doi:10.1111/j.1558-5646.2011.01564.x

PHYLOGENY OF *LEAVENWORTHIA* S-ALLELES SUGGESTS UNIDIRECTIONAL MATING SYSTEM EVOLUTION AND ENHANCED POSITIVE SELECTION FOLLOWING AN ANCIENT POPULATION BOTTLENECK

Adam C. Herman,[1,2] Jeremiah W. Busch,[3] and Daniel J. Schoen[1]

[1]*Department of Biology, McGill University, Montreal, Quebec H3A 1B1, Canada*

[2]*E-mail: adam.herman@mail.mcgill.ca*

[3]*School of Biological Sciences, Washington State University, Pullman, Washington 99164*

Received August 23, 2011

Accepted December 16, 2011

Data Archived: Dryad: doi: 10.5061/dryad.38jf742m

The adoption of self-fertilization from an ancestral outcrossing state is one of the most common evolutionary transitions in the flowering plants. In the mustard family, outcrossing is typically enforced by sporophytic self-incompatibility (SI), but there are also many self-compatible species. The genus *Leavenworthia* contains taxa that either possess or lack SI. Here, we present data showing that SI is associated with strict outcrossing and that there is widespread trans-specific sequence polymorphism at the locus involved in the recognition of self-pollen (the S-locus). This ancestral polymorphism is consistent with the presence of an outcrossing mating system in the common ancestor of *Leavenworthia* species, and suggests that there have been several independent losses of SI in the group. When compared with other mustard species, the bulk of *Leavenworthia* S-allele sequences are highly diverged from those found in other Brassicaceae and show relatively low levels of nucleotide diversity, a pattern that suggests the common ancestor of the genus likely underwent a strong population bottleneck. The hypothesis of postbottleneck S-locus rediversification is supported by tests showing stronger positive selection acting on S-alleles from *Leavenworthia* than those found in other Brassicaceae.

KEY WORDS: **Mating system evolution, positive selection, self-incompatibility, trans-specific polymorphism.**

In close to half of flowering plant species, outcrossing is enforced by one of several different systems of self-pollen recognition, commonly referred to as self-incompatibility (SI) (Igic et al. 2008). Yet throughout the history of the angiosperms, SI has been lost repeatedly, giving rise to self-compatibility (SC) in a large number of taxa. The loss of SI is an important step in the evolution of selfing, which is among the most common transitions in the higher plants (Stebbins 1974), but the mechanisms that explain this repeated trend are not well understood. Advances in the molecular biology and genetics of SI have led to new insights about the evolution of plant reproductive systems, both at and above the species level (Kondo et al. 2002; Schierup et al. 2006; Goldberg et al. 2010). In the case of sporophytic SI (SSI), detailed molecular-level studies of the mechanism of SI in the mustard genera *Brassica*, *Arabidopsis*, and *Capsella*, coupled with population genomic investigations, have identified the mutations causing the breakdown of SI, shed light on the timing of these events, and suggested possible selective factors involved in the evolution of selfing from outcrossing (Bechsgaard et al. 2006; Foxe et al. 2009; Guo et al. 2009; Tsuchimatsu et al. 2010). The

mustard genus *Leavenworthia* has also served as an important plant group for understanding the ecological and evolutionary aspects of the shift from outcrossing to selfing (Lloyd 1965; Liu et al. 1998; Busch 2005; Koelling et al. 2010), and in recent years has emerged as a new system for investigating mating system evolution using population genomics approaches (Busch et al. 2008, 2010a, 2011; Joly and Schoen 2011).

The locus controlling the first steps of the SI reaction in mustards (the S-locus) has been shown to consist of two tightly linked genes encoding proteins that function together as a receptor–ligand system (Fobis-Loisy et al. 2004). These are the S-locus receptor kinase (*SRK*), which codes for a membrane-bound, stigma-expressed protein, and the S-locus cysteine-rich gene (*SCR* in *Arabidopsis*; *SP11* in *Brassica*), which codes for a pollen expressed protein ligand (Kachroo et al. 2001; Takayama et al. 2001). Binding of *SCR* to *SRK* initiates a signal cascade that acts to prevent pollen tube hydration and penetration of the stigma surface (Murase et al. 2004; Liu et al. 2007).

The S-domain of the putative *SRK* ortholog in *Leavenworthia* (originally designated as *Lal2*), which codes for the specific portion of the protein involved in recognition in the stigma, was recently cloned and characterized in *Leavenworthia alabamica* (Busch et al. 2008). *Lal2* shares many of the characteristics reported for the S-domain of *SRK* as in other SI mustards. These include cosegregation with the SI reaction (as observed in diallel crosses), high expression levels in stigmas but not leaves, extremely high levels of nucleotide polymorphism, three hypervariable subdomains that contain a number of codons exhibiting evidence of past diversifying selection, and a within-species phylogenetic pattern of sequence variation characterized by deep coalescence (Busch et al. 2008). In an earlier study, despite the similarities between *Lal2* and *SRK*, phylogenetic analysis of the mustard S-locus indicated that most of the *L. alabamica* sequences form a distinct clade separated by a long branch from the clades containing *SRK* sequences of *Arabidopsis* and *Brassica* (Busch et al. 2008). In view of this pattern, and because of the failure at the time to amplify other regions of *SRK* such as the transmembrane and kinase domains, it was tentatively concluded that while *Lal2* might indeed be the *SRK* ortholog in *Leavenworthia*, the possibility could not be ruled out that it instead represented a separate S-linked locus, or a duplicated and diverged paralogous copy of *SRK* (Busch et al. 2008). More recently, we have succeeded in cloning and sequencing the full genomic region containing *Lal2* in *L. alabamica*. *Lal2* can now be confirmed to contain all seven exons present in the *SRK* locus of other mustards (i.e., those coding for the S-domain, transmembrane domain, and kinase domain). Further, *Lal2* resides less than 10 kb away from an *SCR*-like locus containing a putative exon coding for a small cysteine-rich protein, as well as a second putative exon coding for a conserved signal peptide, as seen in other mustard *SCR*s (S.C. Chantha, et al., unpubl. data), and henceforth we refer to this locus in *Leavenworthia* as *SRK*.

Past studies of S-locus evolution in *Leavenworthia* have mostly focused on one taxon, *L. alabamica*, a species where both SI and SC populations co-occur—this despite the existence of pronounced mating system variation between species across the entire genus (Rollins 1963; Lloyd 1965; Beck et al. 2006). In this article, we analyze S-locus variation in all *Leavenworthia* species, except for two threatened sister species, *L. aurea* and *L. texana*. These studies were conducted to test hypotheses that are central to understanding the evolutionary lability of the SI mechanism and the origins and ultimate fate of selfing taxa (Igic et al. 2008; Goldberg et al. 2010). First, we test whether SI is associated with a high level of outcrossing within two species of *Leavenworthia*. Although SI is generally thought to prevent self-pollination, leaky SI (De Nettancourt 2000) or biparental inbreeding could give rise to significant levels of inbreeding in populations. Because outcrossing rates of SI *Leavenworthia* taxa have not before been estimated from progeny array data (Ritland 2002), such estimates can provide new quantitative information to test the hypothesis that strict outcrossing in SI species is associated with the long-term maintenance of S-locus variation. Second, we test the idea that SI species are polymorphic for S-alleles that are descendants of the same copy in a recent common ancestor (i.e., the S-alleles exhibit trans-specific polymorphism). If this pattern is observed, it is consistent with a long and uninterrupted history of SI, as it is highly unlikely that such similar sequences would evolve independently in each SI lineage (Igic et al. 2006). Third, given the uninterrupted history of SI among SI taxa, we use relationships among the SC and SI species to estimate the minimum number of times that the pathway from outcrossing to selfing has been taken in *Leavenworthia*.

In addition to these inquiries that focus on evolutionary shifts in the mating system, we test a fourth hypothesis that pertains to divergence observed between *Leavenworthia SRK* sequences and those of other mustard species. In particular, the observation of a distinct S-locus clade within the mustards formed by the majority of *Leavenworthia* sequences studied to date suggests that a strong population bottleneck occurred within the evolutionary history of the *Leavenworthia* S-locus. Following recovery from such a bottleneck in which S-alleles are lost to genetic drift, selection for new specificities is expected to be especially strong, leaving a signature of elevated levels of positive selection in the hypervariable domains (compared with populations that have not passed through such a bottleneck). We examine the strength of positive selection inferred using codon-specific models of selection in *Arabidopsis, Brassica, Capsella,* and *Leavenworthia*, and discuss these results in an effort to better explain broader patterns of S-locus evolution in Brassicaceae.

Materials and Methods

SPECIES AND POPULATIONS STUDIED

Leavenworthia is a genus of winter annuals that is largely restricted to cedar glade habitats in the southeastern United States (Rollins 1963). We sampled S-alleles from six of the eight extant *Leavenworthia* species: *L. alabamica*, *L. crassa*, *L. exigua*, *L. stylosa*, *L. torulosa*, and *L. uniflora* (Table S1). *Leavenworthia stylosa* maintains populations comprised mainly of SI plants. In *L. alabamica* and *L. crassa*, some populations have SI plants whereas others are composed of SC plants (Koelling et al. 2010). We analyzed S-alleles found in both types of populations in *L. alabamica*, however, we only analyzed S-alleles from a single SI population of *L. crassa*. In the remaining taxa, plants are entirely SC (Rollins 1963; Lloyd 1965), and we sequenced S-alleles from across the entire range of these species. Plant tissue and/or seeds were collected from these populations in 2008 and 2009.

MATING SYSTEM ANALYSES

To assess whether populations consist mainly of individuals with SI, plants were grown from seed collected in the field from separate seed parents. Following emasculation, flowers were self-pollinated, stigmas were harvested the next day, preserved in 10% acetic acid–ethanol overnight at 4°C, then cleared in NaOH for 1 h and stained with a 0.1% aniline blue solution. Fluorescence microscopy with a 4′-6′ diamidino-2-phenylindole filter and a digital camera was used to visualize pollen tube growth. Depending on the plant, two to four flowers were selfed in this way. The plant was considered self-compatible if a majority of self-pollinations yielded stigmas with a least five visible pollen tubes per stigma. This procedure likely underestimates the strength of SI in the field, where pollen from both self- and cross-pollination may co-occur on the same stigma. Populations were characterized as SI when >75% of plants showed incompatibility reactions with self-pollen.

We estimated outcrossing rates for populations of SI plants in both *L. alabamica* (Population 28) and *L. crassa* (Population 31). In each population, 5 seeds from each of 50 random parents were collected in the field. In addition, silique tissue was collected from each parent. Maternal parent DNA was extracted from the silique tissue, and progeny DNA was extracted from embryos, to avoid potential bias in outcrossing rate estimation associated with low germination rate and early-acting inbreeding depression. DNeasy plant mini kits were used for all extractions (Qiagen, Inc. Valencia, CA). For Population 28, genotypes at two microsatellite loci (C3 and C109) were obtained for all progeny and parents whereas for Population 31 genotypes at a single microsatellite locus (C3) were obtained–see Busch et al. (2010b) for description of the microsatellite loci and amplification methods. Because of amplification failures or genotyping ambiguities, family size varied between two and five individuals in each population sample. A total of 42 families were used in estimating outcrossing rate in *L. alabamica* whereas 50 were used in *L. crassa*. Mating system estimation under the mixed mating model was conducted using the procedures in Ritland (2002).

S-LOCUS SEQUENCE DATA

S-locus sequence data for *L. alabamica* and *L. crassa* were assembled from previously published reports (Busch et al. 2008, 2010a; Joly and Schoen 2011). For the remaining *Leavenworthia* species (*L. exigua, L. stylosa, L. torulosa,* and *L. uniflora*), genomic DNA was extracted from leaf or silique tissue that was either frozen or dried in silica gel. All extractions were carried out using DNeasy plant mini kits (Qiagen Inc.). Allelic variation in the S-domain of *SRK* was sampled in *L. exigua, L. torulosa,* and *L. uniflora* using previously published primer pairs and methods described in Busch et al. (2010a). These primers gave weak amplification products for *L. stylosa* samples, therefore a separate primer pair was used to amplify *SRK* in that species (Lal-Sdomain-F1 5′-ACCTTTGGTGGCAG AGCTTC-3′, SLGR 5′-ATCTGACATAAAGATCTTGACC-3′, 51°C annealing temperature).

To ensure that all of the different S-alleles present in a population sample were sequenced, we first screened the PCR products (S-alleles) using single-strand conformational polymorphism (SSCP) according to previously published methods (Busch et al. 2010a). Briefly, denatured PCR amplicons were electrophoretically separated using polyacrylamide gels, and bands corresponding to single strands of alleles were excised from the gel and reamplified to obtain individual allele products that were then sequenced. These procedures allowed detection of all S-alleles present in each population sample. Allele sampling for *L. alabamica* and *L. crassa* is detailed in Busch et al. (2008, 2010a) and Joly and Schoen (2011). The 13 *L. stylosa* alleles used in this study were sampled from 17 individuals from six populations (Table S1). Amplification of the paralogous locus *Lal1* (Busch et al. 2008) was common in *L. stylosa* samples, and as such, the number of alleles detected in this sample may be an underestimate. Sampling in each of the SC species involved single plants sampled from six populations from across the species range, so these collections reflect S-locus sequence diversity at the species level. To examine relationships among *Leavenworthia* S-alleles and those from other mustard species, we assembled a dataset of previously published *SRK* sequences from *Arabidopsis*, *Brassica*, and *Capsella* (Table S2). These sequences were aligned with the *Leavenworthia* sequences using the software package Geneious (version 5.4.2, Auckland, New Zealand), and trimmed to the shortest allele's length. Genus-specific alignments were then created from the larger alignment and used to estimate nucleotide polymorphism, as well as estimate positive selection

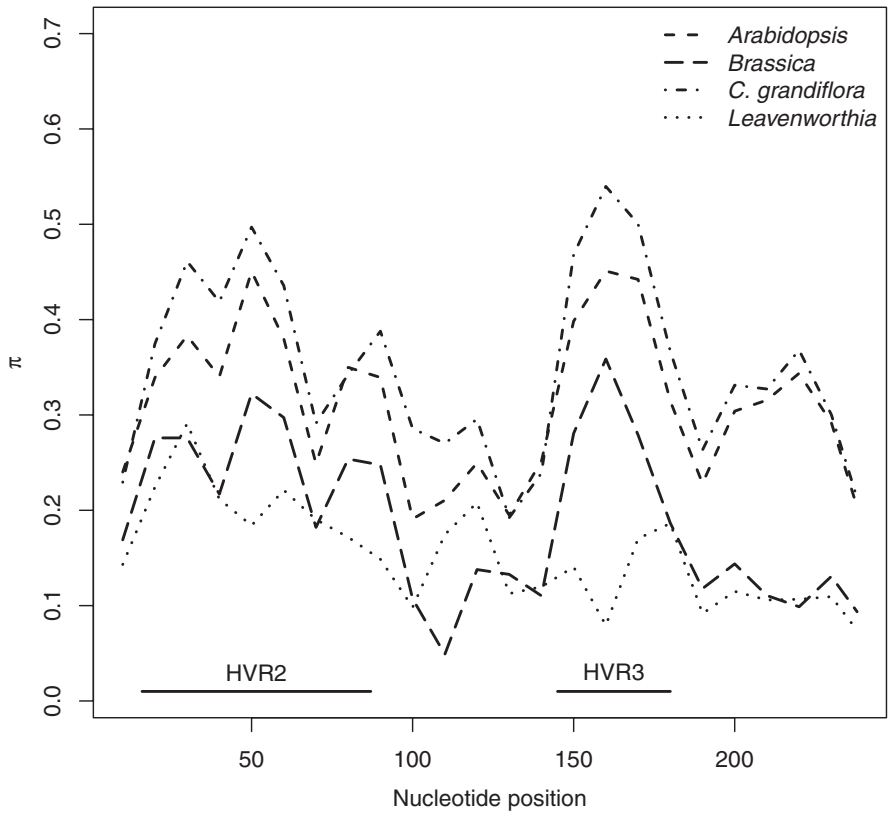

Figure 1. Sliding window analysis of nucleotide diversity (π) in the portion of the S-domain of SRK encompassing hypervariable regions 2 and 3. The numbers of sequences used to estimate π for each taxon are shown in Table 1. *Leavenworthia* sequences unique to this study are deposited in GenBank under accession numbers (JQ397415-JQ397449; JQ714259-JQ714260, accession numbers for all other Leavenworthia alleles are available in the TreeBASE archive for this study (see Figure 2B legend).)

(see below). Analyses of nucleotide diversity, including sliding window analyses, were carried out in DnaSP Version 5 (Librado and Rozas 2009).

PHYLOGENETIC ANALYSES OF S-ALLELES

Prior to conducting phylogenetic analysis of S-allele variation, the best-fit nucleotide substitution model was determined using the Aikake Information Criterion in jModeltest v.0.1.1 (Guindon and Gascuel 2003; Posada 2008). Bayesian phylogenetic inference for *Leavenworthia* S-alleles was carried out using MrBayes v. 3.1.2 (Huelsenbeck and Ronquist 2001) under the SYM + I + Γ model with estimation of all parameters. Two independent analyses of six Markov chain Monte Carlo (MCMC) chains each were run simultaneously for 4 million generations, with sampling every 1000th generation. The first 250 trees were discarded as burn-in and the remaining 751 were used to generate a consensus phylogeny for the *SRK* alleles.

To examine the phylogenetic relationships among S-alleles from *Leavenworthia*, *Arabidopsis*, *Brassica*, and *Capsella*, Bayesian phylogenetic inference was carried out on this larger dataset under the GTR+I+Γ model (determined using jModeltest, as above), again estimating all parameters. Two analyses of six MCMC chains were run for 4 million generations with a burn-in of 250 trees. The phylogenies presented below are 50% majority-rule consensus trees. Additional phylogenetic analyses of these sequences are presented as supplemental results (Figs. S1, S2).

ANALYSIS OF SELECTION OF CODONS IN S-ALLELE SEQUENCES

Frequency-dependent selection at the S-locus is expected to favor mutations conferring new mating specificities (Wright 1939). The portion of *SRK* analyzed in this study spans 246 nucleotides in the first exon encompassing two hypervariable regions, HVR2 and HVR3, and surrounding nucleotides, as detailed in *Brassica* (Nishio and Kusaba 2000; Fig. 1). There is evidence that these regions are involved in recognition of pollen specificities (Kusaba and Nishio 1999), and accordingly, selection for new specificities should lead to high ratios of nonsynonymous to synonymous substitution in these regions (Castric and Vekemans 2007). We analyzed *SRK* sequence variation in the SI taxa of *Leavenworthia* with tools designed to estimate the strength of positive selection on a codon-by-codon basis using the CODEML program in the PAML software package (Yang 2007; Castric and Vekemans

Table 1. Diversity statistics for the *SRK* alleles studied[1]. The sequences were used for all phylogenetic and molecular evolutionary analyses.

Taxon	N[2]	S[3]	π	θ_w
Arabidopsis	50	203	0.30958	0.184
Brassica	53	170	0.18837	0.15228
Capsella grandiflora	17	196	0.34753	0.23567
Leavenworthia	73	200	0.14263[4]	0.16679

[1]All sequences are 246 nucleotides in length.
[2]Number of sequences.
[3]Number of segregating sites.
[4]The two most divergent *Leavenworthia* alleles are not included in this estimate.

2007). We were especially interested in testing the hypothesis that positive selection at the *Leavenworthia SRK* sequences is strong compared to other mustard genera, as might be expected if *Leavenworthia* had undergone a particularly strong bottleneck.

The likelihood of three alternative models of selection at codon sites was compared using likelihood ratio tests (LRTs). The general M7 model allows the dN/dS ratio, ω, at each codon site to vary between 0 and 1 according to the estimated parameters of the Beta distribution. A more specific M8 model adds a class of sites in which ω may be 1 or greater (and is referred to below as the "positive selection model"). A third model, M8a, resembles model M8, but with an added class of sites constrained to $\omega = 1$ (Swanson et al. 2003). LRTs comparing model M7 with M8, and M8 with M8a, with 2 and 1 degrees of freedom, respectively, were used to assess the mode of selection operating on codons in each genus studied. When the positive selection model (M8) provided the best fit, the Bayes Empirical Bayes method (Yang et al. 2005) was used to identify the individual codon sites under positive selection (Table 2). An average ω was obtained for all such sites with ≥95% posterior probability.

To test whether estimates of ω for the *Leavenworthia* (and other mustard family genera) departed from the family average, we used the parametric bootstrapping approach of Castric and Vekemans (2007). First, model M8 was run on data comprising *Arabidopsis, Brassica, Capsella,* and *Leavenworthia* sequences to estimate the parameters of the model for the family. M8 was also used to infer an ML tree with branch lengths. Next, the estimated M8 parameters, codon counts and tree were used as inputs to the simulation program "Evolver" (also in the PAML package), and 100 replicate datasets of *SRK* evolution for the family were simulated. For each replicate, the simulated sequences were divided into their respective genera and model M8 was applied on a genus-by-genus basis to identify sites under positive selection, as described above. For each of the genus-specific datasets obtained from simulation, the average ω was calculated over all sites with ≥95% probability of belonging to the $\omega > 1$ class. Finally, the distributions of these average ω values were compared with the ω estimated from the actual data for each genus.

CODEML estimates the codon-specific ω parameter via maximum likelihood given a particular phylogeny. If recombination has taken place in the history of the sequences from which ω is estimated, such a phylogenetic approach is inappropriate because individual sites within the sequences may have independent phylogenies. This has the effect of increasing the type I error rate with respect to the LRT, although the M7 and M8 comparison is relatively robust to false positives (Anisimova et al. 2003). Although the recombination rate in S-loci is thought to be low, instances of intragenic recombination have been documented (Castric et al. 2010). We tested for the presence of recombination in *Leavenworthia* sequences using the permute program included in the omegaMap program package (McVean et al. 2002; Wilson and McVean 2006). Bayesian inference of the posterior distribution of the ω parameter for *Leavenworthia* sequences was performed using omegaMap. The Bayesian framework of the program allows

Table 2. Number, identity, and average ω for sites identified as evolving under positive selection in the taxa studied, according to Model M8 implemented in CODEML.

Taxon	N[1]	Sites[2]	ω[3]
Arabidopsis	13	274 D*, 285 K*, 286 S*, 287 I*, 303 V*, 305 K*, 306 V*, 320 L, 338 G*, 341 A*, 353 S, 358 T, 359 R	1.49
Brassica	12	274 D*, 285 K*, 287 I*, 288 L*, 303 V*, 306 V*, 320 L, 330 W*, 332 M*, 339 E*, 340 A*, 341 A*	2.49
Capsella grandiflora	1	320 L	1.53
Leavenworthia	12	273 S*, 274 D*, 278 Q*, 280 L*, 289 W*, 303 V*, 306 V*, 319 P, 322 N, 329 P*, 342 S*, 353 S	3.49

[1]Number of sites identified as $\omega > 1$ with ≥95% probability.
[2]Codon position and amino acid with respect to first start codon of *BoSRK60.*
[3]Averaged over all sites identified as $\omega > 1$ with ≥95% probability.
* Indicates codons within hypervariable regions.

effective incorporation of phylogenetic uncertainty into the estimation of ω. We ran two independent MCMC chains for 500,000 generations each, with sampling every 100th generation. Improper inverse prior distributions on the parameters were used and the posterior distribution of ω was estimated independently for each codon. From those posterior distributions, each codon's mean ω was calculated.

Results

OUTCROSSING RATES IN SELF-INCOMPATIBLE TAXA

Estimates of mating system parameters for populations of SI plants in *L. alabamica* and *L. crassa* indicate that both populations exhibit a high rate of outcrossing, although a small amount of selfing may take place ($t = 0.934 \pm 0.033$ and 0.860 ± 0.059, respectively). It should be noted that single-locus estimates of the outcrossing rate, as conducted for *L. crassa*, cannot discriminate apparent selfing due to biparental inbreeding and this may explain the reduced rate in that species in comparison to *L. alabamica*. However, estimates of the parental inbreeding coefficient with standard errors are indistinguishable from zero in each population (*L. alabamica*: 0.003 ± 0.018, *L. crassa*: 0.036 ± 0.045). Taken together, these data indicate that the prohibition of selfing by the SI mechanism is effective in enforcing outcrossing within SI populations.

SEQUENCE VARIATION AT THE *LEAVENWORTHIA* S-LOCUS

As expected for a locus under strong balancing selection, a large amount of sequence variation was detected among different *Leavenworthia SRK* alleles. In general, the level of variation is at least 10-fold higher than that observed at the non-S-linked nuclear loci surveyed in this genus by Joly and Schoen (2011), although compared with other mustard genera (*Arabidopsis*, *Brassica,* and *Capsella*), variation for this same region of *SRK* is somewhat reduced (Table 1). This is also apparent in the sliding window analysis of diversity in this region, where peaks of variability are detectable in the HVR2 and HVR3 regions, paralleling the locations of those seen in the other mustard genera (Fig. 1).

PHYLOGENY OF *LEAVENWORTHIA* S-ALLELES

Leavenworthia S-alleles exhibit a phylogeny characterized by long terminal branch lengths, as has been observed in other S-locus phylogenies (Fig. 2A; Uyenoyama 1997; Schierup et al. 1998). The SI taxa each possess a large number of S-alleles (see also Busch et al. 2010a; Joly and Schoen 2011) and in several instances, the same or similar (<5 nucleotide differences) S-alleles are found in several of these taxa, in other words, a pattern of trans-specific polymorphism is observed. In contrast, each of the SC populations and species are fixed for a distinct and single class

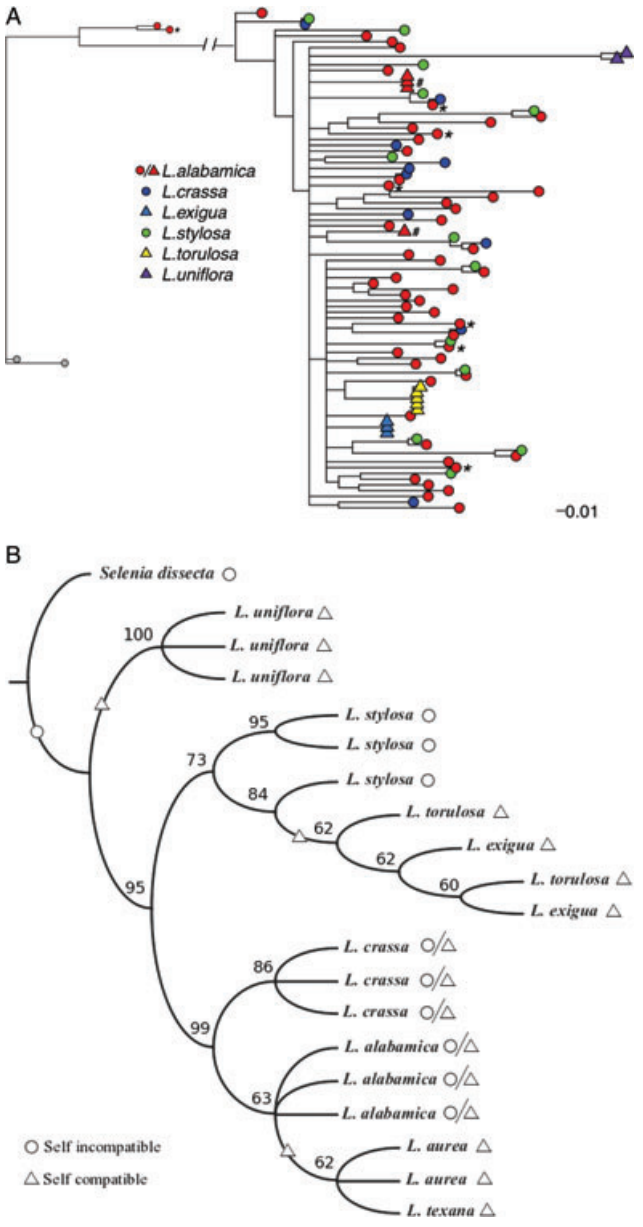

Figure 2. **(A) Majority-rule (50%) consensus phylogeny for *Leavenworthia SRK* alleles. Circles and triangles represent SI and SC species, respectively. An asterisk (*) indicates sequences that have been demonstrated to cosegregate with the self-incompatibility (SI) reaction in *L. alabamica*. A pound sign (#) indicates S-alleles from SC populations of *L. alabamica*. The total distance of the branch leading from the divergent *L. alabamica* S-alleles to the clade formed by the rest of the alleles is 1.3767 substitutions per site. *Arabidopsis ARK1* and *ARK3* sequences (gray circles) were used to root the tree. The tree, along with the alignment used to generate it, is deposited in TreeBASE (http://purl.org/phylo/treebase/phylows/study/TB2:S12225). (B) cpDNA phylogeny of the genus *Leavenworthia* (after Beck et al. 2006). The mating systems of the extant species are indicated following the species names. Inferred ancestral mating systems (see Results) are indicated along the branches. Bootstrap values are indicated at the nodes. Note that *L. alabamica* and *L. crassa* contain both SI and SC populations.**

of S-alleles. Importantly, the S-alleles that are found in each of the SC species are related to alleles from species with functional SI systems that enforce outcrossing (see below). The sequences from *L.exigua* and the SC populations of *L. alabamica*, do not contain stop codons, suggesting that the *SRK* product is functional, although this does not rule out mutations elsewhere in the locus. In *L. torulosa*, two of the five sequences presented in this study do not contain stop codons and neither does one of the two *L. uniflora* sequences. However, nucleotide ambiguities in one of those *L. torulosa* sequences and the single *L. uniflora* sequence preclude a definite conclusion about the lack of stop codons.

Two *Leavenworthia* S-alleles are separated from most of the alleles sampled in this genus by an especially long branch (Fig. 2A); these share close similarity with *Arabidopsis SRK* sequences (see below). Several lines of evidence suggest that these are indeed functional alleles rather than nonfunctional paralogs. First, one of these alleles, which had been detected in an earlier study, showed cosegregation with the SI reaction (Busch et al. 2008). Second, among the 11 *L. alabamica* individuals that harbored these *Arabidopsis*-like *SRK* sequences, each also possessed only one other S-allele, as predicted given the expectation that most individuals should be heterozygous at *SRK*. We have not observed *Arabidopsis*-like *SRK* alleles in the other SI species (*L. crassa* and *L. stylosa*). These alleles are quite rare in *L. alabamica* and our sampling in the other SI species is likely not complete enough to rule out their presence.

PHYLOGENY OF ARABIDOPSIS, BRASSICA, CAPSELLA, AND *LEAVENWORTHIA* S-ALLELES

Most *Leavenworthia* S-alleles occupy a distinct clade in the family-wide phylogeny and are related to S-alleles in the other genera through a long branch (Fig. 3). As mentioned above, two divergent *Leavenworthia* S-alleles were detected in this study, and when compared to S-alleles of other mustards, these are found to be closely related to S-alleles from *Arabidopsis*. These alleles appear to represent the only cases of transgeneric polymorphism between *Leavenworthia* and the other mustard family members. In contrast, there is extensive trans-specific and transgeneric polymorphism within and among *Arabidopsis*, *Brassica,* and *Capsella*, although S-alleles from *Brassica* species are often more closely related to each other than to S-alleles from either *Arabidopsis* or *Capsella*. The reduced polymorphism among *SRK* sequences in *Leavenworthia*, together with their family-wide phylogenetic relationships, suggest a bottleneck during which SI was retained followed by diversification of *SRK* alleles.

ANALYSIS OF SELECTION OF CODONS IN S-ALLELE SEQUENCES

For each taxon studied, LRTs indicated that a model incorporating sites under positive selection (model M8) was most appropriate

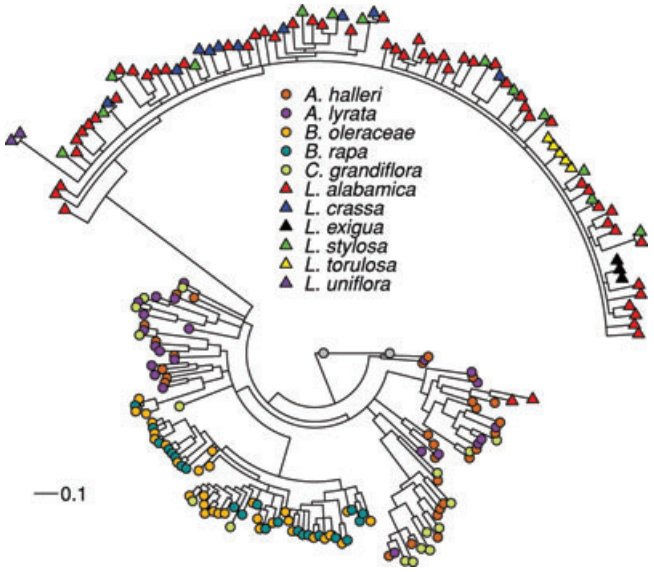

Figure 3. **Majority-rule (50%) consensus phylogeny for all *Arabidopsis*, *Brassica*, *Capsella grandiflora*, and *Leavenworthia SRK* alleles used in this study. *Arabidopsis ARK1* and *ARK3* (gray circles) were used to root the tree. The tree, along with the alignment used to generate it, are deposited in TreeBASE (DOI: X).**

(Table S3). Sites identified as evolving under positive selection in model M8 are detailed in Table 2, and codon positions are referenced from the start codon of *BoSRK60* (GenBank: AB032474). A similar number of positively selected sites were identified in *Arabidopsis*, *Brassica,* and *Leavenworthia* (i.e., residues 13, 12 and 12, respectively), whereas only one positively selected site was identified in *Capsella grandiflora* (this is likely an artifact of the small number of sequences available for the latter taxon). As has been observed previously (Castric and Vekemans 2007), there is low concordance among taxa in positively selected sites (four sites shared among *Arabidopsis*, *Brassica,* and *Leavenworthia*), although as expected most sites are located in the hypervariable regions of the sequence. The average ω for positively selected sites in *Leavenworthia* (3.49) is nearly 150% as large as *Brassica* (2.49) and more 200% that of *Arabidopsis* and *C. grandiflora* (1.49 and 1.53, respectively). Although not as extreme as estimates from the Solanaceae (Miller et al. 2008), which exhibit gametophytic SI (GSI), this is to our knowledge the largest estimate of positive selection acting on S-loci in SSI systems.

The distribution of the average values of ω for the taxa as obtained by CODEML with data simulated using Evolver under the family-based parameter estimates can be summarized with histograms (Fig. 4). For *Arabidopsis*, *Brassica,* and *C. grandiflora*, the ω for each taxon based on the actual data (dashed line) falls above the lower 2.5% and below the higher 97.5% of the ω values obtained from simulated data, indicating that the ω for these taxa could not be distinguished from the family average

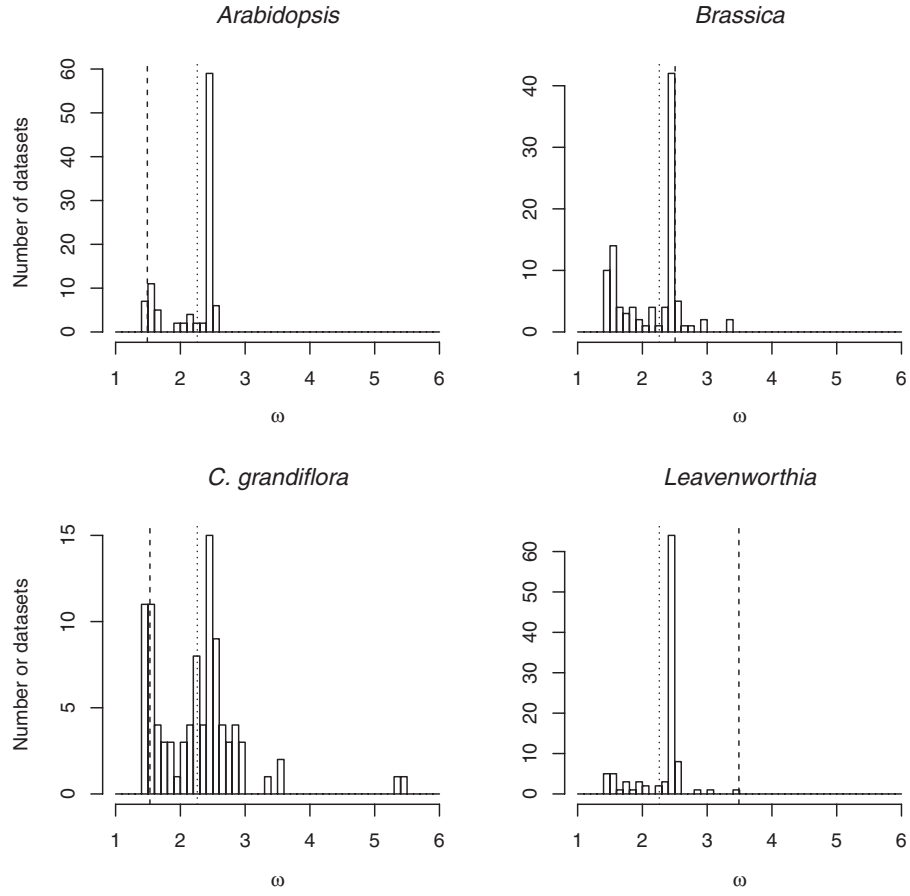

Figure 4. Summary of CODEML runs on data simulated in Evolver. Shown are frequency distributions for average estimated ω from 100 simulated datasets. The vertical dotted line indicates the family-wide average estimate of ω. The dashed line indicates taxon-specific average ω estimates (see Table 2).

(dotted line). In contrast, of the 100 simulated datasets corresponding to *Leavenworthia* sequences, none exhibited an average ω as large as that found for the genus (Fig. 4). We therefore reject the null hypothesis that the *Leavenworthia* sequences evolved under the family-based estimated parameters, and conclude that these S-alleles have experienced significantly stronger recent positive selection than those of the other genera included in this study.

Using permutation tests, we found no significant evidence for the presence of intragenic recombination in *Leavenworthia* sequences (data not shown). As such, when estimating the recombination parameter in OmegaMap, the entire sequence was treated as a single block (as opposed to permitting recombination between individual codons or blocks of codons). Results nearly identical to those presented below were obtained when recombination between codons was allowed (data not shown). With respect to the individual codons inferred to be experiencing positive selection in *Leavenworthia* sequences, omegaMap identified 10 of the codons identified by CODEML and four additional codons not identified by that program (Fig. 5). Only two codons identified by CODEML as evolving under positive selection were not identified by omegaMap. The general concordance between the programs in terms of codons inferred to be under positive selection indicates that our CODEML results are not likely the result of false-positive inference of positive selection.

Discussion

SI AND ITS IMPACT ON OUTCROSSING RATE

Model-based estimates of the mating system from progeny array genotype data show that SI populations of *L. alabamica* and *L. crassa* exhibit outcrossing rates that are high (maximally 97% and 92%, respectively). These estimates contrast with the lower outcrossing rates estimates (ranging from 10% to 45%) in the SC populations of *L. alabamica* (Busch et al. 2010b, 2011). That SI promotes a high outcrossing rate is not a novel insight, although we do provide the first molecular-based estimates of outcrossing for SI taxa in *Leavenworthia*. Additionally, these results lend further support to the suggestion that the maintenance of trans-specific variation at the S-locus in SI species, and its total collapse

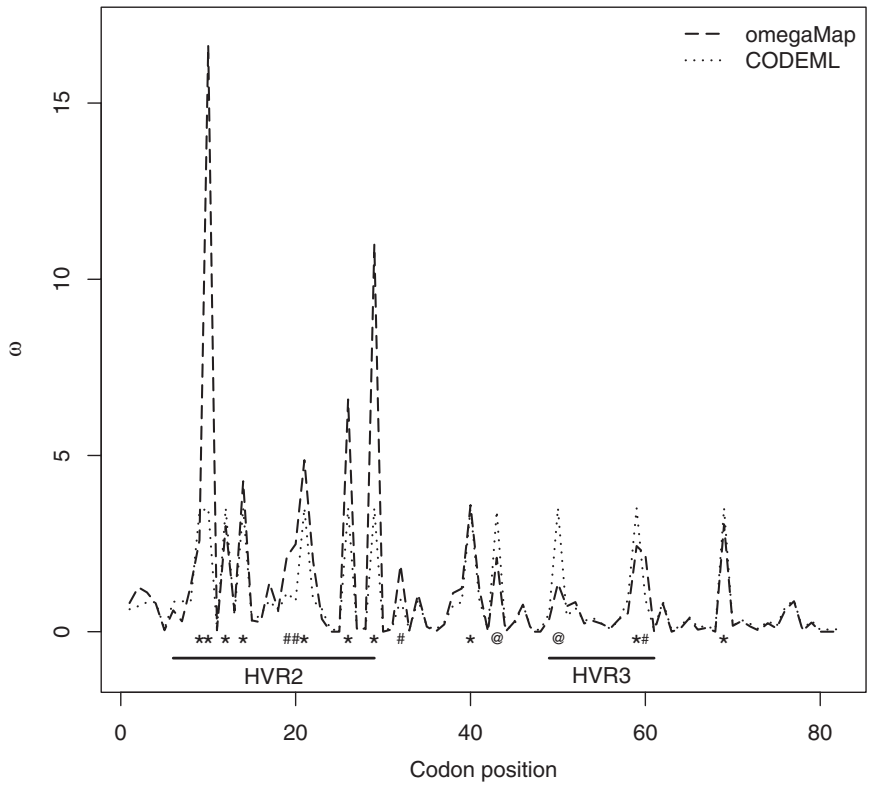

Figure 5. Codon-by-codon estimates of ω for *Leavenworthia* sequences from CODEML and omegaMap. Symbols denote codons with ≥95% probability of ω > 1 identified by both programs (*), CODEML only (@) or omegaMap only (#).

in SC species, is associated with contrasting mating systems in these taxa.

THE EXTENT OF ANCESTRAL S-LOCUS POLYMORPHISM IN SI SPECIES

S-alleles within the genus *Leavenworthia* exhibit a pattern of widespread trans-specific polymorphism. It is highly unlikely that the polymorphism shared among the SI species is the result of convergent evolution. Instead, the shared polymorphism is most likely the result of the persistence of SI, and consequently functional S-alleles, since the last common ancestor of those species. Moreover, some of the S-alleles in *Leavenworthia* appear to be closely related to those in the genus *Arabidopsis*. As pointed out by Igic et al. (2006), when SI breaks down, polymorphism at the S-locus becomes selectively neutral, and is thus expected to be lost within four N_e generations, or perhaps sooner if loss is due to the active selection of a nonfunctional S-allele that causes SC. The presence of trans-specific S-allele polymorphism in *L. alabamica*, *L. crassa*, and *L. stylosa* is therefore consistent with a long and uninterrupted history of SI in each of these lineages.

Reduced S-locus polymorphism would prevent successful reproduction with fewer than three separate S-alleles (mustard SI systems examined to date contain both dominant and codominant S-alleles). If SI were to be regained in a taxon following such a loss by the evolution of new functional S-alleles, it is expected that those alleles would form a unique clade because of their recent diversification (Igic et al. 2006; Paape et al. 2008). The fact that the S-alleles maintained in the currently SI taxa (*L. stylosa*, *L. alabamica*, and *L. crassa*) were not lost before or during the diversification events separating *L. uniflora* from the remainder of the genus implies that each of these species has retained homologous polymorphism from a common ancestor. Thus, there is no evidence for reversal of mating system evolution (from SC to SI) within the genus. Because each of the SC taxa is typified by a unique S-allele, it is much more likely that each loss of the SI mechanism in the genus occurred in parallel.

Arguments have been made elsewhere as to why the re-establishment of SI in SC populations is unlikely, and therefore a well-supported reversal from SC to SI has yet to be reported within either GSI or SSI systems (Igic et al. 2008). When SC evolves in a formerly SI population, there is not only a collapse of S-allele polymorphism, but other loci involved in downstream reactions that are part of normal SI functioning may be rendered selectively neutral and prone to accumulation of loss-of-function mutations (Kondo et al. 2002; Tantikanjana et al. 2010). Any one of these mutational events in a SC lineage would make the re-evolution of SI dependent upon a specific and seemingly unlikely backward mutation. Furthermore, purging of partially recessive,

deleterious mutations that underlie the expression of inbreeding depression may occur, and this should result in relaxed selection for mechanisms that enforce outcrossing (Lande and Schemske 1985).

MULTIPLE TRANSITIONS FROM SI TO SC AND THE EVOLUTIONARY FATE OF SC TAXA

Beck et al. (2006) recently published a phylogeny of *Leavenworthia* based on relationships among chloroplast DNA characters (Beck et al. 2006) (Fig. 2B). Given this phylogeny and the evidence presented for the sequence of the S-locus, if SI is accepted to be the ancestral mating system, it would appear that there have been at least five separate transitions from SI to SC within the genus. One of these occurred in the lineage leading to *L. uniflora*, another leading to the species pair of *L. exigua* and *L. torulosa*, a third leading to the *L. aurea* and *L. texana* species pair, in addition to two within *L. alabamica* (Fig. 2B; Busch et al. 2011). Three of these transitions were also inferred by Beck et al. (2006), but due to the lack of necessary information on trans-specificity of S-alleles required to infer the ancestral mating system in the genus, they underestimated the number of mating system transitions using a parsimony approach to minimize shifts from SI to SC. Although higher than that of Beck et al. (2006), our estimate of five transitions from SI to SC likely represents the minimum number of such transitions in *Leavenworthia* because our analysis does not include alleles from the SC populations of *L. crassa* described in Koelling et al. (2010). In addition, this number may be even higher if there exist as yet undiscovered selfing populations in SI species, if *L. texana* and *L. aurea* independently evolved SC, and if there existed now extinct SC populations or species.

The ancestral mating system of SI, a lack of support for mating system reversals from SC to SI, and the evidence for a high transition rate from SI to SC in *Leavenworthia*, collectively suggest an evolutionary pattern in which the adoption of SC is unidirectional. Interestingly, each of the independently derived SC *Leavenworthia* lineages has been shown to harbor very little nucleotide polymorphism within populations (Liu et al. 1998; Charlesworth 2003; Busch et al. 2011). Apart from the often lower levels of genetic variation, which could constrain adaptation in selfing populations (Schoen and Brown 1991; Hamrick and Godt 1996; Bartkowska and Johnston 2009), it has been suggested that the reduction of recombination in populations of selfers may lower the efficacy of natural selection, which may further impair adaptation, and lead to the accumulation of deleterious mutations (reviewed in Glemin et al. 2006 and Wright et al. 2008). This could lead to an elevated rate of extinction in selfers, as suggested by recent phylogenetically based estimates of extinction rates of SI and SC taxa in the Solanaceae (Goldberg et al. 2010). In turn, the occurrence of higher extinction rates in SC taxa has been postulated as a factor that may help to account for patterns of mating system diversity at and above the species level (Igic et al. 2008; Schoen and Busch 2008).

S-LOCUS BASED EVIDENCE OF A POPULATION BOTTLENECK

Although we have found two *Leavenworthia* S-alleles that share close similarity with those present in *Arabidopsis* (indicating persistence of SI since their common ancestor), the majority of S-allele sequences in the genus form a distinct clade separated from other mustard S-alleles by many nucleotide substitutions. Coupled with evidence of reduced nucleotide polymorphism, such a pattern suggests that the common ancestor of *Leavenworthia* underwent a population bottleneck. Similar evidence has been used to infer bottlenecks in the Solanaceous genera *Physalis* and *Witheringia*, which like *Leavenworthia*, seem to contain S-alleles that are not found elsewhere within the family (Richman et al. 1996; Paape et al. 2008). Similarly, in the genus *Lycium*, an ancient dispersal event across the Atlantic Ocean has reduced the extent of transgeneric polymorphism at the S-locus, yet SI has been maintained (Miller et al. 2008). The evolution of novel S-alleles and reduction in the diversity among S-alleles lineages are both expected to be preserved for long periods of time. This makes the S-locus an excellent candidate for addressing biogeographic or climatic hypotheses concerning ancient events in flowering plants (Miller et al. 2008; Paape et al. 2008).

The conclusion that a population bottleneck occurred in the recent history of *Leavenworthia* receives further support from the CODEML- and omegaMap-based analyses of selection at codons within the hypervariable region of the S-domain of *SRK*. Because population bottlenecks are expected to reduce S-locus diversity, and because the S-locus in general is thought to be subject to negative frequency-dependent selection (i.e., new S-alleles that arise via mutation should be highly favored when rare), there is the expectation that following a bottleneck, S-alleles will exhibit a strong signature of positive selection, particularly in portions of the *SRK* and *SCR* genes coding for amino acids involved in recognition. The presence of strong positive selection (stronger than that observed from analyses of other mustard S-locus alleles) further supports the hypothesis that there was a population bottleneck in the ancestor of the extant species of *Leavenworthia*. In an analysis that included only *Brassica* and *Arabidopsis* S-alleles, a bottleneck was invoked by Castric and Vekemans (2007) to account for stronger positive selection acting on *Brassica* S-alleles. Such a result has also been reported in the S-RNase locus of the Solanaceous genus *Physalis* (the gene that encodes the stylar specificity component of SI in this family), a taxon that has also apparently undergone recent diversification (Paape and Kohn 2011).

The relationships among the recently diversified *Leavenworthia* S-alleles appear to resemble a star phylogeny nested within

the larger phylogenetic tree of *SRK* sequences from the Brassicaceae. Because the process of diversification is expected to produce a pattern of structure (i.e., clearer resolution of relationships) among S-allele sequences and the descendant specificities they produce, the lack of phylogenetic structure has important implications for the process of postbottleneck S-allele diversification in *Leavenworthia*. It is possible that mutations at relatively few amino acid residues are necessary to generate novel specificities. This hypothesis appears possible, as very few amino acids influence specificity in the *SRK* molecule (Boggs et al. 2009). Most of the sites under positive selection reside within HVR2, which has been shown to harbor many amino acid residues that determine specificity (Sainudiin et al. 2005; Boggs et al. 2009). It would also be necessary for such novel *SRK* specificities to rapidly become associated with similarly novel *SCR* specificities, which also appear to involve few amino acids (Chookajorn et al. 2004). If this process had, in fact, occurred following the bottleneck inferred from our data, subsequent mutational change accumulating over time in each S-allele lineage would produce the long branch lengths characteristic of a star phylogeny. The fact that little phylogenetic structure remains following this process of rediversification appears consistent with such a process, and implies at the very least that postbottleneck rediversification occurred at a rapid rate. The observation of pronounced phylogenetic structure among S-alleles in most other species may therefore be generated by the stochastic extinction of S-allele sequences over long periods of evolutionary time.

It is important to note, however, that there may be an alternative explanation for the observed phylogenetic pattern and selective intensity. Specifically, Schierup et al. (1998) have shown that for taxa exhibiting SSI characterized by asymmetric dominance relationships among alleles (i.e., alleles that act codominantly in the stigma but show dominance in pollen), trans-specific polymorphism of S-alleles may be less likely because dominant S-alleles have reduced life spans. Such asymmetric dominance relationships are well known in the Brassicaceae and have been shown in *L. alabamica* (Schierup et al. 1998; Busch et al. 2008). Although dominant S-alleles are lost stochastically at a faster rate than more recessive alleles in a natural turnover process, it is unclear over what time scale we would expect trans-specific or transgeneric polymorphism under such a model, as well as to what degree S-alleles in *Leavenworthia* conform to the modeled dominance relationships (Schierup et al. 1997). Further examination of the dominance hierarchy of S-alleles and their expected turnover rates in *Leavenworthia* could help resolve this possibility.

Although we do not examine the nature of the potential bottleneck in *Leavenworthia*, there are reasons to expect demographic fluctuations in the ancestor of the genus. All extant *Leavenworthia* are edaphic endemics adapted to limestone cedar glades that are quite small in area and may therefore limit population size (Rollins 1963; Busch and Urban 2011). Additionally, these habitats are naturally fragmented and potentially susceptible to extinction/recolonization events, which may lead to bottlenecks.

Conclusions

By examining population genetic diversity at the S-locus in *Leavenworthia*, we have provided evidence that the loss of SI in *Leavenworthia* species and populations has been unidirectional, and that SC has likely evolved independently at least five times within the genus. Additionally, we have provided evidence that SI enforces high rates of outcrossing in *Leavenworthia* species. Both of these results are consistent with a growing body of evidence suggesting that the shift from an outcrossing to a selfing mating system leads lineages to become evolutionary dead-ends and that these taxa are often of recent origin (Igic et al. 2006; Goldberg et al. 2010). Attempts to understand why selfing is an evolutionary dead-end have so far been inconclusive (Wright et al. 2008; Escobar et al. 2010). For studies that have focused on testing the hypothesis that selection is inefficient in selfers, it is possible that the taxa examined are of sufficiently recent origin such that differences with their outcrossing relatives are difficult to detect. Another is that large effective population sizes in selfers or effective purging of deleterious mutations may mask signatures of the reduced efficacy of selection potentially underlying higher extinction rates in selfers (Wright et al. 2008). It is important to understand the reasons for the reduced longevity of selfing lineages, given the increasing amount of evidence that the adoption of selfing ultimately elevates the probability of lineage extinction.

In addition to inferences about ancestral mating system in the genus, we have shown that *Leavenworthia* S-alleles experience significantly stronger positive selection than those from other Brassicaceae species studied so far. We suggest that this accelerated rate of amino acid substitution is the result of an ancient bottleneck that severely reduced S-allele number but did not result in the breakdown of SI. A bottleneck may also explain the phylogenetic characteristics and elevated levels of positive selection observed in the S-alleles of *Brassica*. Within other members of the mustard family, strong bottlenecks have been shown to be associated with the transition to SC (Foxe et al. 2009; Guo et al. 2009). It would be interesting to understand when such bottlenecks lead to the loss of SI, and when they allow persistence of SI and preservation of trans-specific and transgeneric polymorphisms at the S-locus (Miller et al. 2008).

A major feature of the family-wide phylogeny is the long branch separating the majority of *Leavenworthia* S-alleles from those of the other Brassicaceae represented in this study. This pronounced divergence, coupled with the lack of phylogenetic structure among *Leavenworthia* S-alleles, raises long-standing questions about how (and how quickly) novel S-alleles evolve.

As S-locus data from other mustard species accumulate, it should be possible to assess the generality of the phylogenetic patterns observed here and to gain insight into the broader forces governing evolution at the S-locus in the Brassicaceae.

ACKNOWLEDGMENTS

This work was supported by an NSERC Discovery Grant to DJS and by the NSERC CANPOLIN Strategic Network program. The authors thank M. Boileau and D. Leung for laboratory work. M. Thompson conducted laboratory work supported by a National Science Foundation (NSF) UBM fellowship (DMS 0531870). We thank C. Herlihy for help in identifying species and collecting seeds. Three anonymous reviewers made suggestions that improved early versions of the manuscript. JWB thanks the College of Sciences and the School of Biological Sciences at WSU for funding. This is CANPOLIN paper number 28.

LITERATURE CITED

Anisimova, M. R., R. Nielsen, and Z. Yang. 2003. Effect of recombination on the accuracy of the likelihood method for detecting positive selection at amino acid sites. Genetics 164:1229–1236.

Bartkowska, M. P., and M. O. Johnston. 2009. Quantitative genetic variation in populations of *Amsinckia spectabilis* that differ in rate of self-fertilization. Evolution 63:1103–1117.

Bechsgaard, J. S., V. Castric, D. Charlesworth, X. Vekemans, and M. H. Schierup. 2006. The transition to self-compatibility within S-haplotypes over 10 Myr. Mol. Biol. Evol. 23:1741–1750.

Beck, J. B., I. A. Al-Shebaz, and B. A. Schaal. 2006. *Leavenworthia* (Brassicaceae) revisited: testing classic systematic and mating system hypotheses. Syst. Bot. 31:151–159.

Boggs, N. A., K. G. Dwyer, M. E. Nasrallah, and J. B. Nasrallah. 2009. In vivo detection of residues required for ligand-selective activation of the S-locus receptor in *Arabidopsis*. Curr. Biol. 9:786–791.

Busch, J. W. 2005. The evolution of self-compatibility in geographically peripheral populations of *Leavenworthia alabamica* (Brassicaceae). Am. J. Bot. 92:1503–1512.

Busch, J. W., and D. J. Schoen. 2008. The evolution of self-incompatibility when mates are limiting. Trends Plant Sci. 13:128–136.

Busch, J. W., and L. Urban. 2011. Insights gained from 50 years of studying the evolution of self-compatibility in *Leavenworthia* (Brassicaceae). Evol. Biol. 38:15–27.

Busch, J. W., J. Sharma, and D. J. Schoen. 2008. Molecular characterization of *Lal2*, an *SRK*-like gene linked to the S-locus in the wild mustard *Leavenworthia alabamica*. Genetics 178:2055–2067.

Busch, J. W., S. Joly, and D. J. Schoen. 2010a. Does mate limitation in self-incompatible species promote the evolution of selfing? The case of *Leavenworthia alabamica*. Evolution 64:1657–1670.

Busch, J. W., C. R. Herlihy, L. Gunn, and W. J. Werner. 2010b. Mixed mating in a recently derived self- compatible population of *Leavenworthia alabamica* (Brassicaceae). Am. J. Bot. 97:1005–1013.

Busch, J. W., S. Joly, and D. J. Schoen. 2011. Demographic signatures accompanying the evolution of selfing in *Leavenworthia alabamica*. Mol. Biol. Evol. 28:1717–1729.

Castric, V., and X. Vekemans. 2007. Evolution under strong balancing selection: how many codons determine specificity at the female self-incompatibility gene *SRK* in Brassicaceae? BMC Evol. Biol. 7:132.

Castric, V., J. S. Bechsgaard, S. Grenier, R. Noureddine, M. H. Schierup, and X. Vekemans. 2010. Molecular evolution within and between self-incompatibility specificities. Mol. Biol. Evol. 27:11–20

Charlesworth D. 2003. Effects of inbreeding on the genetic diversity of populations. Philos. Trans. R. Soc. Lond. B 358:1051–1070.

Chookajorn, T., A. Kachroo, D. R. Ripoll, A. G. Clark, and J. B. Nasrallah. 2004. Specificity determinants and diversification of the *Brassica* self-incompatibility pollen ligand. Proc. Natl. Acad. Sci. U.S.A. 101:911–917.

De Nettancourt, R. 2000. Incompatibility and incongruity in wild and cultivated plants. Springer-Verlag, Berlin.

Escobar, J. S., A. Cenci, J. Bolognini, A. Haudry, S. Laurent, J. David, and S. Glemin. 2010. An integrative test of the dead-end hypothesis of selfing evolution in Triticeae (Poaceae). Evolution 64:2855–2872.

Fobis-Loisy I., C. Meige, and T. Gaude. 2004. Molecular evolution of the S locus controlling mating in the Brassicaceae. Plant Biol. 6:109–118.

Foxe, J. P., T. Slotte, E. A. Stahl, B. Neuffer, H. Hurka, and S. I. Wright. 2009. Recent speciation associated with the evolution of selfing in *Capsella*. Proc. Natl. Acad. Sci. U.S.A. 106:5241–5245.

Guindon, S., and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52:696–704.

Glemin, S., E. Bazin, and D. Charlesworth. 2006. Impact of mating systems on patterns of sequence polymorphism in flowering plants. Proc. R. Soc. Lond. B 273:3011–3019.

Goldberg, E. E., J. R. Kohn, R. Lande, K. A. Roberson, S. A. Smith, and B. Igic. 2010. Species selection maintains self-incompatibility. Science 330:493–495.

Guo, Y-L., J. S. Bechsgaard, T. Slotte, B. Neuffer, M. Lascoux, D. Weigel, and M. H. Schierup. 2009. Recent speciation of *Capsella rubella* from *C. grandiflora*, associated with loss of self-compatibility and an extreme bottleneck. Proc. Natl. Acad. Sci. U.S.A. 106:5246–5251.

Hamrick J. L., and M. J. W. Godt. 1996. Effects of life history traits on genetic diversity in plant species. Philos. Trans. R. Soc. Lond. B 351:1291–1298.

Huelsebeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Igic, B., L. Bohs, and J. R. Kohn. 2006. Ancient polymorphism reveals unidirectional breeding system shifts. Proc. Natl. Acad. Sci. U.S.A. 103:1359–1363.

Igic, B., R. Lande, and J. R. Kohn. 2008. Loss of self-incompatibility and its evolutionary consequences. Int. J. Plant Sci. 169:93–104.

Joly, S., and D. J. Schoen. 2011. Migration rates, frequency-dependent selection and the self-incompatibility locus in *Leavenworthia* (Brassicaceae). Evolution 65:2357–2369.

Kachroo A., C. R. Schopfer, M. E. Nasrallah, and J. B. Nasrallah. 2001. Allele-specific receptor-ligand interactions in *Brassica* self-incompatibility. Science 293:1824–1826.

Koelling V., J. L. Hamrick, and R. Mauricio. 2010. Genetic diversity and structure in two species of *Leavenworthia* with self-incompatible and self-compatible populations. Heredity 106:310–318.

Kondo, K., M. Yamamoto, R. Itahashi, T. Sato, H. Egashira, T. Hattori, and Y. Kowyama. 2002. Insights into the evolution of self-compatibility in *Lycopersicon* from a study of stylar factors. Plant J. 30:143–153.

Kusaba, M., and T. Nishio. 1999. Comparative analysis of S haplotypes with very similar *SLG* alleles in *Brassica rapa* and *B. oleracea*. Plant J. 17:83–92.

Lande, R., and D. W. Schemske. 1985. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. Evolution 39:24–40.

Librado, P., and J. Rozas. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452.

Liu, F., L. Zhang, and D. Charlesworth. 1998. Genetic diversity in *Leavenworthia* populations with different inbreeding levels. Proc. R. Soc. Lond. B 265:293–301.

Liu, P., S. Sherman-Broyles, M. E. Nasrallah, and J. B. Nasrallah. 2007. A cryptic modifier causing transient self-incompatibility in *Arabidopsis thaliana*. Curr. Biol. 17:734–740.

Lloyd, D. G. 1965. Evolution of self-compatibility and racial differentiation in *Leavenworthia* (Cruciferae). Contr. Gray Herb. Harv. Univ. 195:1–134.

McVean, G., P. Awadalla, and P. Fernhead. 2002. A coalescent-based method for detecting and estimating recombination from gene sequences. Genetics 160:1231–1241.

Miller, J. S., R. A. Levin, and N. M. Feliciano. 2008. A tale of two continents: Baker's rule and the maintenance of self-incompatibility in *Lycium* (Solanaceae). Evolution 62:1052–1065.

Murase, K. H., H. Shiba, M. Iwano, F. S. Che, M. Watanabe, A. Isogai, and S. Takayama. 2004. A membrane-anchored protein kinase involved in *Brassica* self-incompatibility signalling. Science 303:1516–1519.

Nishio, T., and M. Kusaba. 2000. Sequence diversity of *SLG* and *SRK* in *Brassica oleracea* L. Ann. Bot. 85:141–146.

Paape, T., and J. R. Kohn. 2011. Differential strengths of selection on S-RNases from *Physalis* and *Solanum* (Solanaceae). BMC Evol. Biol. 11:243.

Paape, T., B. Igic, S. D. Smith, R. Olmstead, L. Bohs, and J. R. Kohn. 2008. A 15-Myr-old genetic bottleneck. Mol. Biol. Evol. 25:655–663.

Posada, D. 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25:1253–1256.

Ritland, K. 2002. Extensions of models for the estimation of mating systems using *n* independent loci. Heredity 88:221–228.

Rollins, R.C. 1963. The evolution and systematics of *Leavenworthia* (Cruciferae). Contr. Gray Herb. Harv. Univ. 192:1–98.

Schierup, M. H., X. Vekemans, and F. B. Christiansen. 1997. Evolutionary dynamics of sporophytic self-incompatibility alleles in plants. Genetics 147:835–846.

———. 1998. Allelic genealogies in sporophytic self-incompatibility systems in plants. Genetics 150:1187–1198.

Schierup, M. H., J. S. Bechsgaard, L. H. Nielsen, and F. B. Christiansen. 2006. Selection at work in self-incompatible Arabidopsis lyrata: mating patterns in a natural population. Genetics 172:477–484.

Schoen, D. J., and A. H. D. Brown. 1991. Intraspecific variation in population gene diversity and effective population size correlates with mating system in plants. Proc. Natl. Acad. Sci. U.S.A. 88:4494–4497.

Schoen, D. J., and J. W. Busch. 2008. On the evolution of self-fertilization in a metapopulation. Int. J. Plant. Sci. 169:119–127.

Stebbins, G. L. 1974. Flowering plants: evolution above the species level. Belknap Press of Harvard Univ., Cambridge, MA.

Swanson, W. J., R. Nielsen, and Q. Yang. 2003. Pervasive adaptive evolution in mammalian fertilization proteins. Mol. Biol. Evol. 20:18–20.

Tantikanjana, T., M. E. Nasrallah, and J. B. Nasrallah. 2010. Complex networks of *self*-incompatibility signaling in the Brassicaceae. Curr. Opin. Plant Biol. 13:520–526.

Tsuchimatsu, T., K. Suwabe, R. Shimizu-Inatsugi, S. Isokawa, P. Pavlidis, T. Stadler, G. Suzuki, S. Takayama, M. Watanabe, and K. K. Shimizu. 2010. Evolution of self-compatibility in *Arabidopsis* by a mutation in the male specificity gene. Nature 464:1342–1346.

Uyenoyama, M. K. 1997. Genealogical structure among alleles regulating self-incompatibility in natural populations of flowering plants. Genetics 147:1389–1400.

Wright, S. 1939. The distribution of self-sterility alleles in populations. Genetics 169:2255–2265.

Wright, S. I., R. W. Ness, J. P. Foxe, and S. C. H. Barrett. 2008. Genomic consequences of outcrossing and selfing in plants. Int. J. Plant Sci. 169:105–118.

Yang, Z. 2007. PAML 4: a program package for phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24:1586–1591.

Yang, Z., W. S. W. Wong, and R. Nielsen. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. Mol. Biol. Evol. 22:1107–1118.

Associate Editor: J. Vamosi

Supporting Information

The following supporting information is available for this article:

Figure S1. Majority-rule (50%) consensus phylogeny for *Leavenworthia SRK* alleles with statistical support for branches.

Figure S2. Majority-rule (50%) consensus phylogeny for all *Arabidopsis, Brassica, C. grandiflora,* and *Leavenworthia SRK* alleles used in this study.

Table S1. Locality information for species and populations of *Leavenworthia* used in this study.

Table S2. GenBank accession numbers for *SRK* alleles from the Brassicaceous taxa used in this study.

Table S3. Likelihood ratio test results and significance levels for comparison of the positive selection model M8 with alternative general models (M7 and M8a).

Supporting Information may be found in the online version of this article.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.