

Patterns of Variation Within Self-Incompatibility Loci

Naoki Takebayashi,* Philip B. Brewer,† Ed Newbigin,† and Marcy K. Uyenoyama*

*Department of Biology, Duke University; †School of Botany, University of Melbourne, Victoria, Australia

Diverse self-incompatibility (SI) mechanisms permit flowering plants to inhibit fertilization by pollen that express specificities in common with the pistil. Characteristic of at least two model systems is greatly reduced recombination across large genomic tracts surrounding the *S*-locus, which regulates SI. In three angiosperm families, including the Solanaceae, the gene that controls the expression of gametophytic SI in the pistil encodes a ribonuclease (*S*-RNase). The gene that controls pollen SI expression is currently unknown, although several candidates have recently been proposed. Although each candidate shows a high level of polymorphism and complete allelic disequilibrium with the *S*-RNase gene, such properties may merely reflect tight linkage to the *S*-locus, irrespective of any functional role in SI. We analyzed the magnitude and nature of nucleotide variation, with the objective of distinguishing likely candidates for regulators of SI from other genes embedded in the *S*-locus region. We studied the *S*-RNase gene of the Solanaceae and *48A*, a candidate for the pollen gene in this system, and we also conducted a parallel analysis of the regulators of sporophytic SI in Brassica, a system in which both the pistil and pollen genes are known. Although the pattern of variation shown by the pollen gene of the Brassica system is consistent with its role as a determinant of pollen specificity, that of *48A* departs from expectation. Our analysis further suggests that recombination between *48A* and *S*-RNase may have occurred during the interval spanned by the gene genealogy, another indication that *48A* may not regulate SI expression in pollen.

Introduction

Many hermaphroditic plants throughout the angiosperms express self-incompatibility (SI), a genetic barrier to self-fertilization. Among homomorphic SI systems, in which different mating types have similar floral morphologies, several evolutionarily distinct mechanisms serve to cause the rejection of pollen by pistils when both express the same specificity. Pollen specificity is determined by the *S*-locus genotype of the haploid gametophyte in gametophytic SI (GSI) systems and of the diploid sporophyte in sporophytic SI (SSI) systems. Analysis of mutational experiments in a number of SI systems has established a bipartite model for the *S*-locus, with distinct genes encoding pistil and pollen specificities (Golz, Clarke, and Newbigin 2000; Nasrallah et al. 2000).

In the SSI system of Brassica, pollen specificity is vested in a small, cysteine-rich protein embedded in the outer coat of pollen grains (Schopfer, Nasrallah, and Nasrallah 1999) and pistil specificity resides in a receptor protein kinase bound in the plasma membrane of epidermal stigmatic cells (Takayama et al. 2001). A functional ribonuclease (*S*-RNase) determines pistil specificity in the solanaceous GSI system (Lee, Huang, and Kao 1994; Murfett et al. 1994). *S*-RNase genes also cosegregate with *S*-loci within the Rosaceae (Sassa, Hirano, and Ikehashi 1992) and Scrophulariaceae (Xue et al. 1996). Phylogenetic analyses of the sequence and structure of the *S*-RNase genes indicate a common origin of GSI in these distantly related families (Igic and Kohn 2001; Steinbachs and Holsinger 2002). Identification of the gene, provisionally designated *pollen-S*, that controls pollen specificity in *S*-RNase-based GSI systems remains a prime objective. Here we assess a candidate for *pollen-S* by analyzing the pattern of nucleotide variation.

Key words: self-incompatibility, *S*-locus, *S*-RNase.

E-mail: marcy@duke.edu.

Mol. Biol. Evol. 20(11):1778–1794, 2003

DOI: 10.1093/molbev/msg209

Molecular Biology and Evolution, Vol. 20, No. 11,

© Society for Molecular Biology and Evolution 2003; all rights reserved.

Recombination Suppression Within the *S*-Locus

Suppressed recombination between the pollen and pistil genes is considered essential to SI function, as crossing-over would presumably generate a recombinant haplotype that fails to reject the specificity expressed by its own pollen. In the SSI system of Brassica, the extensive structural differences observed among segregating *S*-alleles (Boyes et al. 1997; Suzuki et al. 1999; Nasrallah 2000) would appear to preclude the generation of functional haplotypes by conventional crossing-over in a region spanning perhaps hundreds of kilobases.

Recent studies have greatly improved the resolution of the *S*-locus region in *S*-RNase-based GSI (Li et al. 2000; Dowd et al. 2000; Lai et al. 2002). As in the SSI system of Brassica (Suzuki et al. 1999), the abundance of repetitive sequences in regions flanking the pistil-expressed *S*-RNase gene in the Solanaceae (Coleman and Kao 1992; Matton et al. 1995), Rosaceae (Ushijima, Sassa, and Hirano 1998), and Scrophulariaceae (Lai et al. 2002) suggests suppressed recombination. In solanaceous species, the *S*-locus lies in a centromeric region in which recombination appears to be suppressed over at least one megabase and possibly several (McCubbin and Kao 1999; McCubbin, Wang, and Kao 2000). In *Brassica rapa*, the *S*-locus region is gene-rich, containing approximately one expressed gene per 5.4 kb (Suzuki et al. 1999). If gene density within the *S*-RNase-based GSI *S*-locus is also of this order, hundreds or even thousands of genes may exist for which recombination with *S*-RNase would be undetectable in samples of practicable size. Indeed, McCubbin, Wang, and Kao (2000) have isolated several pollen-expressed genes near *S*-RNase in *Petunia inflata* (Solanaceae), and Lai et al. (2002) predicted 11 genes within a 63-kb region surrounding a functional *S*-RNase allele of *Antirrhinum hispanicum* (Scrophulariaceae). A number of floral traits involved in pollination biology have been shown to be linked to the *Lycopersicon hirsutum* *S*-locus (Bernacchi and Tanksley 1997).

Table 1
Solanaceous *S*-Locus Sequences

Species	<i>S</i> -RNase	48A
<i>Lycopersicon peruvianum</i>	<i>S</i> ₃ (X76065)	
	<i>S</i> ₅ (S61768)	
	<i>S</i> ₆ (Z26583)	
	<i>S</i> ₇ (Z26582)	
	<i>S</i> ₁₁ (U28796)	
	<i>S</i> ₁₂ (U28795)	
	<i>S</i> ₁₃ (D17325)	
<i>Nicotiana alata</i>	<i>S</i> _{5r} ^a	
	<i>S</i> ₇₁₁ ^b	48A-1 (AY159325)
	<i>S</i> ₂ (U08860)	48A-2 (AY159326)
	<i>S</i> ₃ (U66427)	48A-3 (AY159327)
	<i>S</i> ₆ (U08861)	48A-6 (AY159328)
	<i>S</i> ₇ (U13255)	48A-7 (AY159329)
	<i>S</i> _{A2} (U45957)	
	<i>S</i> _{MS2} (D63888)	
<i>Petunia axillaris</i>	<i>S</i> ₁ (AF239908)	
	<i>S</i> ₁₃ (AF239909)	
	<i>S</i> ₁₅ (AF239910)	
<i>Petunia inflata</i>	<i>S</i> ₁ (M67990)	
<i>Solanum chacoense</i>	<i>S</i> ₂ (X56896)	
	<i>S</i> ₃ (X56897)	
	<i>S</i> ₁₁ (S69589)	
	<i>S</i> ₁₂ (AF191732)	
	<i>S</i> ₁₄ (AF232304)	
<i>Solanum tuberosum</i>	<i>S</i> ₂ (X62727)	

^a Manually transcribed from Tsai et al. (1992).^b Kindly provided by Teh-hui Kao.

Candidates for *pollen-S*

Several genes isolated from the *S*-locus region of *S*-RNase-based GSI systems exhibit characteristics considered fundamental to *pollen-S*: pollen-specific expression, complete allelic linkage disequilibrium with *S*-RNase (a distinct restriction fragment length pattern for each *S*-haplotype examined), and no detectable recombination with *S*-RNase. These include 48A in *Nicotiana alata* (Solanaceae; Li et al. 2000), 13 genes in *P. inflata* (Solanaceae; McCubbin, Zuniga, and Kao 2000), and F-box protein genes in species of *Prunus* (Rosaceae; Entani et al. 2003; Ushijima et al. 2003).

A direct assessment of the involvement of a gene in SI function might entail examining whether transgenic plants show a change in SI expression (for example, Lee, Huang, and Kao 1994; Schopfer, Nasrallah, and Nasrallah 1999). However, with respect to the *pollen-S* gene in an *S*-RNase-based system, the analysis of such experiments may not be straightforward, as it would rely on an interpretation of the complex pollination phenotype seen in self-compatible mutant plants that carry an extra copy of all or part of an *S*-allele (Golz, Clarke, and Newbigin 1999; Golz et al. 2001). It is not known whether an appropriately engineered *pollen-S* gene construct would produce this phenotype when introduced into the genome of a self-incompatible host plant, or indeed whether this phenotype

is displayed by all plants with an extra copy of an *S*-allele. Moreover, a transgenic approach may be impossible in some species, either because genetic transformation is far from routine or because plants must be grown for long periods before starting to flower.

We here present an analysis of nucleotide sequence variation designed to assess whether 48A, derived from *N. alata* (Li et al. 2000), is a likely candidate for the determinant of pollen specificity in the *S*-RNase-based system of GSI. Results of a parallel analysis of the SSI system of Brassica, for which both the pollen and stigmatic determinants are known, appear to verify our interpretations.

First, we applied maximum-likelihood (ML) methods developed by Yang and colleagues to identify targets of positive selection in both systems. Second, we examined whether the pollen-expressed gene and the *S*-locus have had a common evolutionary history. Indications of recombination include departures from proportionality of the lengths of corresponding branches in the genealogical trees of the two genes and differences in the numbers of segregating sites. Because these indicators would be influenced by the acceleration by balancing selection of substitution at specificity-determining sites (Maruyama and Nei 1981; Takahata 1990; Sasaki 1992), we based our inferences of genealogical history on the pattern of synonymous or untranslated variation in the partition remaining after the removal of all sites that showed even weak indications of positive selection.

Our analysis confirmed that the determinant of pollen specificity in the Brassica SSI system fulfills the expectations of positive selection and common history with the pistil gene. In contrast, 48A showed no evidence of positive selection; further, the highly significant reduction in the number of segregating neutral mutations relative to *S*-RNase suggests historical recombination between the genes. We conclude that 48A may not correspond to the determinant of pollen specificity in the *S*-RNase-based GSI system.

Materials and Methods

Nucleotide Sequences

Table 1 provides GenBank accession numbers for the sequences studied. The 48A gene transcripts were obtained by reverse transcriptase polymerase chain reaction (RT-PCR) from *N. alata* pollen RNA (Brewer 1998). The near full-length sequences, 48A-1, 48A-2, 48A-3, 48A-6, and 48A-7, generated from lines maintained as homozygotes for *S*-alleles *S*₇₁₁, *S*₂, *S*₃, *S*₆, and *S*₇, respectively, comprised the entire coding region (87 codons) and the 3' UTR (236 bp). Reverse transcription of pollen RNA was primed using an oligo dT adapter (5' CTGAG AGAAC TAGTC TCGAG CTCTA GAACA AGCTT TTTT TTTT TTTT), and the 48A sequences were amplified from this template by the PCR using a 48A-specific primer 48A-F (5' GTAAA ATGTC ACCAA TAGTG CAC) and AD2 (5' TCTCG AGCTC TAGAA CAAGC), which targeted the adapter portion of the oligo dT adapter. Amplified DNA fragments were subsequently isolated from agarose gels and cloned into pGEM-T (Promega) for sequencing.

To identify all sites within solanaceous *S-RNase* genes that may contribute to SI, we expanded our analysis of *S-RNase* sequences beyond those within the five haplotypes for which *48A* sequences were determined. Of the more than 150 solanaceous *S-RNase* sequences available in GenBank, we restricted consideration to the few dozen that include most of the coding region. From these, we excluded the five sequences derived from *Petunia hybrida*, an artificial hybrid formed during the 19th century for which the *S*-locus has been mapped to two different chromosomes (ten Hoopen et al. 1998; Strommer et al. 2000). We also excluded all six *P. inflata* sequences in the two triads identified by Wang et al. (2001) as having been involved in recombination events. In total, we examined 25 *S-RNase* sequences (table 1). Within the five two-locus haplotypes, the shortest *S-RNase* sequence corresponds to *S*₃, which lacks the first 22 codons. We restricted our analysis to sites homologous to those in *S*₃; 214 codons, including gaps inserted to improve alignment (see figure 1).

We applied a parallel analysis to the genes which control the pollen (*SP11* or *SCR*; Schopfer, Nasrallah, and Nasrallah 1999; Takayama et al. 2001) and pistil (*SRK*; Takasaki et al. 2000) components of the sporophytic SI system expressed in Brassica. We excluded class II *S*-alleles, which tend to show recessive expression in pollen, because pollen-recessivity imposes different evolutionary constraints (Schierup, Vekemans, and Christiansen 1997; Uyenoyama 2000). We also excluded a *B. oleracea* *SRK* sequence (*SRK*_{13b}) which appears to encode a specificity functionally identical to another *B. oleracea* allele, despite nonsynonymous differences (Kusaba et al. 2000). Finally, we excluded two *SRK* sequences, one from *B. oleracea* (*SRK*₁₂) and one from *B. rapa* (*SRK*₄₇), which appear to have undergone some form of genetic exchange with the *SLG* gene within the respective haplotype (Sato et al. 2002); we retained the *SP11* alleles corresponding to both specificities.

Our analysis of variation within the Brassica SSI system examined a total of 15 *SRK* and 26 *SP11* class I alleles, including 11 haplotypes for which both the pistil and the pollen components were available (table 2). We restricted our study to the highly polymorphic ectodomain region of *SRK*, which determines pistil specificity, and the entire coding region of *SP11*, excluding the signal sequence. After adjustment of sequence length to permit comparison of all 11 two-locus haplotypes, our data set spanned 421 codons of exon 1 of *SRK* and 80 codons of *SP11*.

Maximum-Likelihood Methods

We used PAUP* 4.0b10 (Swofford 1999) to reconstruct ML genealogies for sequences derived from *S-RNase*, *48A*, *SRK*, and *SP11*, under the substitution models and parameter values indicated by Modeltest 3.04 (Posada and Crandall 1998). An exhaustive search produced genealogies based on the five complete *48A/S-RNase* haplotypes, and an heuristic search with tree bisection-reconnection branch swapping produced the remaining genealogies.

Yang and colleagues developed a ML framework for characterizing the process of substitution at specific codons or sites within one or more genomic regions, given a specified genealogy (Goldman and Yang 1994; Yang 1996; Nielsen and Yang 1998; Yang et al. 2000; Yang and Swanson 2002). Aspects of the substitution process examined include nonsynonymous/synonymous substitution rate ratio ($\omega = d_n/d_s$), transition/transversion substitution rate ratio (κ), codon frequencies (π), rate of substitution, and the shape parameter of gamma-distributed variation in substitution rate among sites. To identify possible targets of positive selection, individual codons can be assigned to one of a number of classes that differ, for example, with respect to the relative rates of synonymous and nonsynonymous substitution within a gene (Nielsen and Yang 1998; Yang et al. 2000). An alternative to this random-sites model is the fixed-sites model (Yang 1996; Yang and Swanson 2002), in which groups of sites may be defined by the user on the basis of prior information, including protein structure and function. This implementation permits comparisons of genomic regions between or within genes with respect to various aspects of the evolutionary process. Within this ML framework, comparisons of log-likelihood ratios provide the basis for statistical testing of nested evolutionary hypotheses (summarized in Table 2 of Yang 1996; Table 2 of Yang et al. 2000; Table 1 of Yang and Swanson 2002).

We used the CODEML module of PAML (Yang 1997) to conduct our codon-based analyses, which include random-sites models for the identification of targets of positive selection and fixed-sites comparisons among designated partitions. We used the BASEML module to compare the 3' UTR of *48A* and third codon position sites within *S-RNase* and *48A*.

Excluding Possible Targets of Positive Selection

We based our reconstruction of evolutionary history on variation at sites that show no evidence of positive selection. In an effort to exclude all sites subject to positive selection, we adopted a liberal definition of such sites.

Model 3 of Yang et al. (2000) estimates for each of a specified number of classes a value of the relative rate of nonsynonymous to synonymous substitution ($\omega = d_n/d_s$). Each codon triplet is assigned to the class for which the posterior probability of its having come from the class is highest. At one extreme, one might set the number of classes equal to the number of codon sites and designate any site with an ω value exceeding unity as a possible target of positive selection. In practice, we began with a one-class model (Model 0) and incremented the number of classes until the estimated ω values for all codons appeared to stabilize, with any additional class assigned an ω value identical to that of an existing class. We then adopted the first model (lowest number of classes) that gave the stable values, designating the union of all classes with ω values greater than unity as the positive selection partition and the complementary set of sites as the conservative evolution partition. Because our primary objective is to infer evolutionary history from the pattern

Table 2
Brassica S-Locus Sequences

Species	Haplotypes	SRK	SP11/SCR
<i>B. oleracea</i>	<i>SRK</i> ₃ (X79432); <i>SCR</i> ₃ (AJ278643)	<i>SRK</i> ₁₈ (AB032473)	<i>SCR</i> ₆ (AF195625)
	<i>SRK</i> ₇ (AB070624); <i>SP11</i> ₇ (AB070623)	<i>SRK</i> ₂₃ (AB013720)	<i>SP11</i> ₁₂ (AB050479)
	<i>SRK</i> ₁₃ (AB024419); <i>SCR</i> ₁₃ (AF195626)	<i>SRK</i> ₂₉ (Z30211)	
	<i>SRK</i> ₃₂ (AB050482); <i>SP11</i> ₃₂ (AB050480)	<i>SRK</i> ₆₀ (AB032474)	
<i>B. rapa</i>	<i>SRK</i> ₈ (D38563); <i>SP11</i> ₈ (AB035504)		<i>SP11</i> ₂₁ (AB039754)
	<i>SRK</i> ₉ (D88193); <i>SP11</i> ₉ (AB022078)		<i>SP11</i> ₂₆ (AB039755)
	<i>SRK</i> ₁₂ (D38564); <i>SP11</i> ₁₂ (AB035503)		<i>SP11</i> ₃₂ (AB039756)
	<i>SRK</i> ₄₅ (AB012106); <i>SP11</i> ₄₅ (AB039763)		<i>SP11</i> ₃₃ (AB039757)
	<i>SRK</i> ₄₆ (AB013717); <i>SP11</i> ₄₆ (AB070625)		<i>SP11</i> ₃₄ (AB039758)
			<i>SP11</i> ₃₆ (AB039759)
			<i>SP11</i> ₃₇ (AB039760)
			<i>SP11</i> ₃₈ (AB039761)
			<i>SP11</i> ₄₁ (AB039762)
			<i>SP11</i> ₄₇ (AB039765)
		<i>SP11</i> ₄₈ (AB039766)	
		<i>SP11</i> ₄₉ (AB039767)	
		<i>SP11</i> ₅₂ (AB035505)	
<i>B. napus</i>	<i>SRK</i> ₉₁₀ (M97667); <i>SCR</i> ₉₁₀ (AJ250857)		
	<i>SRK</i> _{A14} (U00443); <i>SCR</i> _{A14} (AJ250856)		

first partition and y_i the corresponding quantity for the second partition. To test whether the observed base or codon frequencies differ significantly between partitions, we determined the fraction of 10,000 random resamplings that showed a between-partition distance equal to or greater than the observed distance. This test may be generalized in various ways to accommodate more than two partitions.

Expression (1) was originally proposed as an index of distance between genetic samples obtained from different populations (Nei 1987, chapter 9). For x_i and y_i following Gaussian distributions, (1) follows a χ^2 distribution. Nei (1987) noted that the simple Euclidean distance between partitions has the undesirable property that completely complementary sets of bases or codons show only moderate distance. The weighting imposed by the denominator of equation (1) removes this feature.

We restricted likelihood comparisons to models that specified identical or different codon frequencies across partitions, in accordance with the results of the bootstrapping tests. Our analyses used the $F3 \times 4$ model, which determines codon frequencies as products of the base frequencies observed at the three codon positions.

A Test of Relative Numbers of Segregating Sites

For individual partitions within haplotypes, we estimated the total number of neutral substitutions which have occurred since the most recent common ancestor (MRCA). We developed the procedure outlined in the *Appendix* to compare the relative numbers of substitutions between two partitions within the sampled haplotypes. Our null hypothesis proposes absolute linkage and equal rates of substitution between partitions. Under this hypothesis, two partitions share a common genealogical history, with differences in substitution number attributed entirely to random events. We compared relative numbers of synon-

ymous substitutions between coding regions or substitutions between third codon positions of one gene and a noncoding region of another gene.

In accordance with the null hypothesis, we used PAUP* 4.0b10 to obtain an ML estimate of the joint genealogy from concatenated partitions and the CODEML module of PAML to estimate per-site substitution numbers for each branch of the joint genealogy. We then summed these estimates over the entire genealogy and multiplied by the number of sites to obtain the total number of substitutions occurring since the MRCA of the haplotypes (the number of segregating sites under the infinite sites model).

Under the standard assumption that the number of substitutions occurring within a specified time period follows a Poisson distribution, we tested whether the numbers of segregating sites in two partitions are consistent with a common value for the Poisson parameter. Rejection of the null hypothesis might reflect (1) different rates of substitution between partitions with a shared genealogical history or (2) identical substitution rates but different genealogies. In accordance with the first explanation, we estimated the Poisson parameter for each partition separately under the joint genealogy and used likelihood ratio tests to determine whether the parameters differ between partitions. In accordance with the second explanation, we obtained an ML genealogy for each partition separately and used estimates of the numbers of segregating sites to determine ML estimates of the Poisson parameters.

Results

Gene Genealogies

Modeltest 3.04 (Posada and Crandall 1998) indicated that the process of nucleotide substitution in the 25 *S-RNase* sequences studied corresponds most closely to the

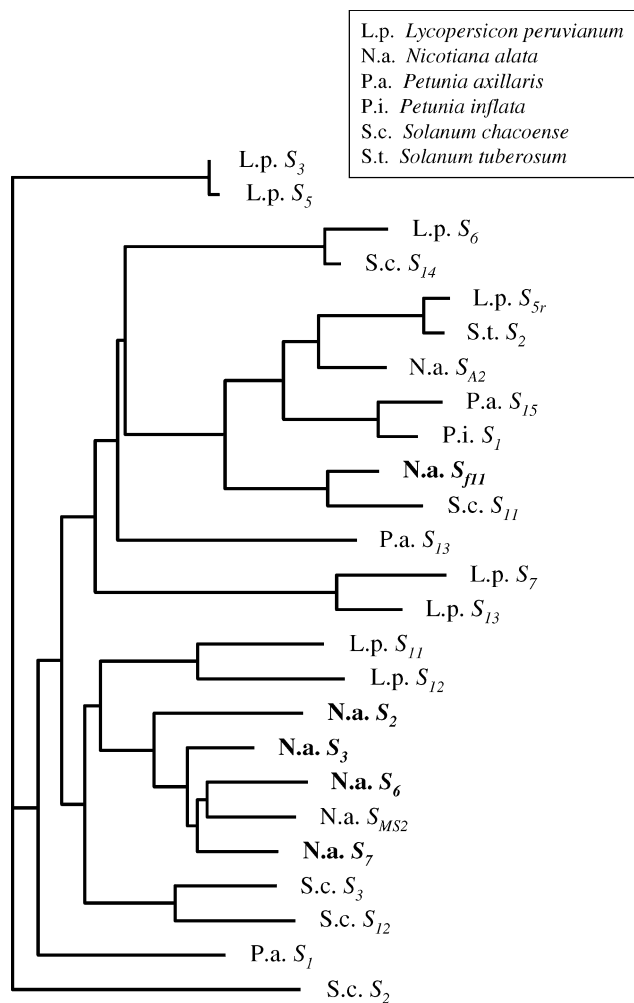


FIG. 2.—Maximum likelihood genealogy estimated from the *S-RNase* region of 25 haplotypes. Boldface indicates alleles for which the *48A* sequence borne on the same haplotype was obtained.

Tamura and Nei (1993) model with invariable sites (proportion 0.15) and gamma-distributed substitution rates (shape parameter 2.4). Figure 2 shows the ML genealogy obtained for *S-RNase* under this specification. The Felsenstein (1981) model was chosen for the joint coding region and 3' UTR of *48A*. Figure 3 compares the ML genealogy obtained by an exhaustive search using this specification to the genealogy estimated for the *S-RNase* regions borne on the same haplotypes. The genealogy estimated for the coding region of *48A* alone showed an identical topology.

Both the pistil (*SRK*) and pollen (*SP11*) components of the Brassica SSI system appear to have evolved under the Hasegawa-Kishino-Yano (HKY; 1985) model with gamma-distributed substitution rates (with shape parameter 0.5241 for *SRK* and 2.3171 for *SP11*). *SP11* included a proportion (0.1124) of invariable sites.

Positive Selection and Conservative Evolution Partitions

Figure 1 indicates the amino acid residues corresponding to codon sites assigned to the positive selection partition

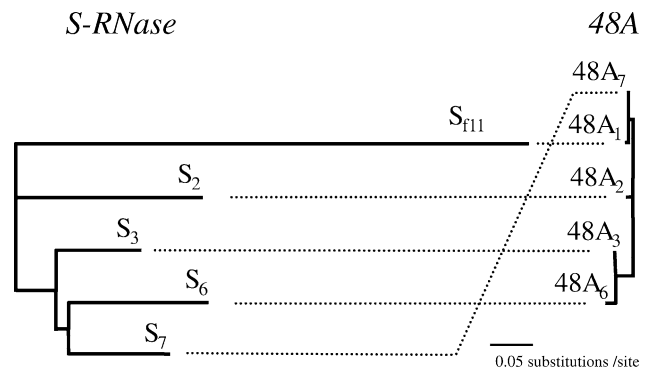


FIG. 3.—Topological differences between the genealogies estimated separately from the *S-RNase* and *48A* regions of the five available haplotypes.

of *S-RNase*. Stabilization of ω estimates and site assignments occurred at six classes, with no new ω values or site assignments indicated even in a model permitted 16 classes (table 3). Estimated ω values exceeded unity for two of the six classes, comprising 52 of the 214 codons studied. In contrast, no sites within *48A* were identified as possible targets of positive selection, under any assignment of class number from one through seven. Figure 1 compares the sites assigned to the positive selection partition of *S-RNase* with regions HVa and HVb, designated as hypervariable by Ioerger et al. (1991), and with regions PS1, PS2, PS3, and PS4, in which the sliding window analysis of Ishimizu et al. (1998) detected an excess of nonsynonymous substitutions in rosaceous *S*-alleles.

We applied a similar analysis to the pistil (*SRK*) and pollen (*SP11*) components of the Brassica SSI system (figs. 4 and 5). As expected, both genes showed evidence of positive selection. Sites within both *SRK* and *SP11* were assigned to five classes with distinct ω values (table 3); specification of 20 classes produced no new ω estimates for any site. The positive selection partition comprises over 26% (111 of 421) of the *SRK* codons studied and over 61% (49 of 80) of *SP11*. Of the 31 codons assigned to the conservative evolution partition of *SP11*, 18 represent indels that occur in only one sequence (*SP11*₂₆ of *B. rapa*; see fig. 5). Of the remaining 13 codons in this partition, one corresponds to an invariable Glu residue and seven to highly conserved Cys residues, only one of which is variable (replaced by Trp in two of the 26 *SP11* sequences studied).

Table 4 compares nucleotide diversity levels among various partitions, computed using MEGA (Kumar et al. 2001) with the Tamura and Nei (1993) distance correction and standard errors estimated by bootstrapping with 1,000 replications. *48A* shows much less diversity than *S-RNase*, even in the comparison of the 3' UTR of *48A* and the conserved partition of *S-RNase*. In contrast, the pollen-expressed *SP11* gene of the Brassica SSI system shows significantly higher diversity than the pistil-expressed *SRK*.

Base Composition and Codon Usage

Table 5 summarizes results of bootstrap comparisons of base composition or codon usage. Listed for each pair of

partitions is the fraction of 10,000 random resamplings that showed a between-partition distance (codon position 1) equal to or greater than the observed distance.

A comparison of the positive selection and conservative evolution partitions of *S-RNase* indicated similar codon usage; significant differences in base composition were found only in the first codon position. We detected no significant differences between partitions within *SRK*.

Base composition and codon usage in the coding region of *48A* differed highly significantly from those in *S-RNase*. In particular, the first and second codon positions of *48A* exhibited a very strong bias toward A and G in contrast with the third codon position and 3' UTR, and the third codon position was more G/C-rich than the 3' UTR. Differences in base composition between the 3' UTR of *48A* and the third codon positions of either *S-RNase* partition were nonsignificant.

Every comparison involving the conserved partition of *SP11* showed highly significant differences in both base composition and codon usage. Both aspects reflect the predominance of Cys residues (codons TGT and TGC).

Variation Within Pistil-Expressed *S-RNase*

We conducted fixed-site comparisons (Yang and Swanson 2002) of the positive selection and conservative evolution partitions of *S-RNase* (fig. 1). Model B of Yang and Swanson (2002) is described as nested within Model C, differing only in permitting different codon frequencies between partitions. However, our observation of a higher estimated likelihood of B confirms that B is not in fact nested within C.

Our bootstrap test of base composition (see *Materials and Methods*) offers an alternative method for comparing codon usage between genes or regions within genes. This analysis indicated significant differences between partitions at the first codon position but nonsignificant differences in codon usage (table 5). In accordance with the significant result, we adopted model F3 × 4 and permitted different codon frequencies between partitions in our fixed-site comparisons of the positive selection and conservative evolution partitions of *S-RNase*.

Table 6 summarizes the results of likelihood ratio comparisons among models. A fixed-site comparison between the positively selected and conserved partitions of *S-RNase* indicated significant differences in ω and κ (Model C versus Model E: $P = 5.7 \times 10^{-29}$). A model that treated the two partitions as separate data sets and permitted differences between partitions in substitution rate, codon frequencies, ω , and κ had significantly higher likelihood than one that assumed a common genealogy, indicating rejection of proportionality of lengths of corresponding branches between partitions, though at a much lower level of significance (E versus F: $P = 0.01$). This most general model (F) indicated a 4.3-fold higher substitution rate in the positive selection partition.

To examine further the nonproportionality of branch lengths, we compared just the third codon positions between the positive selection and conserved partitions. This analysis confirmed the significantly higher rate of substitution in the positive selection partition ($P = 2.3 \times$

Table 3
Class Assignments

System	Gene	ω (d_n/d_s)	Proportion of Sites
GSI	<i>S-RNase</i>	0.000	0.106
		0.079	0.178
		0.371	0.277
		0.778	0.203
		1.535	0.219
		2.089	0.018
SSI	<i>48A</i>	0.325	1.000
SSI	<i>SRK</i>	0.000	0.300
		0.417	0.396
		1.193	0.199
		2.983	0.083
		5.565	0.022
	<i>SP11</i>	0.000	0.098
		0.202	0.115
		1.220	0.241
		3.136	0.421
		4.677	0.124

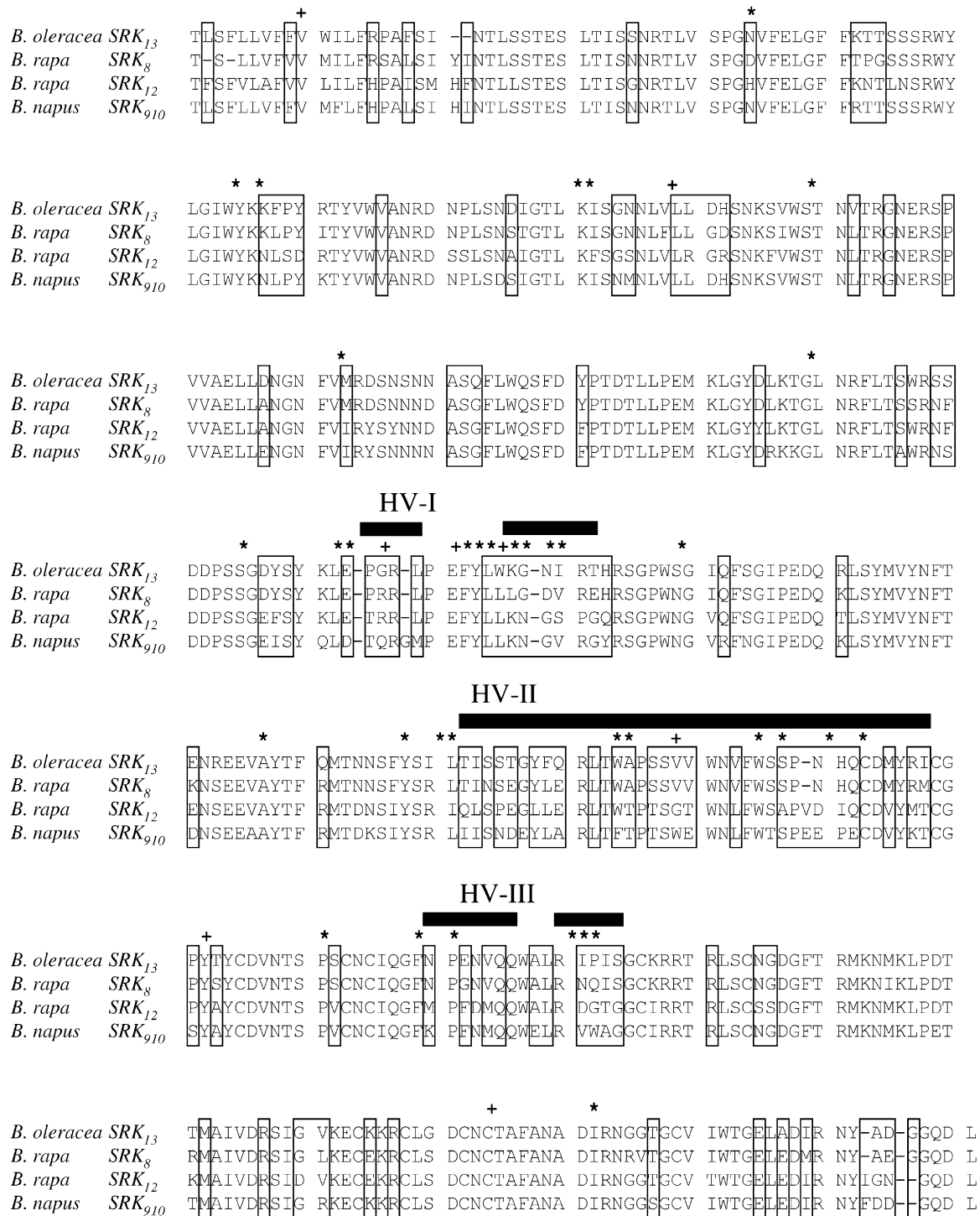
10^{-22}), although at a lower level of acceleration (table 6). In contrast, this comparison did not indicate non-proportional branch lengths, a finding consistent with a shared evolutionary history between the two *S-RNase* partitions.

Variation Within Pollen-Expressed *48A*

We conducted a fixed-site comparison between third codon positions and the 3' UTR of *48A*. Because our null hypothesis asserts absolute linkage between *S-RNase* and *48A*, we specified the *S-RNase* topology (fig. 2) for these partitions within *48A*. We also compared the partitions under the topology estimated from variation at *48A* alone.

Under either topology, comparison among models that assumed gamma-distributed substitution rates among sites within partitions and permitted different base frequencies between partitions indicated no significant differences between partitions in the shape parameter, rate of substitution, κ , or lengths of corresponding branches.

Under the topology determined from *48A* alone (fig. 3), a model that permitted differences between partitions in base composition and overall substitution rate but not in κ appeared to fit the data best. This model indicated an L-shaped distribution of substitution rate among sites within partitions (gamma-distribution shape parameter = 0.061). The log likelihood (lnL) of this best-fit model under the *48A* tree considerably exceeded that under the *S-RNase* tree ($2\Delta\ln L = 2(-541.5 + 564.4) = 45.8$). Because the topology hypotheses are not nested, this value cannot be directly tested against a χ^2 distribution; however, its magnitude provides another indication that the *48A* and *S-RNase* regions may have had distinct evolutionary histories. We tested this issue directly by comparing the numbers of segregating synonymous substitutions in the two genes (see next section).

FIG. 4.—Positive selection and conservative evolution partitions within *SRK*, with notations as in figure 1.

Between-Gene Comparisons

We first made codon-based comparisons between the entire coding region of *48A* and the conservative evolution or positive selection partition of *S-RNase*. We then compared the 3' UTR of *48A* to the third codon positions of each of the *S-RNase* partitions. Each comparison used the

topology estimated from the particular sequences being compared.

Codon-Based Analysis of *S-RNase* and *48A*

We detected highly significant differences in substitution rate between the conservative evolution partition

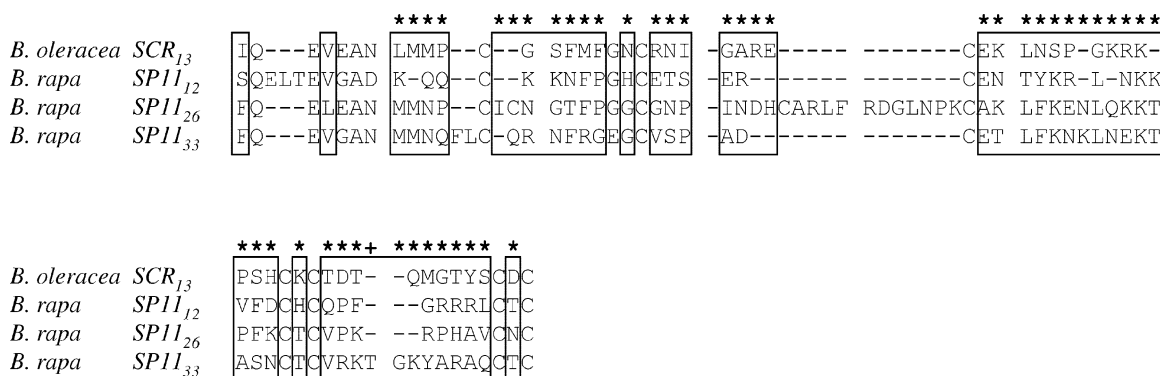


FIG. 5.—Positive selection and conservative evolution partitions within *SP11*, with notations as in figure 1.

of *S-RNase* and the coding region of *48A* ($P = 1.5 \times 10^{-57}$). A comparison of models that allowed differences in codon frequency (π) and substitution rate between partitions indicated nonsignificant differences in ω and κ . A test of proportionality of the numbers of substitutions in the *S-RNase* and *48A* partitions along corresponding branches indicated nonsignificant differences.

A comparison between the positive selection partition of *S-RNase* and the coding region of *48A* gave similar results: highly significant differences in substitution rate ($P = 1.2 \times 10^{-117}$) and proportional branch lengths. Although the ML estimates of ω appeared to differ strongly between genes (1.6 for *S-RNase* and 0.3 for *48A*), the difference was not significant. The low level of polymorphism within the coding region of *48A* (eight variable nucleotide sites among 261) suggests limited statistical power.

A general model, permitting differences between partitions in ω , κ , π , and substitution rate, indicated that substitution within the conserved partition of *S-RNase* proceeds at a rate nearly 19-fold higher than in the coding region of *48A*, with the positive selection partition of *S-RNase* showing even greater acceleration (nearly 80-fold).

3' UTR of *48A* and Third Codon Positions of *S-RNase*

We compared the more polymorphic 3' UTR of *48A* to third codon positions of the conserved partition of *S-RNase*. Comparison of models that assumed uniform substitution rate among sites within partitions and identical κ values between partitions indicated highly significant differences in substitution rate between partitions ($P = 3.5 \times 10^{-38}$). Corresponding branch lengths of the *S-RNase* and *48A* partitions showed significant departures from proportionality, suggesting differences in evolutionary history between the genes.

Comparison of the *48A* 3' UTR to the third codon positions of the positive selection partition of *S-RNase* gave similar results. We found a highly significant ($P = 2.7 \times 10^{-44}$) acceleration of substitution rate in *S-RNase*, even at the third codon position, and confirmed significant nonproportionality of corresponding branch lengths.

Relative Numbers of Substitutions in *S-RNase* and *48A*

Comparisons of the numbers of segregating sites in *S-RNase* and *48A* (table 7) indicated highly significant departures from expectation under the hypothesis of absolute linkage and equal rates of substitution between genes.

We first considered synonymous substitutions in the conservative evolution partition of *S-RNase* and in the entire coding region of *48A*. A model which permits different substitution rates between loci while preserving the assumption of absolute linkage provides a significantly better fit to the data than a model which specifies identical substitution rates. This model indicated a 16.8-fold higher substitution rate in *S-RNase*, comparable to the relative rate estimate of 18.9 obtained from the fixed-site comparison of these regions (table 6).

Alternatively, the apparent deficiency of segregating sites in *48A* may indicate a shallower genealogy, reflecting historical recombination. Comparison of the numbers of segregating sites estimated from independent genealogies for *S-RNase* and *48A* indicated 18.7-fold more synonymous substitutions in *S-RNase*. This result suggests a considerable reduction in total tree length under the

Table 4
Nucleotide Diversity

System	Partition	Diversity
GSI	<i>S-RNase</i>	0.527 (0.021 ^a)
	<i>S-RNase</i> P ^b	1.408 (0.038)
	<i>S-RNase</i> C ^c	0.396 (0.019)
	<i>48A</i>	0.015 (0.005)
	<i>48A</i> 3' UTR	0.033 (0.009)
SSI	<i>SRK</i>	0.141 (0.006)
	<i>SRK</i> P	0.383 (0.021)
	<i>SRK</i> C	0.074 (0.005)
	<i>SP11</i>	0.652 (0.053)
	<i>SP11</i> P	1.037 (0.055)
	<i>SP11</i> C	0.082 (0.027)

^a Bootstrap estimate of standard error (1,000 replications).

^b Positive selection partition.

^c Conservative evolution partition.

Table 5
Base Composition and Codon Usage

Comparison	Codon Position			Codon Usage	Noncodon
	1	2	3		
Within-gene					
<i>S-RNase</i> P ^a	<i>S-RNase</i> C ^b	0.0215	0.1398	0.3301	0.3288
<i>SRK</i> P	<i>SRK</i> C	0.0706	0.8027	0.1526	0.2195
<i>SP11</i> P	<i>SP11</i> C	0.0024	0.0202	0.3176	0 ^c
Between-gene					
<i>S-RNase</i> P	<i>48A</i>	0.0488	0	0	0
<i>S-RNase</i> C	<i>48A</i>	0	0	0	0
<i>S-RNase</i> P1 ^d	<i>48A</i> 3' UTR				0.0116
<i>S-RNase</i> P2	<i>48A</i> 3' UTR				0.0887
<i>S-RNase</i> P3	<i>48A</i> 3' UTR				0.6741
<i>S-RNase</i> C1	<i>48A</i> 3' UTR				0.4607
<i>S-RNase</i> C2	<i>48A</i> 3' UTR				0.6141
<i>S-RNase</i> C3	<i>48A</i> 3' UTR				0.0938
<i>48A3</i>	<i>48A</i> 3' UTR				0
<i>SRK</i> P	<i>SP11</i> P	0.3155	0.1684	0.5031	0.3815
<i>SRK</i> C	<i>SP11</i> C	0.0782	0.0498	0.0176	0.0010
<i>SRK</i> P	<i>SP11</i> C	0.0058	0.0040	0.0251	0
<i>SRK</i> C	<i>SP11</i> P	0.4130	0.2309	0.1078	0.3546

^a Positive selection partition.^b Conservative evolution partition.^c No observations in 10,000 resamplings.^d Codon position.

assumption that synonymous substitution proceeds at comparable rates in the two genes.

SSI Sequences

Fixed-site comparisons between the positively selected and conserved partitions of the pistil (*SRK*) and pollen (*SP11*) components of the Brassica SSI system confirmed significantly higher ω and substitution rate in the positively selected partition (table 6). The lengths of corresponding branches between partitions within *SRK* departed significantly from proportionality.

To examine further the nature of the nonproportionality of branch lengths, we compared third codon positions between the positive selection and conserved partitions of *SRK*. This analysis again indicated a highly significant, though smaller, acceleration in substitution rate. However, departures from proportionality of corresponding branch lengths became nonsignificant (table 6), consistent with the absence of recombination within the pistil gene.

Comparison of the positive selection partitions of the pollen (*SP11*) and pistil (*SRK*) genes indicated a significantly higher rate of substitution in *SP11* ($P = 3.7 \times 10^{-22}$) and significant departures from proportionality of corresponding branches. In contrast, no significant differences were detected between the conserved partitions of these genes. Although the finding of apparent proportionality of branch lengths is in accordance with our expectation of absolute linkage between the genes, it may also reflect low statistical power in comparisons involving the small conserved partition of *SP11*, which shows very little variability (six of 31 codons).

Restricting the between-gene comparison to the third codon position across the entire coding region confirmed a higher substitution rate in *SP11*. However, nonpro-

portionality of corresponding branch lengths remained significant, albeit at a lower level (table 6). Comparison of the third positions of the positive selection partitions between the pollen and pistil genes gave similar results, including significant nonproportionality of corresponding branch lengths. Comparison of the third codon positions in the conservative evolution partitions indicated no significant differences between the genes.

Discussion

Among the likely characteristics of *pollen-S*, the determinant of pollen specificity in *S-RNase*-based GSI systems, are very high levels of diversity and complete linkage disequilibrium with *S*-specificity. However, other genes embedded in the region of suppressed recombination that spans the *S*-locus may also exhibit such features, irrespective of any functional role in SI. We have described an analysis of nucleotide sequence variation designed to distinguish functional components of SI systems from the possibly great many other genes embedded within the massive *S*-locus region. Such studies may permit targeting of direct experimental study to the most likely candidate genes, with other genes serving as useful markers for the genetic dissection of regions tightly linked to targets of strong balancing selection.

Our study of the pattern of nucleotide variation at *48A* (Li et al. 2000) suggests that this candidate gene may not correspond to *pollen-S*. Three lines of evidence support this conclusion: (1) absence of indications of positive selection, (2) departures from the pattern exhibited by the pollen component of the system of SSI expressed in Brassica, and (3) apparent historical recombination with *S-RNase*, the pistil component of this GSI system.

Table 6
Fixed-Site Comparisons

Comparison	$\omega = d_n/d_s$		Relative Rate ^a	Nonproportional Branches ^b	
	Region 1	Region 2			
Within-gene					
<i>S-RNase P</i> ^c	<i>S-RNase C</i> ^d	1.4	0.3	4.3	0.01 ^e
<i>S-RNase P3</i> ^f	<i>S-RNase C3</i>			2.1	0.09
SRK P	SRK C	1.8	0.2	4.7	0.01
<i>SRK P3</i>	<i>SRK C3</i>			2.0	0.26
<i>SP11 P</i>	<i>SP11 C</i>	2.2	0.5	11.2	0.84
Between-gene					
<i>S-RNase P</i>	<i>48A</i>	1.4 ^g	1.4 ^g	81.8	1.00
<i>S-RNase C</i>	<i>48A</i>	0.3 ^g	0.3 ^g	18.9	0.75
<i>S-RNase P3</i>	<i>48A3</i>			32.2	0.85
<i>S-RNase C3</i>	<i>48A3</i>			18.7	0.15
<i>S-RNase P3</i>	<i>48A 3' UTR</i>			34.4	6.2×10^{-6}
<i>S-RNase C3</i>	<i>48A 3' UTR</i>			15.1	3.2×10^{-8}
SRK P	SP11 P	1.7 ^g	2.0 ^g	0.4	2.5×10^{-4}
SRK P3	SP11 P3			0.4	0.01
<i>SRK C</i>	<i>SP11 C</i>	0.3 ^g	0.3 ^g	1.0 ^h	0.70
<i>SRK C3</i>	<i>SP11 C3</i>			1.1 ^h	0.79
SRK3	SP113			0.3	0.01
SRK C3	SP11 P3			0.2	4.6×10^{-3}
<i>SRK P3</i>	<i>SP11 C3</i>			0.6	0.71

^a Ratio of substitution rates in region 1 and region 2.^b Approximate significance level from log likelihood ratio test.^c Positive selection partition.^d Conservative evolution partition.^e Boldface indicates significant differences in both relative rate and nonproportionality.^f Third codon position.^g Not significantly different between regions.^h Not significantly different from 1.

Distinguishing Between Selection and History

Our assessment addressed whether the candidate for *pollen-S* has been subject to positive selection and shows absolute linkage to the determinant of pistil specificity, both features considered essential to the function of the determinant of pollen specificity. Our examination of these two aspects used complementary regions within *S*-locus genes, the positive selection and conservative evolution partitions. Because we base inferences of historical recombination on the relative numbers of substitutions in the pistil and pollen genes, we require a view of history that is largely free from changes in substitution rate induced by selection.

Self-incompatibility engenders an intense form of balancing selection (Vekemans and Slatkin 1994), reflecting the transmission advantage of pollen genes that express rarer specificities. Balancing selection among functionally distinct specificities maintains large numbers of specificities and greatly expands coalescence times (Takahata 1990). In accordance with expectation, *S*-loci typically show extraordinary polymorphism (Lawrence 2000); furthermore, genes that determine pistil specificity in model systems of both GSI and SSI show very ancient divergence among lineages (Ioerger et al. 1990, Dwyer et al. 1991), in excess of 30 Myr in both the solanaceous GSI system (Ioerger, Clarke, and Kao 1990) and the Brassica SSI system (Uyenoyama 1995). Whereas balancing selection accelerates substitution at specificity-determining sites (Maruyama and Nei 1981; Takahata 1990;

Sasaki 1992), the rate of substitution of neutral variants reflects only the rate at which they arise, irrespective of linkage to sites subject to selection (Birky and Walsh 1988).

Hughes and Nei's (1988) seminal approach to the detection of positive selection used information obtained from crystal structure analyses of MHC molecules to identify a candidate target region and documented significantly higher ratios of nonsynonymous to synonymous substitutions in that region compared to other regions of the molecule. Ishimizu et al. (1998), in their sliding window analysis of rosaceous *S-RNase* genes, inverted this approach, identifying regions that showed significant excess nonsynonymous substitutions as candidates for targets of positive selection. We applied the Nielsen/Yang ML method (Nielsen and Yang 1998; Yang et al. 2000), which assigns individual codons to classes that differ with respect to the ratio of nonsynonymous to synonymous substitution rates.

To address the second objective, of assessing whether the pollen-expressed gene shares its evolutionary history with the pistil-expressed gene, we restricted our analysis to sites that showed no indications of positive selection. We assigned to the positive selection partition any site for which the posterior probability of having come from a class with a higher rate of nonsynonymous substitution than synonymous substitution ($\omega = d_n/d_s > 1$) exceeded that for other classes. We accepted even low posterior probabilities, because our intention was to make a conservative determination of the complementary set of

Table 7
Relative Numbers of Segregating Sites

Comparison		Numbers of Sites (<i>n</i>) and Substitutions (<i>x</i>)				One-Rate Model		Two-Rate Model	
Region 1	Region 2	<i>x</i> ₁	<i>n</i> ₁	<i>x</i> ₂	<i>n</i> ₂	θT	<i>P</i> ^a	$(\theta T)_1/(\theta T)_2$	<i>P</i> ^b
Joint genealogy									
<i>S-RNase</i> C3	48A 3' UTR	155.3	155.0	15.7	236.0	0.44	10 ⁻²⁶	15.06	10 ⁻⁴⁴
<i>S-RNase</i> C <i>d</i> _s ^c	48A <i>d</i> _s	201.3	113.1	2.3	21.8	1.51	10 ⁻¹²	16.77	10 ⁻¹³
<i>SRK</i> C <i>d</i> _s	<i>SP11</i> <i>d</i> _s	207.9	231.0	112.1	48.1	1.15	10 ⁻¹¹	0.39	10 ⁻¹⁴
Independent genealogies									
<i>S-RNase</i> C3	48A 3' UTR	153.9	155.0	15.7	236.0			14.97	
<i>S-RNase</i> C <i>d</i> _s	48A <i>d</i> _s	205.8	114.2	2.2	22.8			18.77	
<i>SRK</i> C <i>d</i> _s	<i>SP11</i> <i>d</i> _s	206.1	233.1	112.1	48.1			0.38	

^a Fit of one-rate Poisson model.^b Fit of one-rate Poisson model relative to two-rate Poisson model.^c Synonymous substitutions.

sites. Even within the conservative evolution partition, we further restricted consideration whenever possible to synonymous variation or variation at third codon positions.

We interpreted significant differences in numbers of segregating sites and departures from proportionality of corresponding branch lengths in the genealogies of the pistil- and pollen-expressed genes as evidence of historical recombination. Recombination with specificity-determining sites would reduce the expansion of coalescence times at neutral sites induced by balancing selection (Hudson and Kaplan 1988). Separate genealogical histories of different regions within haplotypes can also cause apparent differences in substitution rates along genealogical branches that would, in the absence of recombination between the regions, correspond to equal time intervals. Actual differences between regions in substitution rate may of course also account for differences in sequence divergence.

Selection Within the *S*-Locus Region

Figures 1, 4, and 5 indicate the positive selection and conservative evolution partitions of the pistil gene (*S-RNase*) of the solanaceous GSI *S*-locus and of the pistil (*SRK*) and pollen (*SP11*) components of the Brassica SSI *S*-locus. For each gene, fixed-site comparisons confirmed significantly greater ω values and higher overall rates of substitution in the positive selection partition than in the conservative evolution partition (table 6). We detected no indication of positive selection within *48A*.

Figure 1 shows the two regions (HV_a and HV_b) which Iøerger et al. (1991) designated as hypervariable within solanaceous *S-RNase* alleles and the four regions (PS1, PS2, PS3, and PS4) in which Ishimizu et al. (1998) detected an excess of nonsynonymous substitutions in rosaceous *S-RNase* alleles. Most of the sites assigned to a class under positive selection ($\omega > 1$) fall within these regions. The offset between PS4 and the three sites near the C-terminal with high posterior probability ($P > 0.95$) of assignment to a positive selection class may reflect some ambiguity, especially near the terminii, in inferring functional similarity or homology between the rosaceous

S-alleles studied by Ishimizu et al. (1998) and solanaceous *S*-alleles.

Ida et al. (2001) have solved the three-dimensional structure of *S-RNase* *S*_{f11} (*N. alata*), the first sequence shown in figure 1. All but one of the eight codons assigned with the highest posterior probability ($P > 0.99$) to a positive selection class fall within the HV and PS regions. The exceptional residue (Lys46, with numbering as shown in figure 1) lies on the surface of the *S*_{f11} protein, a position that does not exclude its possible involvement in allelic interactions. Ida et al. (2001) reported two clefts near the active site for RNase catalysis that are large enough to act as substrate binding sites. Of the 13 amino acids that make up these clefts, only one (Pro93) was assigned to a positive selection class with high posterior probability ($P > 0.95$).

In contrast with *S-RNase*, no sites within *48A* were assigned to the positive selection partition. Furthermore, our comparison of the relative numbers of segregating sites (table 7) indicates that *48A* has experienced significantly fewer synonymous and third position substitutions than *S-RNase*. None of our various analyses detected evidence of positive selection in the pollen-expressed *48A* gene, although low sample size (five) likely limited statistical power.

Although it lacks any indication of positive selection, the coding region of *48A* shows unusual base composition, including very strong bias toward A and G in the first and second codon positions in comparison to the third position and 3' UTR. A bootstrap comparison confirmed significant differences in base composition between the third codon position and the 3' UTR as well (table 5). Although the function of *48A* in the pollen grain is unknown, the high content of charged amino acids (30% Lys, 16% Asp, 14% Glu) and similarity to a group of equally hydrophilic proteins expressed by plant tissues undergoing some form of dessication (Thomashow 1999) suggest a possible role in preventing damage to pollen grain membranes. Mature *Nicotiana* pollen grains exist in an essentially dry state, reflecting a substantial withdrawal of water during the latter stages of development; upon hydration, a comparable amount of water flows back into the grain (Lush, Grieser, and Wolters-Arts 1998). In the absence of suitable protection, these fluxes of water into and out of the grain

would cause cellular membranes to rupture. Another member of this group of proteins, COR15a, significantly increases freezing tolerance of plastids when overexpressed in *Arabidopsis* (Artus et al. 1996), possibly by associating with the membrane and preventing freeze-induced phase transitions (Steponkus et al. 1998). Like desiccation, freezing damages plant tissues, initially through the removal of water and then, as tissues thaw, by the entry of water back into cells.

Consideration of the evolutionary conflicts that arise between the pistil or pollen components of SI, even under absolute linkage, suggests that the origin of new *S*-specificities may begin with a change in the pollen component (Uyenoyama and Newbiggin 2000; Uyenoyama, Zhang, and Newbiggin 2001), although other possibilities exist (Matton et al. 1999; 2000). Under any scenario, the maintenance of the highly specific recognition reaction between pollen and pistil would appear to require comparable rates of nonsynonymous change in the two components. Unlike *48A*, the pollen component of the Brassica *S*-locus bore out the expectation of positive selection. Almost all sites within *SP11* were assigned to the positive selection partition, with the exceptions corresponding primarily to highly conserved cysteine residues (fig. 5). Moreover, the overall rate of substitution in the positive selection partition of the pollen gene significantly exceeded that of the pistil gene (*SRK* P versus *SP11* P in table 6).

Evidence of Historical Recombination

Hughes (2000) compared closely related class II major histocompatibility complex genes of human and cyprinid fish, finding greater pairwise differences between exons than between introns. Furthermore, distances determined from exons and from introns within the cyprinid sequences showed a positive correlation only in between-genus comparisons and not in within-genus comparisons. He concluded that recombination had separated the genealogical histories of introns and exons within these genes.

In reconstructing evolutionary history, we restricted consideration to synonymous or noncoding substitutions in the conservative evolution partition, the region remaining after exclusion of all sites that showed any indication of positive selection. We chose to base our inference of historical recombination on departures from proportionality of lengths of corresponding branches of genealogies of linked regions and differences in numbers of segregating sites (see *Appendix*). Other methods for detecting recombination (Posada 2002), including, for example, those based on comparisons of topology, have very little power for small samples. We found (1) significant nonproportionality of branch lengths between the genealogies of the *S-RNase* and *48A* regions borne on the same haplotype (table 6) and (2) significant reductions in the number of segregating sites in *48A* (table 7; see also table 4). Both aspects are consistent with historical recombination between *48A* and *S-RNase*.

Nonproportionality of corresponding branch lengths

might also reflect variation in substitution rate over evolutionary time. The rate of substitution of specificity-determining mutations is determined by the rate of origin of new *S*-alleles, effective population size, and the number of *S*-alleles segregating in the population (Vekemans and Slatkin 1994; Uyenoyama and Takebayashi 2003). Fluctuations in such aspects over the dozens of millions of years spanned by *S*-allele genealogies may well have contributed to variation in the rate of selectively driven substitutions. In contrast, the rate of neutral substitution is expected to depend only on the frequency at which neutral variants arise.

In table 6, boldface indicates all fixed-site comparisons that showed both significant differences in substitution rate and significant nonproportionality of corresponding branch lengths. In both the GSI and SSI systems studied, comparisons between the positive selection and conservative evolution partitions of the pistil gene indicated a higher substitution rate in the positive selection partition and nonproportionality of branch lengths. The positively selected partition showed an accelerated substitution rate even at the third codon position, suggesting the operation of positive selection at this position. However, nonproportionality in branch length became nonsignificant upon restriction of consideration to the third codon position, in accordance with neutral expectation.

Variation in the rate of selectively driven substitution is unlikely to account for the highly significant nonproportionality detected in a fixed-site comparison between the 3' UTR of *48A* and the third codon position of the conservative evolution partition of *S-RNase*. In this case, we interpret the pattern as indicative of distinct genealogical histories of the *48A* and *S-RNase* regions within the same haplotype.

Although comparisons between the pistil and pollen components of the Brassica SSI system also showed significant differences in overall substitution rate and nonproportional branch lengths, the pattern differed from that observed between *S-RNase* and *48A* (tables 6 and 7). It is *SP11*, the determinant of pollen specificity, that showed the higher substitution rate. Also, because practically the entire coding region of *SP11* showed evidence of positive or purifying selection, any inferences concerning its evolutionary history must be qualified by the possibility of strong selection across the entire gene. Actual variation in substitution rate across time periods and among sites may account for nonproportionality of branch lengths between partitions of the Brassica *S*-locus.

Genetic exchange within the pistil gene *SRK* may also have contributed to nonproportionality of branch lengths, in spite of our efforts to minimize this effect by excluding from our analysis sequences for which evidence of genetic exchange is strong (see *Materials and Methods*). The *SRK* gene was first isolated by virtue of its high sequence similarity to *SLG*, which was formerly regarded as the determinant of pistil specificity (Stein et al. 1991). Observations of patches showing much greater sequence similarity in paralogous comparisons between *SRK* and *SLG* within the same haplotype than in orthologous comparisons between different haplotypes provide evi-

dence of some form of genetic exchange between the genes (Goring et al. 1993; Watanabe et al. 1994; Kusaba et al. 1997; Sato et al. 2002).

That genetic exchange within *S-RNase* (Wang et al. 2001) rather than between *S-RNase* and *48A* has contributed to the nonproportionality of branch lengths in the *S-RNase*-based system of GSI remains a possibility. To the extent that genetic exchange within *S-RNase* might have served as a mechanism for the generation of new, positively selected *S*-specificities, it might also have contributed to the acceleration of substitution in *S-RNase* relative to *48A*. However, a comparison of the third codon positions of the positive selection and conservative partitions of *S-RNase* indicated nonsignificant departures from proportionality of corresponding branch lengths (table 6). Furthermore, in combination with nonproportional branch lengths, *48A* showed highly significant reductions in number of synonymous substitutions relative to *S-RNase* (table 7). One might consider attributing the difference in synonymous substitutions to acceleration in *S-RNase* rather than to reduction in *48A*. Because neutral substitution depends on the rate of origin of neutral variants alone, and in particular because it is not influenced by linkage to selected sites (Birky and Walsh 1988), hitchhiking explanations would require the simultaneous creation of selected and neutral variants. Such coincident mutation might arise, for example, through genetic exchange of segments spanning both synonymous and non-synonymous sites. Even so, the absence of detectable positive selection within *48A* would appear to support our central conclusion that it may not correspond to *pollen-S*.

It is only comparisons involving the 3' UTR of *48A* that indicated significant nonproportionality with regions within *S-RNase* borne on the same haplotype (table 6). This pattern raises the possibility that genetic transfer events involving *48A* have been restricted to the 3' UTR. For example, proper control of the expression of *48A* throughout the very long periods over which *S*-haplotypes have been maintained might have been achieved not through successive substitution of *48A* alleles but rather through gene conversion targeted specifically to the 3' UTR. Even so, this explanation for nonproportionality of branch lengths between the 3' UTR of *48A* and *S-RNase* would not account for the highly significant reduction in substitution rate relative to *S-RNase* evident in the coding region of *48A* as well as the 3' UTR (table 7).

We suggest that incomplete linkage between *48A* and *S-RNase* provides the most parsimonious interpretation of our results. This conclusion challenges the hypothesis that *48A* serves as the primary determinant of pollen specificity in the solanaceous system of *S-RNase*-based GSI. Our estimate of a 15- to 20-fold reduction in genealogical depth in *48A* provides information about the rate of recombination between *48A* and *S-RNase* (cf. Strobeck 1980; Hudson and Kaplan 1988; Schierup, Charlesworth, and Vekemans 2000). We address this issue in a separate study.

Appendix

To test the hypothesis of absolute linkage to the *S*-locus, we compared estimates of the total number of

neutral substitutions segregating at two loci (partitions) within a sample of haplotypes, assuming an infinite-sites model.

Neutral mutations arise within locus i ($i = 0, 1$) at rate v_i per site per generation. Given the total time in the genealogy since the MRCA of the sampled genes, the number of substitutions within the genealogy follows a Poisson distribution with parameter

$$\lambda_i = \theta_i n_i T_i, \quad (\text{A.1})$$

for n_i the number of sites, T_i time in units of $2N$ generations, and $\theta_i = 4Nv_i$. Because substitutions occur independently across loci, the joint density corresponds to the product of the densities for the individual loci ($f(x_0, x_1) = f(x_0)f(x_1)$).

Identical Substitution Rates at Absolutely Linked Loci

We first addressed whether the relative divergence at neutral sites within locus 0 and locus 1 is consistent with absolute linkage and identical rates of neutral substitution. We obtained per-locus estimates of the number of segregating sites ($x_i = n_i d_i$ for d_i the estimated per-site distance) under an ML joint genealogy determined from sequences concatenated across loci. Absolute linkage entails a single total time with the genealogy for the two loci. Under the hypothesis $\{T_0 = T_1 = T, \theta_0 = \theta_1 = \theta\}$, equation A.1 reduces to

$$\lambda_i = n_i \theta T. \quad (\text{A.2})$$

The maximum likelihood estimate of θT , obtained by maximizing the probability of observing the estimated distances under equation A.2, corresponds to

$$\widehat{\theta T} = \frac{[x_0] + [x_1]}{n_0 + n_1}, \quad (\text{A.3})$$

in which the $[x_i]$ represent the per-locus estimates rounded to integer values.

To examine whether the observed distances are consistent with a common value of θT for the loci, we determined the approximate probability of states showing deviations equal to or greater than observed. Designating locus 1 as the one that shows fewer substitutions than expected, the probability is

$$P = \sum_{k=0}^{[x_1]} f_1(k), \quad (\text{A.4})$$

in which the Poisson density f_1 incorporates $\widehat{\theta T}$ (Eq. A.3).

Different Substitution Rates at Absolutely Linked Loci

For cases in which the hypothesis of a common total time and substitution rate across loci is rejected, we retained the estimate of x_i obtained under the joint genealogy but estimated λ_i for each locus separately. Under the hypothesis of a common total time but different substitution rates for the loci,

$$\lambda_i = n_i \theta_i T. \quad (\text{A.5})$$

The maximum likelihood estimate of λ_i now corresponds to

$$\hat{\lambda}_i = x_i. \quad (\text{A.6})$$

We compared the constrained (Eq. A.2) and unconstrained (Eq. A.5) models using likelihood ratio tests.

Incomplete Linkage

An alternative interpretation of between-locus differences in the numbers of segregating sites is that the loci share a common rate of substitution but have different genealogical histories, perhaps reflecting ancestral recombination events. Accordingly, we estimated for each locus separately an independent genealogy and the number of segregating sites (x_i). Under this model, estimates obtained from (equation A.6) have the interpretation

$$\hat{\lambda}_i = n_i \hat{\theta} T_i. \quad (\text{A.7})$$

Acknowledgments

We thank T.-h. Kao for providing the S_{f11} cDNA sequence, T. Nishio for permission to cite work in press, Z. Yang for assistance with PAML, and S. Palumbi and two anonymous reviewers for comments. E.N. receives funding from the Australian Research Council. U.S. Public Health Service grant GM 37841 to M.K.U. provided support for this study.

Literature Cited

- Artus, N. N., M. Uemura, P. L. Steponkus, S. J. Gilmour, C. T. Lin, and M. F. Thomashow. 1996. Constitutive expression of the cold-regulated *Arabidopsis thaliana* *COR15a* gene affects both chloroplast and protoplast freezing tolerance. *Proc. Natl. Acad. Sci. USA* **93**:13404–13409.
- Bernacchi, D., and S. D. Tanksley. 1997. An interspecific backcross of *Lycopersicon esculentum* \times *L. hirsutum*: linkage analysis and a QTL study of sexual compatibility factors and floral traits. *Genetics* **147**:861–877.
- Birky, C. W., and J. B. Walsh. 1988. Effects of linkage on rates of molecular evolution. *Proc. Natl. Acad. Sci. USA* **85**:6414–6418.
- Boyes, D. C., M. E. Nasrallah, J. Vrebalov, and J. B. Nasrallah. 1997. The self-incompatibility (S) haplotypes of *Brassica* contain highly divergent and rearranged sequences of ancient origin. *Plant Cell* **9**:237–247.
- Brewer, P. 1998. The statistical analysis of a gene tightly linked to the self-incompatibility locus. B.Sc. Honours thesis, University of Melbourne, Australia.
- Coleman, C. E., and T.-h. Kao. 1992. The flanking regions of 2 *Petunia inflata* S-alleles are heterogeneous and contain repetitive sequences. *Plant Mol. Biol.* **18**:725–737.
- Dowd, P. E., A. G. McCubbin, X. Wang, J. A. Verica, T. Tsukamoto, T. Ando, and T.-h. Kao. 2000. Use of *Petunia inflata* as a model for the study of solanaceous type self-incompatibility. *Ann. Bot.* **85**(Suppl. A):87–93.
- Dwyer, K. G., M. A. Balent, J. B. Nasrallah, and M. E. Nasrallah. 1991. DNA sequences of self-incompatibility genes from *Brassica campestris* and *B. oleracea*: polymorphism predating speciation. *Plant Mol. Biol.* **16**:481–486.
- Entani, T., M. Iwano, H. Shiba, F.-S. Che, A. Isogai, and S. Takayama. 2003. Comparative analysis of the self-incompatibility (S-) locus region of *Prunus mume*: identification of a pollen-expressed F-box gene with allelic diversity. *Genes Cells* **8**:203–213.
- Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17**:368–376.
- Goldman, N., and Z. Yang. 1994. A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol. Biol. Evol.* **11**:725–736.
- Golz, J. F., A. E. Clarke, and E. Newbigin. 2000. Mutational approaches to the study of self-incompatibility: revisiting the pollen-part mutants. *Ann. Bot.* **85**(Suppl. A):95–103.
- Golz, J. F., H. Y. Oh, V. Su, M. Kusaba, and E. Newbigin. 2001. Genetic analysis of *Nicotiana* pollen-part mutants is consistent with the presence of an S-ribonuclease inhibitor at the S locus. *Proc. Natl. Acad. Sci. USA* **98**:15372–15376.
- Golz, J. F., V. Su, A. E. Clarke, and E. Newbigin. 1999. A molecular description of mutations affecting the pollen component of the *Nicotiana glauca* S locus. *Genetics* **152**:1123–1135.
- Goring, D. R., T. L. Glavin, U. Schafer, and S. J. Rothstein. 1993. An S receptor kinase gene in self-compatible *Brassica napus* has a 1-bp deletion. *Plant Cell* **5**:531–539.
- Hasegawa, M., K. Kishino, and T. Yano. 1985. Dating the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**:160–174.
- Hudson, R. R., and N. L. Kaplan. 1988. The coalescent process in models with selection and recombination. *Genetics* **120**:831–840.
- Hughes, A. L. 2000. Evolution of introns and exons of class II major histocompatibility complex genes of vertebrates. *Immunogenetics* **51**:473–486.
- Hughes, A. L., and M. Nei. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature* **335**:167–170.
- Ida, K., S. Norioka, M. Yamamoto, T. Kumasaka, E. Yamashita, E. Newbigin, A. E. Clarke, F. Sakiyama, and M. Sato. 2001. The 1.55 Å resolution structure of *Nicotiana glauca* S_{F11} -RNase associated with gametophytic self-incompatibility. *J. Mol. Biol.* **314**:103–112.
- Igic, B., and J. R. Kohn. 2001. Evolutionary relationships among self-incompatibility RNases. *Proc. Natl. Acad. Sci. USA* **98**:13167–13171.
- Ioerger, T. R., A. G. Clark, and T.-h. Kao. 1990. Polymorphism at the self-incompatibility locus in Solanaceae predates speciation. *Proc. Natl. Acad. Sci. USA* **87**:9732–9735.
- Ioerger, T. R., J. R. Gohlke, B. Xu, and T.-h. Kao. 1991. Primary structural features of the self-incompatibility protein in Solanaceae. *Sex. Plant Reprod.* **4**:81–87.
- Ishimizu, T., T. Endo, Y. Yamaguchi-Kabata, K. T. Nakamura, F. Sakiyama, and S. Norioka. 1998. Identification of regions in which positive selection may operate in S-RNase of Rosaceae: implication for S-allele-specific recognition sites in S-RNase. *FEBS Lett.* **440**:337–342.
- Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei. 2001. MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* **17**:1244–1245.
- Kusaba, M., M. Matsushita, K. Okazaki, Y. Satta, and T. Nishio. 2000. Sequence and structural diversity of the S locus genes from different lines with the same self-recognition specificities in *Brassica oleracea*. *Genetics* **14**:413–420.
- Kusaba, M., T. Nishio, Y. Satta, K. Hinata, and D. Ockendon. 1997. Striking sequence similarity in inter- and intra-specific comparisons of class I *SLG* alleles from *Brassica oleracea* and *Brassica campestris*: implications for the evolution and

- recognition mechanism. *Proc. Natl. Acad. Sci. USA* **94**:7673–7678.
- Lai, Z., W. S. Ma, B. Han, L. Z. Liang, Y. S. Zhang, G. F. Hong, and Y. B. Xue. 2002. An F-box gene linked to the self-incompatibility (*S*) locus of *Antirrhinum* is expressed specifically in pollen and tapetum. *Plant Mol. Biol.* **50**:29–42.
- Lawrence, M. J. 2000. Population genetics of the homomorphic self-incompatibility polymorphisms in flowering plants. *Ann. Bot.* **85**(Suppl. A):221–226.
- Lee, H. S., S. S. Huang, and T.-h. Kao. 1994. *S*-proteins control rejection of incompatible pollen in *Petunia inflata*. *Nature* **367**:550–563.
- Lush, W. M., F. Grieser, and M. Wolters-Arts. 1998. Directional guidance in *Nicotiana glauca* pollen tubes in vitro and on the stigma. *Plant Physiol.* **118**:733–741.
- Li, J. H., N. Nass, M. Kusaba, P. N. Dodds, N. Treloar, A. E. Clarke, and E. Newbigin. 2000. A genetic map of the *Nicotiana glauca* *S* locus that includes three pollen-expressed genes. *Theor. Appl. Genet.* **100**:956–964.
- Maruyama, T., and M. Nei. 1981. Genetic variability maintained by mutation and overdominant selection in finite populations. *Genetics* **98**:441–459.
- Matton, D. P., D. T. Luu, D. Morse, and M. Cappadocia. 2000. Reply: establishing a paradigm for the generation of new *S*-alleles. *Plant Cell* **12**:312–315.
- Matton, D. P., D. T. Luu, Q. Xike, G. Laublin, M. O'Brien, O. Maes, D. Morse, and M. Cappadocia. 1999. 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, D. P., S. L. Mau, S. Okamoto, A. E. Clarke, and E. Newbigin. 1995. The *S*-locus of *Nicotiana glauca*: genomic organization and sequence analysis of two S-RNase alleles. *Plant Mol. Biol.* **28**:847–858.
- McCubbin, A. G., and T.-h. Kao. 1999. The emerging complexity of self-incompatibility (*S*-) loci. *Sex. Plant Reprod.* **12**:1–5.
- McCubbin, A. G., X. Wang, and T.-h. Kao. 2000. Identification of self-incompatibility (*S*-) locus linked pollen cDNA markers in *Petunia inflata*. *Genome* **43**:619–627.
- McCubbin, A. G., C. Zuniga, and T.-h. Kao. 2000. Construction of a binary bacterial artificial chromosome library of *Petunia inflata* and the isolation of large genomic fragments linked to the self-incompatibility (*S*-) locus. *Genome* **43**:820–826.
- Murfett, J., T. L. Atherton, B. Mou, C. S. Gasser, and B. A. McClure. 1994. *S*-RNase expressed in transgenic *Nicotiana* causes *S*-allele-specific pollen rejection. *Nature* **367**:563–566.
- Nasrallah, J. B. 2000. Cell-cell signaling in the self-incompatibility response. *Curr. Opin. Plant Biol.* **3**:368–373.
- Nasrallah, M. E., M. K. Kandasamy, M.-C. Chang, Z. Stadler, S. Lim, and J. B. Nasrallah. 2000. Identifying genes for pollen-stigma recognition in crucifers. *Ann. Bot.* **85**(Suppl. A):125–132.
- Nei, M. 1987. *Molecular evolutionary genetics*. Columbia University Press, New York.
- Nielsen, R., and Z. Yang. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* **148**:929–936.
- Posada, D. 2002. Evaluation of methods for detecting recombination from DNA sequences: empirical data. *Mol. Biol. Evol.* **19**:708–717.
- Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**:817–818.
- Sasaki, A. 1992. The evolution of host and pathogen genes under epidemiological interaction. Pp. 247–263 in N. Takahata, ed. *Population paleo-genetics*. Japan Scientific Societies Press, Tokyo.
- Sassa, H., H. Hirano, and H. Ikehashi. 1992. Self-incompatibility-related RNases in styles of Japanese pear (*Pyrus serotina* Rehd). *Plant Cell Physiol.* **33**:811–814.
- Sato, K., T. Nishio, R. Kimura, M. Kusaba, T. Suzuki, K. Hatakeyama, D. J. Ockendon, and Y. Satta. 2002. Co-evolution of the *S*-locus genes, *SRK*, *SLG*, and *SP11/SCR*, in *Brassica oleracea* and *Brassica rapa*. *Genetics* **162**:931–940.
- Schierup, M. H., D. Charlesworth, and X. Vekemans. 2000. The effect of hitch-hiking on genes linked to a balanced polymorphism in a subdivided population. *Genet. Res. Camb.* **76**:63–73.
- Schierup, M. H., X. Vekemans, and F. B. Christiansen. 1997. Evolutionary dynamics of sporophytic self-incompatibility alleles in plants. *Genetics* **147**:835–846.
- Schopfer, C. R., M. E. Nasrallah, and J. B. Nasrallah. 1999. The male determinant of self-incompatibility in *Brassica*. *Science* **286**:1697–1700.
- Stein, J. C., B. Howlett, D. C. Boyes, M. E. Nasrallah, and J. B. Nasrallah. 1991. Molecular cloning of a putative receptor protein kinase gene encoded at self-incompatibility locus of *Brassica oleracea*. *Proc. Natl. Acad. Sci. USA* **88**:8816–8820.
- Steinbachs, J. E., and K. E. Holsinger. 2002. S-RNase-mediated gametophytic self-incompatibility is ancestral in eudicots. *Mol. Biol. Evol.* **19**:825–829.
- Steponkus, P. L., M. Uemura, R. A. Joeseeph, S. J. Gilmour, and M. F. Thomashow. 1998. Mode of action of the *COR15a* gene on the freezing tolerance of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **95**:14570–14575.
- Strobeck, C. 1980. Heterozygosity of a neutral locus linked to a self-incompatibility locus or a balanced lethal. *Evolution* **34**:779–788.
- Strommer, J., A. G. M. Gerats, M. Sanago, and S. J. Molnar. 2000. A gene-based RFLP map of petunia. *Theor. Appl. Genet.* **100**:899–905.
- Suzuki, G., N. Taki, T. Hirose, K. Fukui, T. Nishio, S. Takayama, A. Isogai, M. Watanabe, and K. Hinata. 1999. Genomic organization of the *S* locus: identification and characterization of genes in *SLG/SRK* region of *S9* haplotype of *Brassica campestris* (syn. *rapa*). *Genetics* **153**:391–400.
- Swofford, D. L. 1999. PAUP* phylogenetic analysis using parsimony (and other methods), Version 4.0b. Sinauer Associates, Sunderland, Mass.
- Takahata, N. 1990. A simple genealogical structure of strongly balanced allelic lines and trans-species evolution of polymorphism. *Proc. Natl. Acad. Sci. USA* **87**:2419–2423.
- Takasaki, T., K. Hatakeyama, G. Suzuki, M. Watanabe, A. Isogai, and K. Hinata. 2000. The *S* receptor kinase determines self-incompatibility in *Brassica* stigma. *Nature* **403**:913–916.
- Takayama, S., H. Shiba, M. Iwano, H. Shimosato, F.-S. Che, N. Kai, M. Watanabe, G. Suzuki, K. Hinata, and A. Isogai. 2001. The pollen determinant of self-incompatibility in *Brassica campestris*. *Proc. Natl. Acad. Sci. USA* **97**:1920–1925.
- Tamura, K., and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**:512–526.
- ten Hoopen, R., R. M. Harbord, T. Maes, N. Nanninga, and T. P. Robbins. 1998. The self-incompatibility (*S*) locus in *Petunia hybrida* is located on chromosome III in a region, syntenic for the Solanaceae. *Plant J.* **16**:729–734.
- Thomashow, M. F. 1999. Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**:571–599.
- Tsai, D.-S., H.-S. Lee, L. C. Post, K. M. Kreiling, and T.-h. Kao. 1992. Sequence of an S-protein of *Lycopersicon peruvianum*

- and comparison with other solanaceous S-proteins. *Sex. Plant Reprod.* **5**:256–263.
- Ushijima, K., H. Sassa, A. M. Dandekar, T. M. Gradziel, R. Tao, and H. Hirano. 2003. Structural and transcriptional analysis of the self-incompatibility locus of almond: identification of a pollen-expressed F-box gene with haplotype-specific polymorphism. *Plant Cell* **15**:771–781.
- Ushijima, K., H. Sassa, and H. Hirano. 1998. Characterization of the flanking regions of the *S-RNase* genes of Japanese pear (*Pyrus serotina*) and apple (*Malus × domestica*). *Gene* **211**:159–167.
- Uyenoyama, M. K. 1995. A generalized least-squares estimate for the origin of sporophytic self-incompatibility. *Genetics* **139**:975–992.
- . 2000. Evolutionary dynamics of self-incompatibility alleles in Brassica. *Genetics* **156**:351–359.
- Uyenoyama, M. K., and E. Newbigin. 2000. Evolutionary dynamics of dual-specificity self-incompatibility alleles. *Plant Cell* **12**:310–312.
- Uyenoyama, M. K., and N. Takebayashi. 2003. Genus-specific diversification of mating types, in R. Singh and M. K. Uyenoyama, eds. *The Evolution of Population Biology*. Cambridge University Press, New York (in press).
- Uyenoyama, M. K., Y. Zhang, and E. Newbigin. 2001. On the origin of self-incompatibility haplotypes: transition through self-compatible intermediates. *Genetics* **157**:1805–1817.
- Vekemans, X., and M. Slatkin. 1994. Gene and allelic genealogies at a gametophytic self-incompatibility locus. *Genetics* **137**:1157–1165.
- Wang, X., A. L. Hughes, T. Tsukamoto, T. Ando, and T.-h. Kao. 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.
- Watanabe, M., T. Takasaki, K. Toriyama, S. Yamakawa, A. Isogai, A. Suzuki, and K. Hinata. 1994. A high degree of homology exists between the protein encoded by *SLG* and the *S* receptor domain encoded by *SRK* in self-incompatible *Brassica campestris* L. *Plant Cell Physiol.* **35**:1221–1229.
- Xue, Y. B., R. Carpenter, H. G. Dickinson, and E. S. Coen. 1996. Origin of allelic diversity in *Antirrhinum S* locus RNases. *Plant Cell* **8**:805–814.
- Yang, Z. 1996. Maximum-likelihood models for combined analyses of multiple sequence data. *J. Mol. Evol.* **42**:587–596.
- . 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13**:555–556.
- Yang, Z., R. Nielsen, N. Goldman, and A.-M. K. Pedersen. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **155**:431–449.
- Yang, Z., and W. J. Swanson. 2002. Codon-substitution models to detect adaptive evolution that account for heterogeneous selective pressures among site classes. *Mol. Biol. Evol.* **19**:49–57.

Stephen Palumbi, Associate Editor

Accepted June 5, 2003