the vector primers U19 (5′-GTT TTC CCA GTC ACG ACG T-3′) and R20 (5′-CAG CTA TGA CCA TGA TTA CG-3′) and sequenced on an Applied Biosystems Automated 3730 DNA Analyzer at Cornell University (Ithaca, NY, USA). Nine of the RFLP patterns recovered were sequenced multiple (range, 2–13) times to confirm the reliability of RFLP analysis to identify *S*-alleles. For all of the 49 sequenced colonies corresponding to the nine RFLP patterns, sequences with the same RFLP pattern had identical nucleotide sequences.

L. parishii S-RNase diversity

The number of S-alleles for a diploid, self-incompatible population was estimated from the number of alleles found in a sample of that population by iteratively solving Paxman's (1963) estimator. Given n GSI alleles in a sample of r diploid individuals, the number of alleles in the population, N, is given by Equation (1).

$$n = N\left[1 - \left(1 - \frac{2}{N}\right)^{r}\right] \quad (1)$$

This equation assumes that individuals are under balancing selection and that alleles exist in equal frequencies within the population. To examine the assumption that the frequency distribution of alleles in the sample were drawn from a uniform distribution, Mantel's (1974) test was used (2), where C_j is the number of times an allele occurs, n is the number of alleles found and r is the number of diploid individuals sampled.

$$\chi^2 = (n-1)\left(\sum C_j^2 - \frac{4r^2}{n}\right)\left(2r - \frac{4r}{n}\right)^{-1} \quad (2)$$

Molecular evolution, allele diversification, and terminal branch lengths

L. parishii S-RNase sequences were aligned in Clustal X (Thompson *et al*, 1994) using default settings and adjusted by eye following Ioerger *et al* (1991). Partial S-RNase sequences from *L. andersonii* that spanned regions C2–C5 from Ioerger *et al* (1991) were included in the alignment (Richman, 2000; GenBank accession numbers AF05343–4, AF105347–9, AF105353, AF105355, AF105358–9, and AF105362–3).

Nonsynonymous and synonymous substitution rates were measured in MEGA3 (Kumar *et al*, 2004) using the modified Nei-Gojobori model with Jukes–Cantor correction for the calculated transition/transversion ratio ($R = 0.75$) among *L. parishii* S-RNases. Values of D_n and D_s and their ratio were calculated for the entire S-RNase sequence, as well as for the conserved (C3 and C4) and hypervariable (HVa and HVb) regions separately. Nucleotide substitution patterns were also assessed using site-specific codon models in Phylogenetic Analysis using Maximum Likelihood (PAML, version 3.13d; Yang, 2002). Likelihood ratio tests were used to compare models of neutral (M1) and positive (M2) selection for both *L. parishii* and *L. andersonii*.

To assess patterns of allelic diversification, we identified pairs of closely related alleles using the maximum likelihood (ML) topology (see below in Methods and Figure 3 in Results) and calculated the D_n, D_s, and D_n/D_s ratios for each pair of alleles. This resulted in 14 intraspecific allele pairs within *L. parishii* (S01, S03; S01, S02; S02, S03; S08, S18; S10, S13; S10, S11; S11, S13; S04, S12; S06, S07; S06, S23; S07, S23; S21, S22; S21, S25; S22, S25), five intraspecific allele pairs within *L. andersonii* (S07, S17; S06, S21; S07, S20; S17, S20; S05, S16), and 14 interspecific allele pairs between *L. parishii* and *L. andersonii* (P-S24, A-S13; P-S09, A-S06; P-S09, A-S21; P-S18, A-S11; P-S10, A-S01; P-S11, A-S01; P-S06, A-S16; P-S07, A-S16; P-S23, A-S16; P-S21, A-S05; P-S22, A-S05; P-S25, A-S05). We used non–parametric Mann–Whitney *U*-tests to determine if D_n, D_s, or D_n/D_s for intraspecific pairs of alleles differed from interspecific allele pairs.

We measured terminal branch lengths in *L. parishii* to facilitate comparisons with previous studies in *Lycium* (Richman and Kohn, 2000) and Solanaceae (Richman and Kohn, 2000; Stone and Pierce, 2005). After verifying that the *L. parishii* sequences did not significantly deviate from clocklike behavior using a two-cluster test in LINTREE (Takezaki *et al*, 1995), terminal branch lengths were measured using KITSCH (Felsenstein, 2005). Relative terminal branch lengths were calculated following Uyenoyama (1997) using equation (3), where S is the sum of terminal branch lengths, D is the time since coalescence of all alleles (total depth of the genealogy), and n is the number of alleles in the sample:

$$R_{SD} = S\left(1 - \frac{1}{n}\right) \Big/ D \qquad (3)$$

Genealogical analysis of Solanaceae S-RNases

Initially we constructed a genealogy using the neighbor-joining algorithm (Saitou and Nei, 1987) including all *L. parishii* S-RNases reported here, those mentioned above for *L. andersonii*, *Petunia axillaris* (AF239907–10, AY180048–50), *Petunia integrifolia* (AF301167–80), *Physalis cinerascens* (AF058930–41), *P. crassifolia* (L46653, L46655–63, L46665–6, L46668–9, L46672–3, L46677, L46679–80), *Solanum carolinense* (L40539–48, L40551), *Solanum chacoense* (AF191732, L36666–7, S69589, X56896–7), *Witheringia. maculata* (AF102065–75), and *W. solanacea* (AY454099–107, AY454109–21). The S-like RNase S2 from *Antirrhinum hispanicum* (X96465) was used as the outgroup.

From this genealogy (not shown), one allele representing each TG lineage from each species of *Petunia*, *Physalis*, *Solanum*, and *Witheringia* was selected in order to preserve TG relationships across alleles from these genera plus *Lycium*. The selected alleles, as well as all alleles from *L. parishii* and *L. andersonii* were included in a ML analysis containing 90 alleles sampled across these five genera within Solanaceae. The ML model parameters were determined using the Akaike information criterion in Modeltest version 3.6 (Posada and Crandall, 1998). The best-fit model (GTR + I + G) was used in a ML analysis in PAUP* (Swofford, 2002) using the heuristic search option, TBR branch swapping, MulTrees option in effect, and a single neighbor-joining tree as a starting topology.

We also generated a Bayesian consensus tree to test for differences in branching relationships compared to the ML tree. We included the entire 138-allele data set of *Lycium*, *Petunia*, *Physalis*, *Solanum*, and *Witheringia* S-alleles in the Bayesian tree to determine if any TG lineages were lost by reducing allelic sampling in the ML analysis. Bayesian analysis was run in MrBayes version 3.0b4 (Ronquist and Huelsenbeck, 2003) using four simultaneous Markov chain Monte Carlo chains each starting from a random tree, with a general time reversible substitution model and gamma-distributed rate variation across sites. Two million generations were run with a tree saved every 100 generations, and trees proceeding the stabilization of likelihood values were excluded in the construction of a 50% majority rule consensus tree in PAUP*. The number of TG lineages was determined for *L. andersonii* and *L. parishii* for both the ML and Bayesian analyses; we define a TG lineage as the most recent node including alleles from more than one genus.

Results

L. parishii S-RNase diversity

The 25 sampled *L. parishii* plants contained 24 unique S-RNase sequences ranging in length from 369 to 390 bp (GenBank accession numbers DQ367853–76). Twenty-one of the 25 plants had unique S-RNase genotypes, while genotypes S03 S12 and S01 S04 were each found in two individuals (Table 1). Only one allele was recovered from one of the sampled plants, whereas several plants had either three or four alleles. The possibility of cross-contamination among samples can be eliminated for genotype 04-P13; thus, it is possible that individuals of *L. parishii* can have >2 S-RNases. In addition, further sequencing of genotype 04-P12 found no evidence of contamination (ie, the presence of the second allele from the potentially contaminating individual). While we recognize that this region is susceptible to contamination (due to the necessity of multiple PCRs and cloning), we do not believe it explains individuals with >2 S-RNases in our sample.

The frequency of alleles in the sample did not significantly differ from a uniform distribution, as assessed using Mantel's (1974) test, and so Paxman's (1963) estimator was used to measure population allelic diversity. We estimated population allele number for three combinations of the number of individuals (r, equation (1)) and alleles (n, equation (1)). First, we included all plants ($r = 25$) and alleles ($n = 24$), which returned a population estimate of 29 alleles. We also estimated allele number by treating those individuals with greater than two alleles (Table 1) as multiple individuals ($r = 30$ and $n = 24$); doing so reduced the estimate of the number of alleles to 27. Finally, excluding all of the individuals with greater than two S-alleles ($r = 20$ and $n = 19$), Paxman's estimator returned a population estimate of 23 *S*-alleles.

Table 1 S-RNase genotypes for 25 *Lycium parishii* plants sampled from Organ Pipe National Monument, Pima county, AZ, USA

Plant ID	*S-genotype*
01-P5	S01 S04
01-P6	S03 S04
01-P7	S05 S06
01-P8	S07 S08 S21
01-P10	S09
01-P14	S03 S12
01-P15	S03 S12
01-P16	S06 S08
01-P17	S08 S11
04-P1	S13 S17 S18 S19
04-P2	S01 S04
04-P3	S04 S19
04-P4	S09 S18
04-P5	S02 S18
04-P6	S02 S14
04-P7	S09 S15 S19
04-P8	S11 S17
04-P9	S08 S11
04-P10	S08 S26
04-P11	S04 S08
04-P12	S04 S12 S17
04-P13	S02 S18 S22
04-P14	S05 S23
04-P15	S10 S24
04-P16	S04 S25

Molecular evolution at the *S-RNase* locus in *Lycium*

Among *L. parishii* S-RNases, the ratio of nonsynonymous substitutions per nonsynonymous site to synonymous substitutions per synonymous site (D_n/D_s) was 0.758 for the entire amplified region of the *S-RNase* gene, which included both the conserved and hypervariable regions specified by Ioerger *et al* (1991). Including only the hypervariable regions HVa and HVb, the D_n/D_s ratio increased to 1.068, whereas including only conserved regions returned a very low D_n/D_s of 0.107 (Figure 1).

Previous studies indicating that variation at fewer than five AA positions can be sufficient to result in a new S-allele specificity (eg, Matton *et al*, 1997; Wang *et al*, 2001) suggests that a site-by-site analysis of nucleotide substitutions may be more appropriate for detecting positive selection at this locus. Likelihood ratio tests comparing site-by-site models M1 (neutral evolution) to M2 (positive selection) were conducted in PAML; these analyses found that the model of positive selection fit the data significantly better than the model of neutral evolution for both *L. parishii* ($2\Delta\ell = 68.12$, $df = 2$, $P < 0.0001$) and *L. andersonii* ($2\Delta\ell = 65.97$, $df = 2$, $P < 0.0001$) S-RNases (Table 2). Among *L. parishii* alleles, positive selection was detected at 27% of AA sites ($p_2 = 0.27$, Table 2) for a total of 35 sites for *L. parishii*, and there was a probability of at least 0.90 that the D_n/D_s ratio exceeded one for 17 of these sites. For *L. andersonii*, positive selection was detected at 37.4% of sites, with a probability of at least 0.90 that the D_n/D_s ratio exceeded one at a total of 25 sites (Table 2).

Figure 2 depicts positively selected sites across the different structural regions of the *S-RNase* gene. Eighteen of the 35 positively selected sites among *L. parishii* alleles occurred within the hypervariable regions HVa and HVb. Further, among the 17 sites with a 90% probability or higher that the D_n/D_s ratio exceeded one, 11 were located in the hypervariable regions. In contrast, 20 of 47 positively selected sites, and only 13 of the 25 sites with a 90% probability or higher that the D_n/D_s ratio exceeded one, occurred within hypervariable regions among *L. andersonii* alleles. Positive selection was not detected at any site within conserved regions C3 or C4 for either species (Figure 2).

L. andersonii and *L. parishii* share 24 positively selected sites (underlined in Figure 2). There are eight common hypervariable sites between the two species where there is at least a 90% probability of positive selection for both species, five in HVa (AA sites 15, 18, 23, 26, and 28) and three in HVb (AA sites 43, 44, and 50). Many of the positively selected sites among *L. andersonii* alleles occur in the region beyond C4, which has not been previously designated hypervariable or conserved.

Alleles S08 and S13 recovered from *L. parishii* (this study) were identical to alleles S11 and S01 recovered by Richman (2000) for *L. andersonii*, reducing the number of interspecific comparisons of closely related alleles to 12. There was no difference in D_n, D_s, or D_n/D_s for the comparison of all intraspecific allele pairs ($n = 19$) to interspecific pairs of closely related alleles (D_s, $U = 120$; D_n, $U = 136$; D_n/D_s, $U = 136$; all $n_1 = 19$, $n_2 = 12$, $U_{critical} = 163$, $P > 0.05$; Table 3).

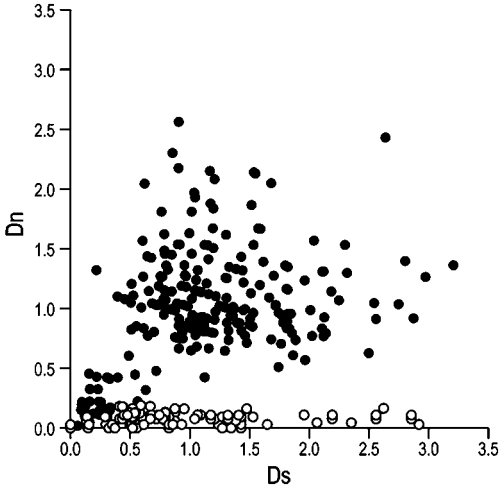

Figure 1 Pairwise comparison of nonsynonymous substitutions per nonsynonymous site (D_n) versus synonymous substitutions per synonymous site (D_s) among *L. parishii* S-RNases. Open symbols indicate comparisons including the conserved regions C3 and C4, whereas closed symbols are for HVa and HVb regions (Ioerger *et al*, 1991).

Table 2 Likelihood ratio tests comparing models of neutral evolution (M1) and positive selection (M2) among S-RNases for *Lycium parishii* and *L. andersonii* as implemented in PAML (Yang, 2002)

	Model	ℓ	$2\Delta\ell$ (*df, P-value*)	*Estimated parameters*	*No. positively selected sites (posterior probability)*
L. parishii	M1: neutral	−5803.94		$p_0 = 0.113$; $p_1 = 0.887$	Not allowed
	M2: positive	−5769.88	68.12 (2, <0.0001)	$p_0 = 0.111$; $p_1 = 0.619$	6 ($P > 0.99$)
				$p_2 = 0.270$ ($\omega_2 = 2.58$)	7 ($P > 0.95$)
					4 ($P > 0.90$)
					18 ($0.90 > P > 0.50$)
					35 Total
L. andersonii	M1: neutral	−3566.58		$p_0 = 0.1107$; $p_1 = 0.893$	Not allowed
	M2: positive	−3533.58	65.99 (2, <0.0001)	$p_0 = 0.099$; $p_1 = 0.526$	8 ($P > 0.99$)
				$p_2 = 0.374$ ($\omega_2 = 3.37$)	13 ($P > 0.95$)
					4 ($P > 0.90$)
					22 ($0.90 > P > 0.50$)
					47 Total

The neutral evolution model assumes two site classes with the D_n/D_s ratio (ω) fixed ($\omega_0 = 0$ and $\omega_1 = 1$) and with the proportions of sites falling into each class (p_0 and p_1) estimated from the data. In the positive selection model, there is a third site class (ω_2) that is also estimated; ω_2 can exceed 1 and is the D_n/D_s ratio of those sites under positive selection. The proportion of sites falling into each class are also estimated from the data (p_0, p_1, and p_2) (see also Yang and Nielson, 2002).

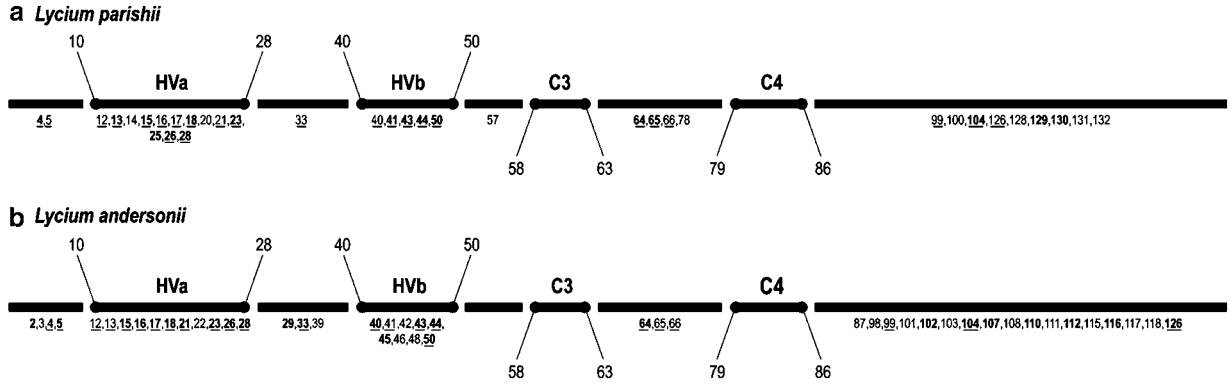

Figure 2 Schematic structure of the *S-RNase* gene between conserved regions C2 and C5 (Ioerger *et al*, 1991) in (**a**) *L. parishii* and (**b**) *L. andersonii*. Regions are drawn to scale and the AA positions of the hypervariable (HVa and HVb) and conserved regions (C3 and C4) are indicated above or below the regions, respectively. Numbers directly below a region are those AA sites reported to be under positive selection using site-specific analyses in PAML. Those sites in bold had a probability of positive selection >90%, whereas others had probabilities between 50 and 90%. Underlined sites represent those shared between *L. parishii* and *L. andersonii*.

Table 3 Average D_s, D_n, and D_n/D_s for closely-related alleles within *Lycium parishii*, within *L. andersonii* and between *L. parishii* and *L. andersonii*

Allele comparison	n	D_s	D_n	D_n/D_s
Intraspecific	19	0.207[a]	0.129[a]	0.572[a]
Lycium parishii	14	0.154	0.091	0.548
Lycium andersonii	5	0.355	0.235	0.639
Interspecific	12	0.212[a]	0.066[a]	0.470[a]

[a]Comparisons of D_s, D_n, and D_n/D_s for intraspecific versus interspecific allele pairs are not significant.

A two-cluster test (Takezaki *et al*, 1995) did not reject a molecular clock for *L. parishii* ($Q = 33.7$, $df = 23$, $P = 0.0694$). Terminal branch lengths relative to the depth of the genealogy were significantly longer than expected ($R_{sd} = 5.22$) when compared to Uyenoyama (1997) simulated frequency distribution for a gene under balancing selection. There was no negative association between the age (measured either by the number of TG lineages or R_{sd}) and number of alleles for *L. parishii* (Table 4).

Solanaceae S-RNase genealogies

S-RNase alleles from *L. parishii* are located across numerous clades in the ML (Figure 3) and Bayesian genealogies (Figure 4). Both the ML and Bayesian topologies cluster *L. parishii* alleles into ten TG lineages (labeled 1–10 in Figures 3 and 4). *L. parishii* allele S15 does not group closely with alleles from any other species, but is so divergent from all conspecific alleles that we consider it a TG lineage and attribute its isolation to incomplete sampling of S-RNase polymorphism in other genera. *L. parishii* alleles are not evenly distributed: seven occur in lineage nine, while the remaining lineages have only one, two or three alleles. In contrast, the *L. andersonii* S-alleles have a relatively even distribution, although substantially fewer alleles (11 S-alleles) were included in the analyses compared to *L. parishii* (24 S-alleles). S-RNases from both species of *Lycium* group together in many lineages (Figures 3 and 4) with only a few exceptions; allele S26 from *L. parishii* is separate from *L. andersonii* alleles, as is the lineage containing alleles S04 and S12 from *L. parishii*.

At the level of genera, *Lycium* exhibits similar levels of TG polymorphism as *Solanum* (11 TG lineages) and *Petunia* (eight TG lineages) (Figure 3). Consistent with Richman and Kohn's (2000) neighbor-joining Solanaceae S-allele genealogy, *Witheringia* and *Physalis* alleles are restricted to three TG lineages (lineages A–C in Figures 3 and 4). There is extensive diversification within two of these three lineages (A and B), but not the third (C) (Figure 4). Similar to *L. andersonii* but not to *L. parishii*, *Solanum* and *Petunia* alleles are generally evenly distributed across lineages, with the exception of one *P. integrifolia* TG lineage that includes 6 alleles.

Table 4 The number of S-RNases sampled from plants within natural populations for seven species of Solanaceae and estimates of allelic diversity in these populations (*n*, number individuals sampled)

Species	n	*No. alleles*		*TG*	R_{sd}
		Sampled	*Estimated*		
Solanum carolinense[a,b]	26	12	15	8	6.27
Physalis crassifolia[b]	22	28	44	2	2.64
Physalis cinerascens[b]	12	12	14	3	2.59
Witheringia maculata[b]	12	10	15	3	1.48
Witheringia solanacea[c]	21	21	37	3	2.73
Lycium andersonii[b]	16	22	38	11	5.42
Lycium parishii	30 (20)	24 (19)	27 (23)	10	5.22

Genealogical information, including the number of TG and relative terminal branch lengths (R_{sd}), are taken from sources listed below. For *Lycium parishii*, the number of TG lineages was estimated from Figures 3 and 4 and calculation of R_{sd} followed Uyenoyama (1997). Two estimates of allele number are included for *L. parishii* depending on how samples with greater than two alleles are treated (see results).
[a]Data taken from Richman (1995).
[b]Data taken from Richman and Kohn (2000).
[c]Data taken from Stone and Pierce (2005).

Discussion

SI in *L. parishii*

One purpose of this study was to verify SI in *L. parishii* using molecular sequence data from the *S-RNase* locus. We recovered a total of 24 S-RNases from *L. parishii*; however, several individuals had >2 alleles. One explanation for the three (or four) allele plants could be that these individuals are polyploid. Chromosome

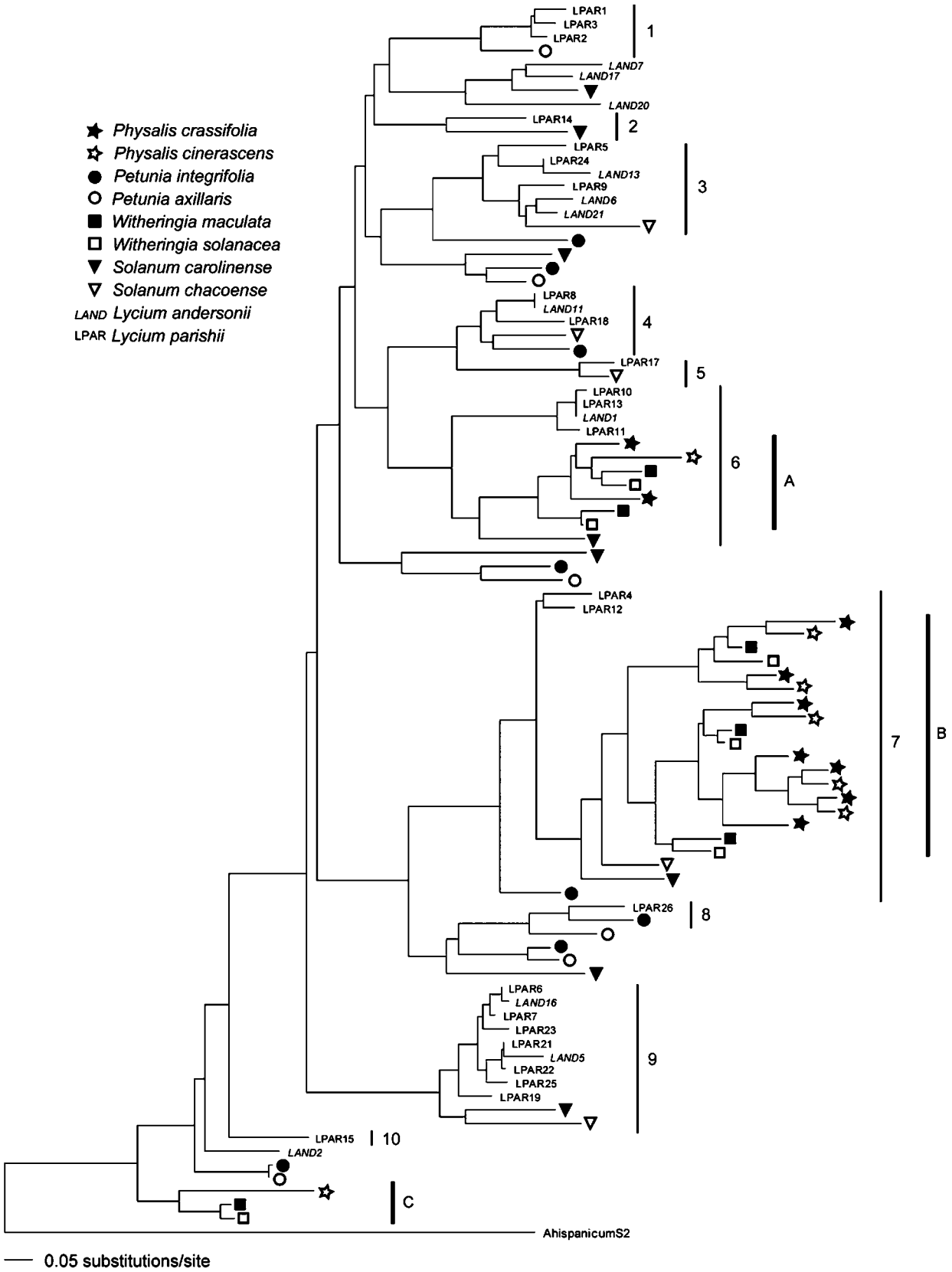

Figure 3 ML genealogy of 90 S-RNase sequences from Solanaceae. The 90 sequences were selected from a larger set of *L. andersonii, Petunia* spp., *Physalis* spp., *Solanum* spp., and *Witheringia* spp. S-RNases obtained from GenBank (see Methods for accession numbers and sampling strategy). Clades 1–10 are *L. parishii* TG lineages and clades A–C are the three lineages including all *Witheringia* and *Physalis* S-RNases.

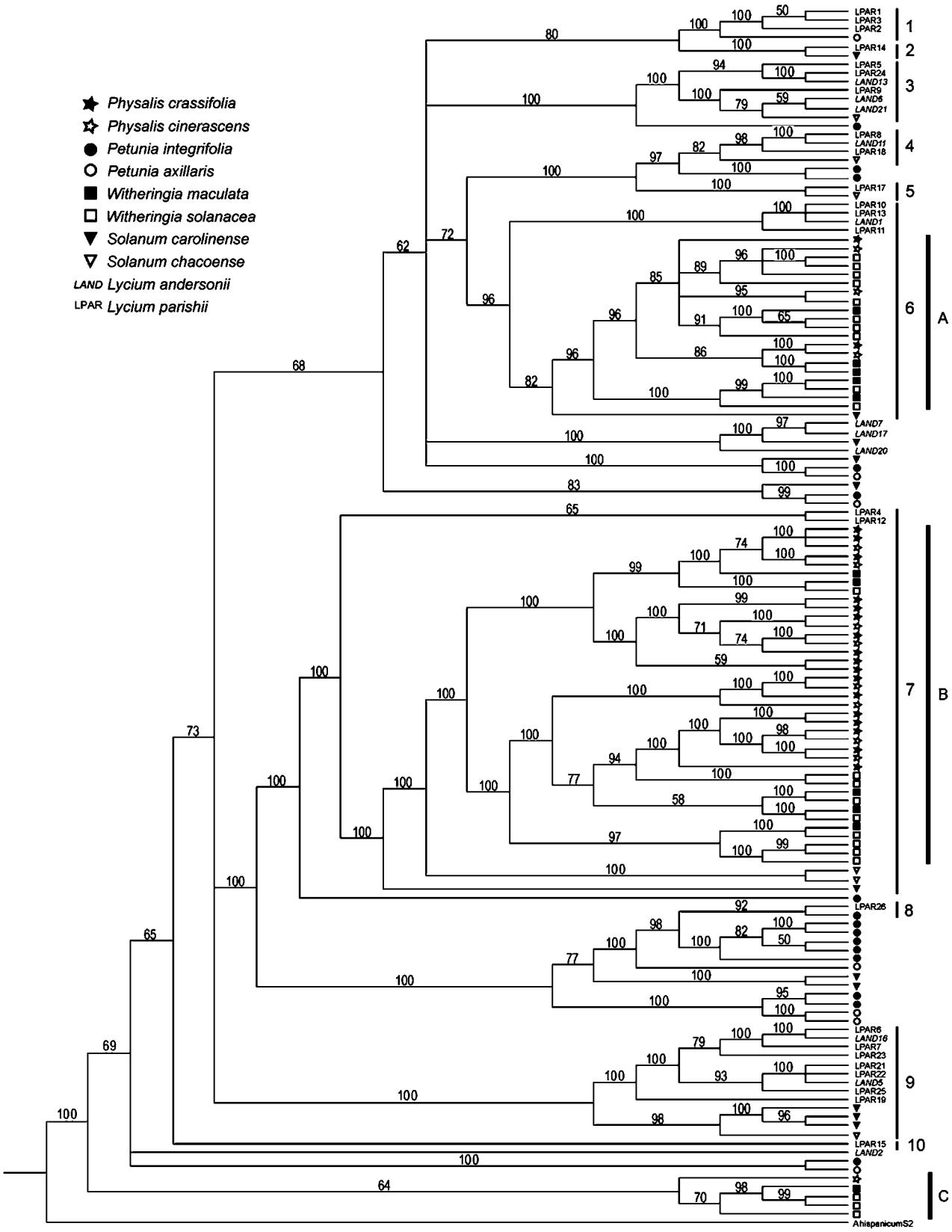

Figure 4 Bayesian genealogy including 138 S-RNase sequences from Solanaceae. Sequences for alleles from *L. andersonii*, *Petunia* spp., *Physalis* spp., *Solanum* spp., and *Witheringia* spp. were taken from GenBank (see Methods for accession numbers). Posterior probabilities are indicated above the branches. Clades 1–10 are *L. parishii* TG lineages and clades A–C are the three lineages including all *Witheringia* and *Physalis* S-RNases.

counts have revealed mixed ploidy (ie, both diploid and polyploid individuals) for several species of *Lycium* (*L. chilense*, Stiefkens and Bernardello, 2000; *L. gariepense*, *L. horridum*, and *L. mascarenense*, Venter, 2000; *L. californicum*, Yeung *et al*, 2005). The *L. parishii* population is located within a national monument, which is home to six additional *Lycium* species; tetraploidy following hybridization between *L. parishii* and one or more of these other species is a possible explanation for plants bearing >2 S-RNases. Whether the putative polyploids are self-incompatible is presently unknown and can only be confirmed via experimental crosses; thus, conclusions on the compatibility status of these plants are premature. In addition, it is unclear if the plants bearing three S-RNases are partially homozygous or whether additional intensive sampling of clones could detect a fourth allele. An alternative interpretation of the >2 S-RNase individuals is that the style samples included pistils collected from two different genotypes. *L. parishii* often grows in dense patches along washes and it is possible that styles from multiple individuals were inadvertently pooled in these samples (see Methods). Treating individuals with three or four alleles as separate genotypes, our population estimate of the number of alleles is 27. Excluding these individuals, the population estimate is 23 alleles. Regardless of which estimate is used, both are within the range of alleles recovered for other SI species (Table 4), and genealogical analysis reveals that TG polymorphism and terminal branch lengths are comparable to other Solanaceae with GSI (Figures 3 and 4).

S-RNases in *L. parishii* were highly polymorphic (average pairwise AA divergence was 52%), and all but a single individual were heterozygous, as expected under GSI. It is unclear why we were not able to retrieve two alleles from 01-P10; perhaps this plant has an extremely divergent S-RNase that we were unable to amplify with the degenerate primers. Alternatively, this plant may be homozygous and possibly self-compatible. This does not necessarily indicate a breakdown of GSI in the *L. parishii* population, as a low frequency of SC individuals has been reported for other populations with GSI (eg, Tsukamoto *et al*, 1999; Stone and Pierce, 2005).

The data presented here, coupled with the experimental crosses in Miller and Venable (2002), confirm the presence of SI in this species. However, it should be noted that Miller and Venable (2002) reported low levels of seed production following self-crosses for *L. parishii*. Given that SI in many Solanaceae appears to be disrupted by polyploidy (reviewed in Stone, 2002), the relationship between the low levels of self-fertility reported by Miller and Venable (2002) and the >2 S-RNase genotypes should be investigated. Indeed, Miller and Venable (2000) have argued that in *Lycium* the loss of SI leads to the evolution of gender dimorphism. Confirmation of GSI in the hermaphroditic close relatives of the dimorphic species is an important first step in evaluating this hypothesis, and an extension of the present study to the dimorphic taxa (Miller and Savage, unpublished data) may shed light on Miller and Venable's suggestion.

Patterns of molecular evolution at the *S-RNase* locus

Despite the classic explanation that balancing selection maintains the extreme allelic polymorphism commonly found at the *S-RNase* locus (Wright, 1939), S-allele studies are often unable to demonstrate such selection using D_n/D_s ratios averaged across the entire amplified *S-RNase* gene (eg, Ma and Oliveira, 2002; Kato and Mukai, 2004). This is not surprising, as the *S-RNase* gene includes five highly conserved regions that are structurally necessary for RNase function and would be expected to evolve primarily under purifying selection. In contrast, hypervariable regions HVa and HVb are located on the surface of the S-RNase protein and are thought to be involved in recognition, making these regions likely candidates for positive selection (Ioerger *et al*, 1991). In *L. parishii*, the low D_n/D_s for conserved regions is consistent with functional constraints, whereas the only comparisons that had a D_n/D_s ratio greater than one were for hypervariable regions (Figure 1). Considering individual sites, many of the positively selected sites in both *L. parishii* and *L. andersonii* were within the hypervariable regions, and there were no positively selected sites detected in the conserved regions (Figure 2). In addition, *L. parishii* and *L. andersonii* S-RNases both fit a model of positive selection significantly better than a model of neutral evolution (Table 2).

Congruence across genera in the hypervariable sites that determine specificity suggests that a single change within targeted regions of the *S-RNase* locus may be capable of generating new alleles. Ida *et al* (2001) compared the hypervariable regions of 10 S-RNases across three genera (*Nicotiana*, *Petunia*, and *Solanum*) in Solanaceae to determine the most variable sites within HVa and HVb. In our alignment, these sites correspond to sites 21 (HVa) and 43 (HVb; Figure 2). There is at least a 95% probability of positive selection at site 21 and at least a 99% probability of positive selection at site 43 for the *L. andersonii* S-RNases. For *L. parishii*, there is only a 74% probability of positive selection at site 21, but a 99% probability of selection at site 43. There are four sites (in our alignment, sites 18, 20, 21, and 46) within the S-RNase hypervariable regions that Matton *et al* (1997) modified in the *S. chacoense* S11 allele to match the S13 allele, resulting in a change to the S13 specificity. When only three of these four sites were modified (sites 18, 21, and 46), the result was a dual-specificity allele that rejected both S11 and S13 pollen (Matton *et al*, 1999). In our analyses, these four sites were found to be under positive selection for both species (Figure 2) with the exception of site 46 in *L. parishii* and site 20 in *L. andersonii*. Overlap across multiple genera in the specific hypervariable sites that differ reinforces the hypothesis that a few targeted AA changes can generate new specificities. This process could enable the rapid regeneration of allelic diversity following a bottleneck event, such as the one suggested in the ancestor of *Witheringia* and *Physalis* (Richman and Kohn, 1999), and may explain the maintenance of GSI in these genera in the period of reduced N_e directly following the bottleneck.

S-RNases recovered from this *L. parishii* population may provide evidence of a new lower limit for the number of AA changes necessary to generate a novel S-RNase specificity. Previous empirical research found a minimum of four AA differences between functionally distinct S-RNases (Matton *et al*, 1997). In *L. parishii*, two alleles, S21 and S22, vary at only two nonsynonymous sites (AA sites 44 and 123, Figure 2) and at a single synonymous site (site 46). Both *L. parishii* alleles were

sequenced twice to verify these nucleotide differences, although functional differences in these alleles can only be confirmed through experimental crosses. Empirical verification that these two sequences are unique and functional S-RNases would reinforce the importance of the hypervariable regions in determining specificity, and suggest that even fewer AA changes may be important to the generation of new alleles.

Solanaceae S-RNase genealogy

Genealogies constructed using neighbor-joining, ML, and Bayesian methods show extensive TG polymorphism among *L. parishii* S-RNases. This is consistent with both theoretical expectations under GSI (Vekemans and Slatkin, 1994) and empirical findings from closely related *L. andersonii* (Richman, 2000). Although there are only six TG lineages for *L. andersonii* in our genealogies (Figures 3 and 4), we restricted our sampling to the 11 *L. andersonii S*-RNase sequences that fully spanned regions C2–C5. Richman's (2000) original analysis included 22 *L. andersonii* alleles and recovered 11 distinct TG lineages. Thus, 10 TG lineages in *L. parishii* is similar to results for the complete set of *L. andersonii* alleles (see also Table 4).

Contrary to the results of Stone and Pierce (2005) for *Witheringia*, we find no evidence for post speciation radiation of S-RNases in *L. parishii* or *L. andersonii*. Estimates of D_s, D_n, and D_n/D_s were equivalent for intraspecific and interspecific allelic comparisons (Table 3), and alleles from both species group within similar TG lineages in the genealogies (Figures 3 and 4). Thus, it would appear that considerable S-RNase diversity was present in the ancestor of these species. It is intriguing then to consider how S-RNases S08 and S13, recovered here for *L. parishii*, could have identical nucleotide sequences to S-RNases S11 (AF10535) and S01 (AF10534) recovered for *L. andersonii* by Richman (2000). Although *L. parishii* and *L. andersonii* are closely related (Miller, 2002; Levin and Miller, 2005), it is unlikely that the S-RNases would have identical DNA sequences. Further, as our laboratory has never worked with *L. andersonii* S-RNases, this finding cannot be explained by cross-contamination. Such a result seems best explained by gene flow between species; additional studies investigating gene flow between *Lycium* are warranted. It is interesting to note that both the S08 and S13 alleles from *L. parishii* were found in individuals with more than two S-RNases (Table 1).

For *Lycium*, *Physalis*, and *Witheringia*, species pairs within genera have similar numbers of TG lineages and terminal branch lengths (Table 4). However, whereas the level of S-RNase diversity between *Lycium* species is comparable, the numbers of alleles are strikingly different for intrageneric comparisons within *Physalis* (44 versus 14 alleles) and *Witheringia* (37 versus 15 alleles; Table 4). Richman and Kohn (1999, 2000) argue that for *Physalis* the difference in allelic diversity is a result of divergent life-history strategies in the two species that affect the long-term stability of populations. Specifically, ephemeral populations of *P. cinerascens* result in small effective population sizes and lower S-RNase diversity compared to larger, stable populations of *P. crassifolia*, which experience strong selection for S-RNase diversification. Similarly, Stone and Pierce (2005) note that the larger altitudinal range of *W. solanacea* (compared to *W. maculata*) may be responsible for maintaining high S-RNase diversity in this species, primarily as a result of seed rain between populations of *W. solanacea*.

In contrast to comparisons within *Physalis* and *Witheringia*, *L. parishii* and *L. andersonii* have similar measures of terminal branch lengths (*L. parishii*, $R_{sd} = 5.22$ versus *L. andersonii*, $R_{sd} = 5.42$; Table 4), indicating similar rates of allelic turnover, and relatively similar and high numbers of S-RNases (38 versus 27 alleles for *L. andersonii* and *L. parishii*, respectively; Table 4). *L. andersonii* has a wider geographic range than does *L. parishii* (Chiang-Cabrera, 1981), and it would be informative to investigate S-RNase diversity in multiple populations of both taxa to obtain species-wide estimates of allele number. Although both species are primarily pollinated by anthophorid bees and seeds are dispersed by birds, the greater range of *L. andersonii* may harbor higher numbers of S-RNases in the species as a whole.

Acknowledgements

This work was supported by the Department of Biology at Amherst College and grants to JSM from NSF (DEB-0343735) and the Amherst College Faculty Research Award Program. We thank RA Levin and JR Cavatorta for assistance with field collections, JL Stone, AD Richman, and JR Kohn for helpful advice and discussion, and anonymous reviewers for comments. Lastly, we thank Organ Pipe National Monument and Sue Rutman and Tim Tibbitts, in particular, for access to the field site.

References

Anderson MA, Cornish EC, Mau SL, Williams EG, Hoggart R, Atkinson A *et al* (1986). Cloning of cDNA for a stylar glycoprotein associated with expression of self-incompatibility in *Nicotiana alata*. *Nature* **321**: 38–44.

Bernardello LM (1986). Revisión taxonómica de las especies sudamericanas de *Lycium* (Solanaceae). *Bol Acad Nac Cienc Córdoba* **57**: 173–356.

Castric V, Vekemans X (2004). Plant self-incompatibility in natural populations: a critical assessment of recent theoretical and empirical advances. *Mol Ecol* **13**: 2873–2889.

Charlesworth D (1985). Distribution of dioecy and self-incompatibility in angiosperms. In: Greenwood PG, Harvey PH, Slatkin M (eds) *Evolution: Essays in Honor of John Maynard Smith*. Cambridge University Press: Cambridge. pp 237–268.

Charlesworth D, Charlesworth B (1987). Inbreeding depression and its evolutionary consequences. *Ann Rev Ecol Syst* **18**: 237–268.

Chiang-Cabrera F (1981). A taxonomic study of the North American species of *Lycium* (Solanaceae). PhD dissertation, University of Texas, Austin, Texas, USA.

Felsenstein J (2005). PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. *Department of Genome Sciences*. University of Washington: Seattle.

Hitchcock CL (1932). A monographic study of the genus *Lycium* of the western hemisphere. *Ann MO Bot Gard* **19**: 179–374.

Ida K, Norioka S, Yamamoto M, Kumasaka T, Yamashita E, Newbigin E *et al* (2001). The 1.55 Å resolution structure of *Nicotiana alata S*F11-RNase associated with gametophytic self-incompatibility. *J Mol Biol* **314**: 103–112.

Ioerger TR, Clark AE, Kao T (1990). Polymorphism at the self-incompatibility locus in Solanaceae predates speciation. *Proc Natl Acad Sci USA* **87**: 9732–9735.

Ioerger TR, Gohlke JR, Xu B, Kao TH (1991). Primary structural features of the self-incompatibility protein in Solanaceae. *Sex Plant Reprod* **4**: 81–87.

Jain SK (1976). The evolution of inbreeding in plants. *Ann Rev Ecol Syst* **7**: 469–495.

Kao T-H, McCubbin AG (1996). How flowering plants discriminate between self and non-self pollen to prevent inbreeding. *Proc Nat Acad Sci USA* **93**: 12059–12065.

Kato S, Mukai Y (2004). Allelic diversity of S-RNase at the self-incompatibility locus in natural flowering cherry populations (*Prunus lannesiana* var *speciosa*). *Heredity* **92**: 249–256.

Kehyr-Pour A, Pernes J (1985). A new *S*-allele and specific *S*-proteins associated with two *S*-alleles in *Nicotiana alata*. In: Mulcahy DL, Mulcahy GB, Ottaviano E (eds) *Biotechnology and ecology of pollen*. Springer-Verlag: New York. pp 191–196.

Kumar S, Tamura K, Nei M (2004). MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings Bioinformatics* **5**: 150–163.

Lawrence MJ (2000). Population genetics of the homomorphic self-incompatibility polymorphisms in flowering plants. *Ann Bot London A* **85**: 221–226.

Levin RA, Miller JS (2005). Relationships within tribe Lycieae (Solanaceae): paraphyly of *Lycium* and multiple origins of gender dimorphism. *Am J Bot* **92**: 2044–2053.

Lu Y (2001). Roles of lineage sorting and phylogenetic relationship in the genetic diversity at the self-incompatibility locus of Solanaceae. *Heredity* **86**: 195–205.

Lu Y (2002). Molecular evolution at the self-incompatibility locus in *Physalis longifolia* (Solanaceae). *J Mol Evol* **54**: 784–793.

Ma RC, Oliveira MM (2002). Evolutionary analysis of *S*-RNase genes from Rosaceae species. *Mol Genet Genomics* **267**: 71–78.

Mantel N (1974). Approaches to a health research occupancy problem. *Biometrics* **30**: 355–362.

Matton DP, Luu DT, Qin X, Laublin G, O'Brien M, Maes O *et al* (1999). The production of an *S*-RNase with dual specificity suggests a novel hypothesis for the generation of new *S*-alleles. *Plant Cell* **11**: 2087–2097.

Matton DP, Maes O, Laublin G, Xike Q, Bertrand C, Morse D *et al* (1997). Hypervariable domains of self-incompatibility RNases mediate allele-specific pollen recognition. *Plant Cell* **9**: 1757–1766.

McClure BA, Haring V, Ebert PR, Anderson MA, Simpson RJ, Sakiyama F *et al* (1990). Style self-incompatibility gene products of *Nicotlana alata* are ribonucleases. *Nature* **342**: 955–957.

Miller JS (2002). Phylogenetic relationships and the evolution of gender dimorphism in *Lycium* (Solanaceae). *Syst Bot* **27**: 416–428.

Miller JS, Venable DL (2000). Polyploidy and the evolution of gender dimorphism in plants. *Science* **289**: 2335–2338.

Miller JS, Venable DL (2002). The transition to gender dimorphism on an evolutionary background of self-incompatibility: an example from *Lycium* (Solanaceae). *Am J Bot* **89**: 1907–1915.

Olmstead RG, Sweere JA, Spangler RE, Bohs L, Palmer JD (1999). Phylogeny and provisional classification of the Solanaceae based on chloroplast DNA. In: Nee M, Symon DE, Lester RN, Jessop JP (eds) *Solanaceae IV*. Richmond Botanic Gardens: Richmond, UK. pp 111–137.

Paxman GJ (1963). Maximum likelihood estimation of the number of self-sterility alleles in a population. *Genetics* **48**: 1029–1032.

Posada D, Crandall KA (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**: 817–818.

Richman AD (1995). *S*-allele sequence diversity in natural populations of *Solanum carolinense*. *Heredity* **75**: 405–415.

Richman AD (2000). *S*-allele diversity in *Lycium andersonii*: Implications for the evolution of *S*-allele age in the Solanaceae. *Ann Bot London A* **85**: 241–245.

Richman AD, Kao TH, Schaeffer SW, Uyenoyama MK (1995). *S*-allele sequence diversity in natural populations of *Solanum carolinense* (Horsenettle). *Heredity* **75**: 405–415.

Richman AD, Kohn JR (1996). Learning from rejection: the evolutionary biology of single-locus incompatibility. *Trends Ecol Evol* **11**: 497–502.

Richman AD, Kohn JR (1999). Self-incompatibility alleles from *Physalis*: implications for historical inference from balanced genetic polymorphisms. *Proc Natl Acad Sci USA* **96**: 168–172.

Richman AD, Kohn JR (2000). Evolutionary genetics of self-incompatibility in the Solanaceae. *Plant Mol Biol* **42**: 169–179.

Richman AD, Uyenoyama MK, Kohn JR (1996). *S*-allele diversity in a natural population of *Physalis crassifolia* (Solanaceae) (ground cherry) assessed by RT-PCR. *Heredity* **76**: 497–505.

Ronquist F, Huelsenbeck JP (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572–1574.

Saitou N, Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**: 406–425.

Sijacic P, Wang X, Skirpan AL, Wang Y, Dowd PE, McCubbin AG *et al* (2004). Identification of the pollen determinant of *S*-RNase-mediated self-incompatibility. *Nature* **429**: 302–305.

Stiefkens L, Bernardello G (2000). Karyotypes and DNA content in diploid and polyploid *Lycium* (Solanaceae). *B Soc Argent Bot* **35**: 237–244.

Stone JL (2002). Molecular mechanisms underlying the breakdown of gametophytic self-incompatibility. *Q Rev Biol* **77**: 17–32.

Stone JL, Pierce SE (2005). Rapid recent radiation of *S*-RNase lineages in *Witheringia solanacea* (Solanaceae). *Heredity* **94**: 547–555.

Swofford DL (2002). *PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods)* Version 4. Sinauer Associates: Sunderland, MA.

Takezaki N, Rzetsky A, Nei M (1995). Phylogenetic test of the molecular clock and linearized trees. *Mol Biol Evol* **12**: 823–833.

Thompson JD, Higgins DG, Gibson TJ (1994). CLUSTAL X: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl Acids Res* **22**: 46–73.

Tsukamoto T, Ando T, Kokubun H, Watanabe H, Masada M, Zhu X *et al* (1999). Breakdown of self-incompatibility in a natural population of *Petunia axillaris* (Solanaceae) in Uruguay containing both self-incompatible and self-compatible plants. *Sex Plant Reprod* **12**: 6–13.

Uyenoyama MK (1997). Genealogical structure among alleles regulating self-incompatibility in natural populations of flowering plants. *Genetics* **147**: 1389–1400.

Vekemans X, Slatkin M (1994). Gene and allelic genealogies at a gametophytic self-incompatibility locus. *Genetics* **137**: 1157–1165.

Venter AM (2000) Taxonomy of the genus *Lycium* L. (Solanaceae) in Africa. PhD dissertation, University of the Free State, Bloemfontein, South Africa.

Wang X, Hughes AL, Tsukamoto T, Ando T, Kao T (2001). Evidence that intragenic recombination contributes to allelic diversity of the *S*-RNase gene at the self-incompatibility (*S*) locus in *Petunia inflata*. *Plant Physiol* **125**: 1012–1022.

Wright S (1939). The distribution of self-sterility alleles in populations. *Genetics* **24**: 538–552.

Yang Z (2002) *Phylogenetic Analysis by Maximum Likelihood (PAML) Version 3.13d*. Distributed by the author. Department of Biology, University College London.

Yang Z, Nielsen R (2002). Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol Biol Evol* **19**: 908–917.

Yeung K, Miller JS, Savage AE, Husband BC, Igic B, Kohn JR (2005). Association of ploidy and sexual system in *Lycium californicum* (Solanaceae). *Evolution* **59**: 2048–2055.