The Plant Cell, Vol. 13, 627–643, March 2001, www.plantcell.org © 2001 American Society of Plant Physiologists

Self-Incompatibility in the Genus *Arabidopsis*: Characterization of the *S* Locus in the Outcrossing *A. lyrata* and Its Autogamous Relative *A. thaliana*

Makoto Kusaba,[a,1] Kathleen Dwyer,[b] Jennifer Hendershot,[b] Julia Vrebalov,[c] June B. Nasrallah,[a] and Mikhail E. Nasrallah[a,2]

[a] Department of Plant Biology, Cornell University, Ithaca, New York 14853
[b] Biology Department, University of Scranton, Scranton, Pennsylvania 18510
[c] Federal Soil and Nutrition Laboratory, United States Department of Agriculture, Ithaca, New York 14853

As a starting point for a phylogenetic study of self-incompatibility (SI) in crucifers and to elucidate the genetic basis of transitions between outcrossing and self-fertilizing mating systems in this family, we investigated the SI system of *Arabidopsis lyrata*. *A. lyrata* is an outcrossing close relative of the self-fertile *A. thaliana* and is thought to have diverged from *A. thaliana* ~5 million years ago and from *Brassica* spp 15 to 20 million years ago. Analysis of two *S* (sterility) locus haplotypes demonstrates that the *A. lyrata S* locus contains tightly linked orthologs of the *S* locus receptor kinase (*SRK*) gene and the *S* locus cysteine-rich protein (*SCR*) gene, which are the determinants of SI specificity in stigma and pollen, respectively, but lacks an *S* locus glycoprotein gene. As described previously in *Brassica*, the *S* haplotypes of *A. lyrata* differ by the rearranged order of their genes and by their variable physical sizes. Comparative mapping of the *A. lyrata* and *Brassica S* loci indicates that the *S* locus of crucifers is a dynamic locus that has undergone several duplication events since the *Arabidopsis–Brassica* split and was translocated as a unit between two distant chromosomal locations during diversification of the two taxa. Furthermore, comparative analysis of the *S* locus region of *A. lyrata* and its homeolog in self-fertile *A. thaliana* identified orthologs of the *SRK* and *SCR* genes and demonstrated that self-compatibility in this species is associated with inactivation of SI specificity genes.

INTRODUCTION

Self-incompatibility (SI) is the major outcrossing mechanism in the family Brassicaceae (de Nettancourt, 1977). Species in this family have been grouped into 19 tribes on the basis of morphological criteria (Schultz, 1936), and SI has been described in all tribes analyzed to date. When Bateman (1955) surveyed 182 species distributed in 11 tribes, he found that approximately half of these species included self-incompatible accessions. In a survey of 59 taxa in the subtribe Brassicineae of the tribe Brassiceae (which includes *Brassica* and *Raphanus*), 50 taxa were self-incompatible (Takahata and Hinata, 1980). In all cases analyzed, SI has been shown to be controlled sporophytically by a single *S* (sterility) locus, with multiple alleles or variants and complex dominance relationships between alleles (Bateman, 1954, 1955; Thompson and Taylor, 1966): in self-incompatible plants, pollen will not develop on a stigma that expresses the same *S* alleles as the pollen parent.

Molecular analysis of the *Brassica S* locus region has shown that this mendelian locus is a gene complex consisting of distinct stigma-expressed and anther-expressed genes that determine SI specificity in stigma and pollen, respectively (reviewed in Nasrallah, 2000). The *SRK* (for *S* locus receptor kinase) gene (Stein et al., 1991) encodes a plasma membrane–spanning receptor serine/threonine kinase specific to the stigma epidermis (Stein et al., 1996) and is the determinant of SI specificity in the stigma (Takasaki et al., 2000). The *SCR* (for *S* locus cysteine-rich protein) gene, which is the determinant of SI specificity in pollen (Schopfer et al., 1999), is expressed specifically in the anther tapetum and in microspores (Schopfer et al., 1999; Schopfer and Nasrallah, 2000; Takayama et al., 2000) and encodes a small secreted cysteine-rich protein predicted to localize to the pollen coat and to function as a ligand for the SRK receptor. As expected for genes involved in self-recognition, the *SRK* and *SCR* genes are highly polymorphic (Stein et al., 1991; Kusaba et al., 1997; Kusaba and Nishio, 1999; Schopfer et al., 1999; Schopfer and Nasrallah, 2000; Watanabe et al., 2000) and appear to have coevolved. A third highly polymorphic gene contained within

[1] Current address: Institute of Radiation Breeding, National Institute of Agrobiological Resources, Ministry of Agriculture, Forestry, and Fisheries, Ohmiya-machi, Naka-gun, Ibaraki 319-2293, Japan.
[2] To whom correspondence should be addressed. E-mail men4@cornell.edu; fax 607-255-5407.

the *S* locus in most *Brassica S* haplotypes is the *SLG* (for *S* locus glycoprotein) gene (Nasrallah et al., 1987; Kusaba et al., 1997). This gene exhibits a high degree of sequence similarity to the extracellular domain or the *S* domain of *SRK* (Stein et al., 1991; Kusaba et al., 1997) and encodes a glycoprotein specific to the papillar cell and localized in its wall (Kandasamy et al., 1989). The function of *SLG* is not well understood, and its role in SI has been questioned (Cabrillac et al., 1999; Nasrallah, 2000; Nishio and Kusaba, 2000). However, recent studies have indicated that *SLG* enhances the intensity of the SI response (Takasaki et al., 2000), possibly by contributing to the proper maturation of the SRK receptor and its accumulation to physiologically relevant levels (Dixit et al., 2000).

Sequence comparisons of *Brassica SRK* alleles have allowed a grouping of *S* haplotypes into two classes. Class I consists of *S* haplotypes that are placed high in the dominance series and that determine a robust SI response; class II consists of *S* haplotypes that are recessive to class I haplotypes in pollen and that determine a relatively leaky SI response (Chen and Nasrallah, 1990; Kusaba et al., 1997). Class I and class II *SRK* alleles diverge by >30%, whereas class I and class II *SCR* alleles are so diverged from one another (Schopfer et al., 1999) that the high number of available class I *SCR* sequences has not allowed the isolation of class II *SCR* alleles.

Comparative mapping of different *S* haplotypes has demonstrated that they vary not only in the sequence of their SI genes but also in the order, relative orientation, and spacing of these genes (Boyes and Nasrallah, 1993; Boyes et al., 1997; Cui et al., 1999; Nasrallah, 2000; Takayama et al., 2000). This structural heteromorphism and the sequence polymorphism of SI genes exhibit *trans*-specific variation (Dwyer et al., 1991; Kusaba et al., 1997; Schopfer et al., 1999; Uyenoyama, 2000). Thus, the *S* haplotypes of *Brassica* are thought to be ancient and to predate speciation in the genus. Estimates of divergence time have indicated that class I and class II *S* haplotypes diverged ~40 million years ago, whereas *S* haplotypes within class I and class II diverged ~25 and 7 million years ago, respectively (Uyenoyama, 1995, 2000).

In view of the ancient age of *Brassica S* haplotypes, an understanding of the origins and evolution of the *S* locus complex and its SI specificity genes in the crucifer family requires comparisons with *S* locus genes from tribes of the Brassicaceae other than the tribe Brassiceae. Because a detailed molecular analysis of SI has been performed only in *Brassica* species, it is not known which of the features elucidated for SI and the *S* locus in *Brassica* are specific to lineages within the Brassiceae and which features are common to the *S* loci of crucifers in general. Application of a phylogenetic approach to the study of SI is likely to provide insight into the structure and composition of the archetypal *S* locus as well as the origins of *S* locus polymorphisms.

As a starting point for this analysis, we elected to investigate SI in the genus *Arabidopsis*. Classically, the genus *Arabidopsis* has been placed in the tribe Sisymbrieae (Al-Shehbaz, 1984), but more recently it has been included in the tribe Arabideae (Price et al., 1994). In any case, *Arabidopsis* is thought to have diverged from *Brassica* ~15 to 20 million years ago (Yang et al., 1999). For our analysis, we used an outcrossing herbaceous perennial plant that was incorporated recently into the genus *Arabidopsis* on the basis of molecular data (O'Kane and Al-Shehbaz, 1997; van Treuren et al., 1997; Koch et al., 1999). This species, *Arabidopsis lyrata* (formerly known as *Arabis lyrata*, *Arabis petraea*, or *Cardaminopsis petraea*), has a genome size approximately twice that of *A. thaliana*. The genes of the two *Arabidopsis* spp share a high degree of sequence similarity, allowing facile transfer of molecular markers and other data generated by the *A. thaliana* genome project to *A. lyrata* (van Treuren et al., 1997). In addition, unlike *Brassica* spp, *A. lyrata* is not cultivated and is not closely related to cultivated species. This feature has made it an attractive model for population genetics and ecological studies (Koch et al., 1999) and will allow a study of SI in natural populations that have not been subjected to selective breeding.

Importantly, a molecular analysis of SI in *A. lyrata* and comparisons with its self-fertilizing close relative *A. thaliana* are likely to provide insight into the evolution of SI and the genetic basis of transitions in mating systems within the crucifer family. In a previous comparative mapping study of the *S* locus region of *Brassica* and its homeolog in *A. thaliana*, we found that the two regions, although largely colinear, differ by the absence of the SI specificity genes in *A. thaliana* (Conner et al., 1998). This result suggested two possibilities. One possibility is that the *A. thaliana* genome does not contain SI specificity genes, which implies that self-compatibility in this species resulted from deletion of these genes. The other possibility is that the SI specificity genes were translocated to different chromosomal locations in the *Brassica* and *Arabidopsis* lineages, which leaves open the question of why *A. thaliana* is self-fertile. Analysis of the *S* locus and SI specificity genes in *A. lyrata* is expected to distinguish between these possibilities.

In this article, we report on the analysis of SI in a population of *A. lyrata* plants segregating for two SI specificities. We describe the isolation of *SRK* and *SCR* genes from the two *S* haplotypes encoding these specificities. The sequence polymorphism exhibited by these genes as well as the complement of genes contained in the *A. lyrata S* locus and their genomic organization are discussed in relation to our current knowledge of SI in *Brassica*. Furthermore, on the basis of the chromosomal location of the *S* locus in *Arabidopsis* spp and the analysis of *A. thaliana* sequences orthologous to the *SRK* and *SCR* genes, the molecular basis for self-compatibility in this species is inferred.

RESULTS

Cytology, Development, and Genetics of SI in *A. lyrata*

A. lyrata has been described as a largely outcrossing species on the basis of the inability of isolated plants to self-

pollinate and set seed (Schierup, 1998). However, cytological and developmental studies of pollen rejection have not been reported for this species. We examined pollen tube behavior on *A. lyrata* stigmas after self- and cross-pollination and at different stages of stigma development. This microscopic analysis showed that self-pollen is inhibited at the surface of the stigma epidermis in mature flowers and that the stigmas of immature buds support the development of pollen tubes in self-pollinations (data not shown). Thus, the SI response in *A. lyrata* is similar to that of other self-incompatible crucifers, both in its site of inhibition and in its developmental regulation. Furthermore, as in *Brassica*, *S* locus homozygotes can be generated and maintained by forced self-pollination in immature young buds before the stigma acquires competence for pollen rejection.

To investigate the genetic control of SI in *A. lyrata*, we started with seed collected from the wild in Michigan and kindly provided by Charles Langley (University of California at Davis), from which we generated a population that segregated for SI phenotype by isolating a single plant and forcing self-pollination on developing floral buds. A population of 54 plants was analyzed by microscopic examination of pollen tube growth in self-pollinations and in reciprocal cross-pollinations among siblings and with the parent plant (see Methods). The data are consistent with the segregation of two SI specificities, *Sa* and *Sb*, and with the single locus control of SI. As shown in Table 1, the plants were grouped into three classes based on pollination phenotype: 18 plants were *SaSa*, 27 plants were *SaSb*, and nine plants were *SbSb*. The observed ratios of the three classes approximate the 1:2:1 ratio expected for segregation of a single locus ($\chi^2 = 3$, $P = 0.25$). In pollen of heterozygotes, *Sb* is dominant to *Sa*. In stigmas of heterozygotes, *Sa* and *Sb* exhibit codominance, with weakening of *Sa* activity. The *Sa* haplotype also conferred a somewhat leaky SI response in *SaSa* homozygotes, as indicated by increased seed set resulting from spontaneous self-pollination (pseudocompatibility) relative to *SbSb* homozygotes and *SaSb* heterozygotes.

Table 1. Pollination Analysis of Plants Segregating for *Sa* and *Sb* Specificities[a]

	Pollen Parent[b]		
Ovule Parent	*SaSa* (18)	*SaSb* (27)	*SbSb* (9)
SaSa (18)	−	+	+
SaSb (27)	±	−	−
SbSb (9)	+	−	−

[a] Numbers within parentheses indicate the number of plants in each phenotypic group.
[b] (−) indicates an incompatible response in which <10 pollen tubes were produced per pollinated stigma; (+) indicates a compatible response in which >300 pollen tubes were produced per pollinated stigma; (±) indicates a variable pollen tube count in which 20 to 100 pollen tubes were produced per pollinated stigma.

Molecular Cloning of *A. lyrata SRKa* and *SRKb* Genes

A cDNA library was constructed in the bacteriophage vector λUniZAP from pistil mRNA isolated from an *SaSb* plant. The library was screened at low stringency with *S* domain probes from *Brassica SLG* and related genes (see Methods). Positive clones that also contained kinase sequences were identified subsequently by hybridization with a probe derived from the kinase domain of *Brassica SRK*. Several clones that hybridized with both probes were sequenced partially and found to be derived from the same gene. Sequence analysis of the clone with the longest insert revealed one open reading frame (ORF) with the features characteristic of a functional *SRK* gene (see below). To determine if this gene was in fact derived from the *A. lyrata S* locus, we performed restriction fragment length polymorphism (RFLP) analysis of plants segregating for the *Sa* and *Sb* haplotypes. As shown in Figure 1A, under high-stringency wash conditions (see Methods), the *S* domain of the putative *SRKa* gene hybridized with a 1.7-kb restriction fragment in EcoRV digests of DNA from plants carrying the *Sa* haplotype. This probe did not hybridize with DNA from *Sb* homozygotes, indicating a high degree of divergence between the two *SRK* alleles. Furthermore, in the 54 plants analyzed, there was perfect correlation between the presence of this 1.7-kb EcoRV fragment and SI phenotype deduced from pollination assays. These results demonstrate that we had isolated the *SRKa* allele of *A. lyrata*. Additional confirmation of this conclusion derives from work done independently by Deborah Charlesworth and co-workers, who amplified an *A. lyrata* genomic sequence nearly identical to the *S* domain of the *SRKa* gene and showed its linkage to the *S* locus by using an independent segregating population (Schierup et al., 2001).

Genomic clones for the *SRKa* gene were isolated from libraries of *A. lyrata SaSb* DNA constructed in the bacteriophage vector λGEM11 and in the bacterial artificial chromosome (BAC) vector pBeloBAC (see Methods). A screen of the λGEM11 library with the *S* domain of *SRKa* resulted in the isolation of several λ clones for *SRKa*. Analysis of sequences flanking the *SRKa* gene in these clones provided single-copy probes that subsequently were used to probe the BAC library, leading to the isolation of BAC clones derived from the *Sa* and *Sb* haplotypes (see Methods).

In our RFLP analysis of plants segregating for the *Sa* and *Sb* specificities by using the *SRKa S* domain probe, we detected an *Sb*-associated 550-bp XhoI fragment under low-stringency wash conditions (data not shown). A BAC clone derived from the *Sb* haplotype was digested with XhoI, and a subclone containing the 550-bp XhoI fragment was obtained. This 550-bp XhoI fragment was used to screen the pistil cDNA library, resulting in the isolation of six independent cDNA clones. All of these clones hybridized with a kinase domain probe, suggesting that they contained *SRK*-like sequences. Further analysis indicated that the six clones were all derived from the same gene. RFLP analysis

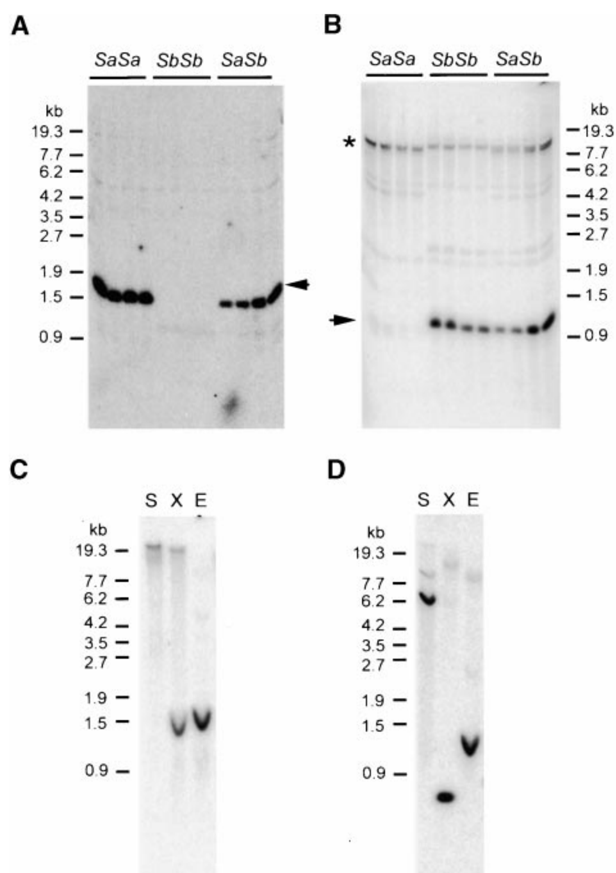


Figure 1. DNA Gel Blot Analysis of the *A. lyrata SRK* Gene.

(A) and **(B)** RFLP analysis of plants segregating for the *Sa* and *Sb* haplotypes. Genomic DNA was digested with EcoRV and probed with the *S* domains of *SRKa* **(A)** and *SRKb* **(B)**. The arrows at ∼1.5 kb in **(A)** and at ∼0.9 kb in **(B)** indicate restriction fragments containing the *SRKa* and *SRKb* genes, respectively. The asterisk at ∼10 kb in **(B)** indicates additional restriction fragments that hybridize with the *S* domain of *SRKb* and appear to be linked to, but reside outside of, the *S* locus. Molecular mass markers in kilobases are shown next to each blot.

(C) and **(D)** DNA gel blot analysis of *A. lyrata Sa* **(C)** and *Sb* **(D)** homozygotes. The DNA was digested with SacI (S), XhoI (X), and EcoRI (E) and probed with the *S* domain of *SRKa* **(C)** and *SRKb* **(D)**.

of our segregating population with the 550-bp XhoI probe confirmed the linkage of this gene to the *S* locus. Under high-stringency wash conditions, the 550-bp XhoI probe detected a 1.3-kb restriction fragment that cosegregated with *Sb* specificity, as shown by the EcoRV digests in Figure 1B.

DNA isolated from *SaSa* and *SbSb* homozygotes was digested with several restriction enzymes and subjected to gel blot analysis by using *S* domain probes from the *SRKa* and *SRKb* genes, respectively. Figure 1C shows that under high-stringency conditions, the *SRKa* probe hybridized with a single restriction fragment in SacI and EcoRV digests. Two restriction fragments were detected in XhoI digests due to the presence of an internal XhoI site within the probe. These results indicate that the *A. lyrata* genome does not contain other genes with a high degree of sequence similarity to *SRKa*. In particular, a homolog of the *SLG* gene, which is detectable in *Brassica* DNA by using *SRK S* domain probes even under high-stringency conditions, appears to be absent from the *Sa* haplotype of *A. lyrata.*

In a similar gel blot analysis of DNA isolated from *SbSb* homozygotes, the *SRKb*-derived 550-bp XhoI probe detected an additional restriction fragment under high-stringency wash conditions (Figure 1D), indicating the presence of a sequence with appreciable sequence similarity to the *SRKb* gene. A restriction fragment of a slightly different size, possibly containing an allelic form of this sequence, also was detected by the 550-bp XhoI probe in *SaSa* plants (Figure 1B). The RFLP for this sequence cosegregated with the *S* locus in the 12 plants analyzed, suggesting its possible linkage to the *S* locus. However, the gene was not located in the *S* locus BAC contig; therefore, it must lie outside of the *S* locus proper.

Sequence Polymorphism of the *A. lyrata SRK* Gene

Database searches showed that the *A. lyrata SRK* genes share appreciable sequence similarity with *Brassica SRK* alleles as well as with several other members of the *S* gene family. The sequence alignment in Figure 2A shows that the *SRKa* and *SRKb* genes exhibit the features characteristic of receptor-like kinases of the *S* gene family: a signal sequence, a predicted extracellular *S* domain containing 12 conserved cysteine residues, a transmembrane domain, and a kinase domain with conserved residues suggestive of serine/threonine kinase activity. Alignment of the cDNA and genomic sequences revealed that *SRKa* and *SRKb* share with *Brassica SRK* genes (Stein et al., 1991) and *A. thaliana S* receptor–like kinase genes (Tobias et al., 1992; Dwyer et al., 1994; Tobias and Nasrallah, 1996) a conserved exon–intron structure in which the entire *S* domain is contained on the first exon, the transmembrane and juxtamembrane domains are encoded by the second exon, and the kinase domain is contained on exons 3 to 7. The six introns of *SRKa* and *SRKb* occur at the same positions as in other *S* receptor kinase genes, but they differ in size between the two genes (arrowheads and numbers in Figure 2A).

The *SRKa* and *SRKb* ORFs are highly diverged and exhibit only 62.6% amino acid sequence similarity. In *Brassica*, sequence similarities in the range of 60 to 65% are observed in comparisons between class I and class II *SRK* alleles, whereas sequence similarities of >80% are typically observed in comparisons of *SRK* alleles within a class. It is possible that the *Sa* and *Sb* haplotypes belong to two distinct classes of *A. lyrata SRK* alleles. However, the *SRKa* and *SRKb* sequences are equally diverged from the class I and class II alleles of *Brassica*: each allele shares 50 to 60%

sequence similarity with the *Brassica* class I *SRK6* and class II *SRK2* alleles, and these *A. lyrata* alleles did not group with either *Brassica* allele in a preliminary phylogenetic analysis (Figure 2B).

Expression Pattern of the *A. lyrata SRK* Gene

Reverse transcription–polymerase chain reaction (RT-PCR) analysis using primers specific for the *SRKa* gene showed that this gene is expressed at its highest level in stigmas (Figure 3A). *SRKa* expression is maximal at 1 to 2 days before flower opening and persists in stigmas of open flowers. The gene also is expressed at low levels in styles but not in anthers at 1 to 2 days before flower opening. This expression pattern is largely identical to that of *Brassica SRK* genes. However, one difference between the two taxa may be noted: in *A. lyrata*, *SRK* expression was detected only in pistils, whereas in *Brassica*, *SRK* also exhibited a low level but physiologically irrelevant expression in anthers (Stein et al., 1991, 1996; Conner et al., 1997; Cui et al., 2000; Takasaki et al., 2000).

A. lyrata RNA isolated from *SaSa* and *SbSb* pistils was examined by gel blot analysis using *S* domain probes derived from the *SRKa* and *SRKb* genes, respectively. This analysis revealed the presence of two transcript species: a 2.8-kb transcript that corresponds to the size of the full-length *SRK* transcript, and a shorter 1.6-kb transcript (Figure 3B) that is ~5 and 10% as abundant as the 2.8-kb transcript in *SaSa* and *SbSb* pistils, respectively. *Brassica* stigmas also produce 2.8- and 1.6-kb transcripts that are detected with *SRK S* domain probes (Stein et al., 1991). However, in most *Brassica* strains analyzed to date, the steady state amount of the 1.6-kb transcripts exceeds that of the 2.8-kb transcripts by as much as two orders of magnitude. The *Brassica* 1.6-kb transcripts have been shown to consist predominantly of transcripts derived from the *SLG* gene and to a lesser extent of *SRK* transcripts that contain the *S* domain–encoding exon and terminate within the first intron of the gene (Stein et al., 1991; Giranton et al., 1995). In *A. lyrata*, the entire 1.6-kb transcript population must be derived from the *SRK* gene, based on of the absence of an *SLG*-like gene in the *Sa* and *Sb* haplotypes we analyzed.

Molecular Cloning and Characterization of the *A. lyrata SCR* Genes

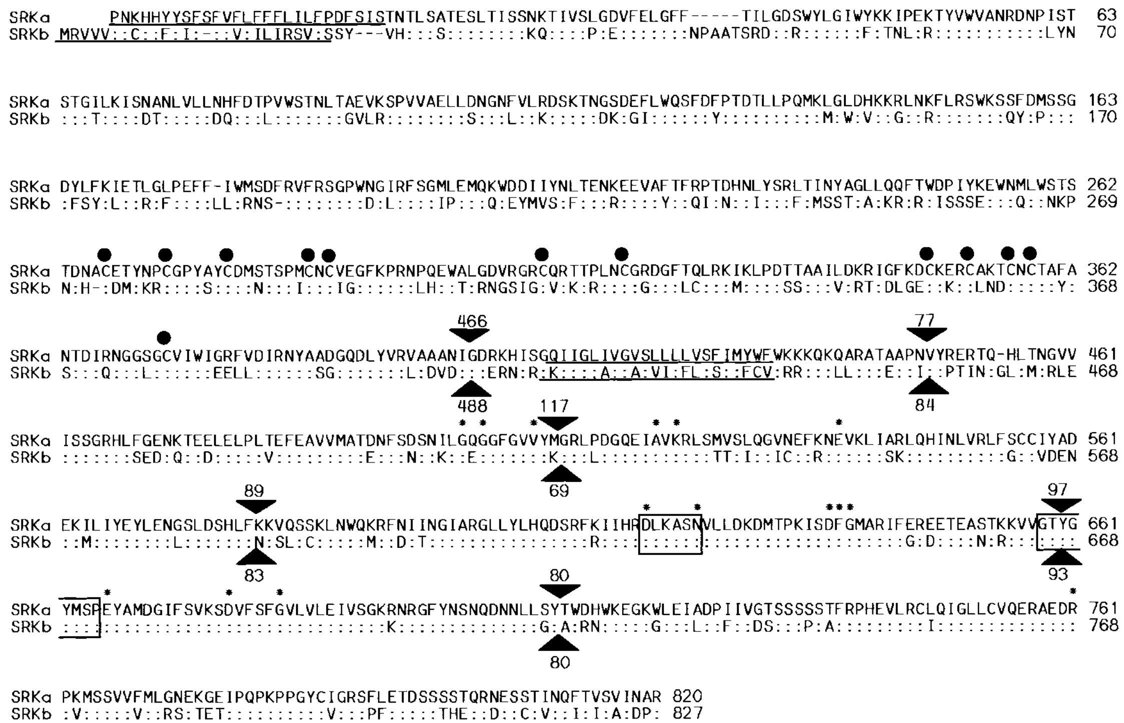
We postulated that the *SCR* gene would be located in close physical proximity to the *SRK* gene in *A. lyrata*, as it is in *Brassica*. Furthermore, the *SCR* alleles of *Brassica* exhibit extensive sequence divergence even when *S* haplotypes are compared, the *SRK* genes of which share >90% sequence similarity. Because of the >30% sequence divergence exhibited by the *SRKa* and *SRKb* genes, we also postulated that a probe containing the *SCRa* gene would not hybridize with transcripts encoded by the *SCRb* gene. Therefore, to identify the *SCR* genes in the *Sa* and *Sb* haplotypes, we used subclones of *SRKa*-containing λ genomic clones and of an *SRKb*-containing BAC clone to probe gel blots of poly(A)$^+$ RNA isolated from *SaSa* and *SbSb* anthers. An 8.8-kb XhoI fragment (Figure 4A) from an *SRKa*-containing λ clone hybridized specifically with an ~500-bp transcript in *SaSa* anthers (data not shown), suggesting that this fragment contained the *SCRa* allele. Similarly, a 14.5-kb EagI fragment from an *SRKb*-containing BAC clone (Figure 4A) hybridized specifically with an ~500-bp transcript in *SbSb* anthers (data not shown), indicating that it contained the *SCRb* allele.

The 8.8-kb XhoI fragment and the 14.5-kb EagI fragment were used to screen anther cDNA libraries prepared from *SaSa* and *SbSb* plants, respectively. Several positive clones were isolated from each library. Comparison of cDNA and genomic sequences revealed that the *A. lyrata SCR* genes have a structure similar to that of the *Brassica SCR* genes. Each gene is interrupted by one intron that is located toward the 3′ end of the signal sequence (Figure 4B) and measures ~1.5 kb in *SCRa* and 78 bp in *SCRb*. The cDNA clones isolated from the *SaSa* library all contained an identical 258-bp ORF with a structure characteristic of the *SCR* gene (Figure 4B). This ORF predicts a polypeptide of 86 amino acid residues containing a potential signal sequence that, when cleaved, would result in a mature protein of 63 amino acids having a molecular mass of 7.3 kD and a pI of 8.3. The cDNA clones isolated from the *SbSb* library also all contained an identical 243-bp ORF that encodes a protein of 81 amino acids containing a potential signal sequence (Figure 4B). The predicted mature SCRb protein is 55 amino acids long and has a molecular mass of 6.5 kD and a pI of 8.3.

The *SCRa* and *SCRb* sequences are highly diverged from each other and from *Brassica SCR* genes. *SCRa* and *SCRb* share only 35.2% amino acid sequence identity, a level of intraspecific allelic polymorphism also exhibited by *Brassica SCR* genes. As shown in Figure 4B, the predicted mature SCRa and SCRb proteins contain the eight cysteines and one glycine residue that are conserved in most *Brassica SCR* alleles analyzed to date (Schopfer et al., 1999; Schopfer and Nasrallah, 2000; Takayama et al., 2000; Watanabe et al., 2000). Also like their *Brassica* orthologs, SCRa and SCRb differ in the sequence and length of the segments between and flanking these conserved residues. Interestingly, the *SCRa* and *SCRb* alleles also have highly divergent signal sequences and differ in the positions of their predicted signal sequence cleavage sites (Figure 4B). In this respect, they differ from the *Brassica SCR* alleles isolated to date, all of which contain relatively conserved signal sequences.

The *SCRa* and *SCRb* genes also differ in their positions within the *S* locus relative to the *SRK* gene. Fine structure mapping of an *Sa*-derived BAC clone indicated that *SRKa* and *SCRa* are located in inverse orientation and within ~1.5 kb of each other (Figure 4A). Unexpectedly, restriction analysis

A



B

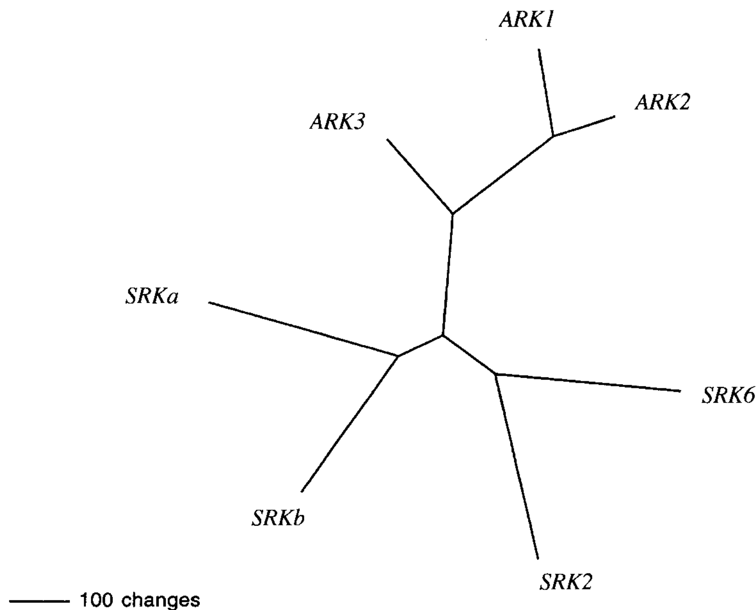


Figure 2. Polymorphism of the *A. lyrata SRK* Gene and Relationship to the *Brassica SRK* Genes and *A. thaliana ARK* Genes.

(A) Alignment of the predicted SRKa and SRKb proteins. The SRKa sequence is missing five residues at its N terminus. The putative signal se-

of *Sb* BAC clones showed that the *Sb* haplotype contained two copies of *SCRb*, designated *SCRb-1* and *SCRb-2* (Figure 4A). These two genes are located within a longer direct repeat that also includes a retroelement-like sequence downstream from *SCRb-2* and a truncated duplicate of the element downstream from *SCRb-1* (Figure 4A). The *SCRb-1* and *SCRb-2* genes are arranged in the same orientation as *SRKb*, with *SCRb-1* positioned ~15.5 kb from *SCRb-2* and ~21.0 kb from *SRKb*. The exons of the *SCRb-1* and *SCRb-2* genes are identical to the *SCRb* cDNA sequence, and the two genes also are identical to each other over the length of sequence analyzed, namely, from 580 bp upstream of the initiation codon to 280 bp downstream from the stop codon. Therefore, it is likely that *SCRb-1* and *SCRb-2* are both functional genes and exhibit identical expression patterns.

RNA gel blot analysis demonstrated that the *SCRa* and *SCRb* genes are expressed specifically in anthers. As shown in Figure 4C for *SCRb*, transcripts were detected in both young and mature anthers as well as in isolated microspores. This expression pattern is similar to that described previously for *Brassica SCR* genes by RNA gel blot analysis (Schopfer et al., 1999), in situ hybridization (Takayama et al., 2000), and reporter gene analysis (Schopfer and Nasrallah, 2000).

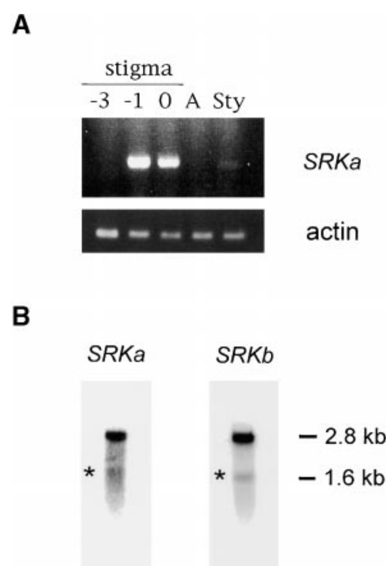


Figure 3. Expression of the *A. lyrata SRK* Gene.

(A) Amplification of *SRK* mRNA from stigmas at different stages of development, from anthers (A), and from styles (Sty). All tissues were collected from *SaSa* plants. Poly(A)$^+$ RNA isolated from each tissue was subjected to RT-PCR by using *SRKa*-specific primers (top) or actin-specific primers (bottom) as described in Methods. The numbers above the stigma lanes refer to the ages of flower buds from which the stigmas were collected: the −3 stage is 3 days before flower opening, the −1 stage is 1 day before flower opening, and the open (O) stage is the day of flower opening.

(B) Gel blot analysis of poly(A)$^+$ RNA isolated from −1 stage stigmas of *SaSa* (left) and *SbSb* (right) plants. The blots were probed with the *S* domain of *SRKa* and *SRKb* as indicated. Each lane contains 0.7 μg of poly(A)$^+$ RNA. The asterisks indicate the short transcripts discussed in the text.

Chromosomal Location of the *S* Locus in *Arabidopsis* spp

Before our construction and screening of a pistil cDNA from *A. lyrata* as described above, we had attempted to isolate *S* locus genes from *A. lyrata* based on the assumptions that the *S* locus occupied a similar chromosomal location in *Brassica* and *A. lyrata* and that the regions flanking the SI genes would be largely colinear in the two taxa. We previously had mapped the *Brassica S* locus to an *ETR1*-linked position and performed a comparative analysis of the *Brassica S* locus region and its *ETR1*-linked homeolog in *A. thaliana* (Conner et al., 1998). This analysis had identified several molecular markers that flank the SI specificity genes in *Brassica* and are conserved in the homeologous region of the *A. thaliana* genome (Conner et al., 1998). From these markers, we selected markers that were single copy in *Brassica* and *A. thaliana* and that were located on either side of the *Brassica SRK–SCR–SLG* complex (see Methods). These markers were used to screen a library of *A. lyrata* genomic DNA constructed in the bacteriophage vector λGEM11, and 50 positive clones were isolated and assembled into contigs of overlapping clones (data not shown). Restriction mapping of the clones confirmed the synteny of the *A. thaliana* and *A. lyrata* homeologous regions. Unexpectedly, however, none

Figure 2. (continued).

quences and transmembrane domains are underlined. The black dots indicate the 12 cysteines that are conserved in the *S* domains of genes in the *S* gene family. The boxes and asterisks indicate residues in the kinase domain that are conserved in protein kinases. The arrowheads indicate the positions of the introns deduced from aligning the cDNA and genomic sequences. The numbers above and below the arrowheads indicate the sizes of the introns in *SRKa* and *SRKb*, respectively. The *SRKa* and *SRKb* cDNA sequences can be found in GenBank under accession numbers AB052755 and AB052756, respectively.

(B) The single most parsimonious tree identified by phylogenetic analysis of DNA sequences of *A. lyrata SRKa* and *SRKb*, *Brassica SRK2* and *SRK6*, and three members of the *A. thaliana S* gene family (*ARK1*, *ARK2*, and *ARK3*). The tree is shown as unrooted because of difficulties in assuming an outgroup and the large amount of divergence of potential outgroups from these sequences.

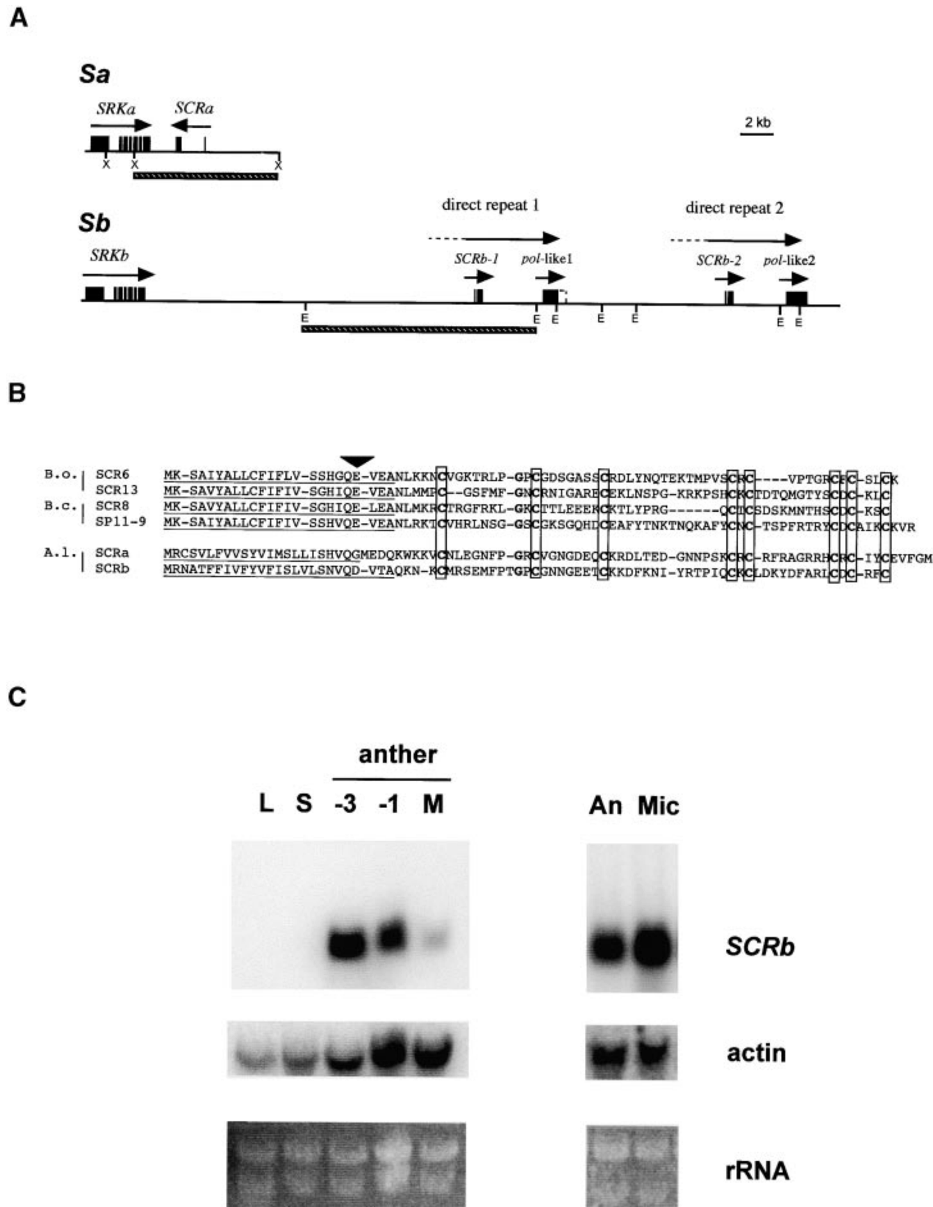


Figure 4. Location, Polymorphism, and Expression of the *A. lyrata SCR* Gene.

(A) Partial maps of the *Sa* and *Sb* haplotypes. The arrows above each gene indicate the direction of transcription. The hatched boxes below the maps indicate the restriction fragments that were used to probe gel blots of anther RNA and to identify the location of the *SCR* gene on genomic λ clones (*Sa*) or BAC clones (*Sb*). The map of the *Sb* haplotype shows the positions of the two *SCRb* genes as well as the position of the larger direct repeat that includes the *SCRb* gene and a *pol*-like sequence. The partially closed rectangle representing the *pol*-like1 sequence indicates that this sequence is truncated relative to the *pol*-like2 sequence. The broken segments of the arrows spanning the direct repeats indicate that the extent of the repeats has not been determined. E, EagI; X, XhoI.

(B) Alignment of the predicted *A. lyrata* (A.l.) *SCRa* and *SCRb* proteins and *SCR* alleles from *Brassica oleracea* (B.o.) and *B. campestris* (B.c.). The following *SCR* alleles (with GenBank accessions numbers) are shown: *SCR6* (AF195625), *SCR13* (AF195626), *SCR8* (AF195627), and *SP11-9* (AB022078). *SCR* is designated *SP11* by Takayama et al. (2000). The *SCRa* and *SCRb* sequences can be found in GenBank under accession numbers AB052753 and AB052754, respectively.

(C) Expression pattern of the *SCRb* gene. Total RNA was isolated from leaves (L), stigmas and styles (S), anthers (An), and microspores (Mic) of *SbSb* plants. The anther samples shown at left were collected from floral buds at the −3 and −1 stages of development (see legend to Figure 3) and from mature anthers (M). The RNA samples shown at right were prepared from whole anthers (An) and isolated microspores (Mic) at the −3 stage of development. Each lane contains ~15 μg of total RNA. The blots were probed with the *SCRb* cDNA and subsequently with an actin probe as a loading control. A photograph of the ethidium bromide gel before RNA transfer is shown.

of the clones hybridized with *Brassica SRK*-derived probes, suggesting that they did not contain *SRK*-related sequences. We subsequently found that these clones did not hybridize with *S* domain probes from *A. lyrata SRK* and that λGEM11 clones containing the *SRKa* gene did not hybridize with the molecular markers derived from the *S* locus region of *Brassica*. These unexpected results indicated that the *S* locus occupies a different region of the *A. lyrata* genome than it does in *Brassica*.

Database searches using the *SRKa* sequence showed that it is most similar to an *A. thaliana SFR2*-like receptor kinase (GenBank accession number CAA20202) that is located on BAC T6K22 in contig fragment 55 of chromosome IV. In addition, database searches with partial DNA sequences of subclones derived from regions flanking the *A. lyrata SRKa* gene produced hits on chromosome IV of *A. thaliana*. Specifically, a subset of these sequences exhibited 94 to 96% sequence identity with the *S* domain of the *ARK3* gene and 90% sequence identity with a *BRLK*-like gene, both of which are contained on BAC T6K22. This *ARK3*-containing region thus is identified as the *A. thaliana* homeolog of the *A. lyrata S* locus region.

Long-Range Comparative Mapping of Two Haplotypes of the *A. lyrata S* Locus and Its Homeolog in *A. thaliana*

A map of BAC T6K22 is shown in Figure 5. The region spanned by the BAC clone contains several predicted genes, including a cluster of five genes that exhibit sequence relatedness to kinase genes. Of these predicted kinase genes, three belong to the *S* receptor kinase class of genes and therefore are candidates for being the *A. thaliana* counterpart of *SRK*. These are T6K22.120 (GenBank accession number CAA20204; designated B120 in Figure 5), a sequence related to the *B. oleracea BRLK* gene; T6K22.110, which is identical to *ARK3* (GenBank accession number CAA20203), a gene characterized previously (Dwyer et al., 1994); and T6K22.100 (GenBank accession number CAA20202), a sequence related to a wound-inducible *SFR2* receptor-like kinase from *B. oleracea* (Pastuglia et al., 1997a).

To compare the organization and gene content of this *A. thaliana* region to the *A. lyrata S* locus region and identify possible *A. thaliana* orthologs of SI specificity genes, we performed a comparative mapping study of *A. thaliana* BAC T6K22 and *A. lyrata* BAC clones corresponding to the *Sa* and *Sb* haplotypes. The availability of the sequence of BAC T6K22 provided us with a collection of molecular markers that allowed us to derive long-range maps for the *A. lyrata Sa* and *Sb* haplotypes and to compare the *A. lyrata S* locus with its homeolog in *A. thaliana*. Restriction mapping and partial sequencing of the *A. lyrata Sa* and *Sb* BAC clones revealed a high degree of microsynteny between the *A. lyrata* and *A. thaliana* homeologous regions within segments that flank the *SRK*-*SCR*–containing segment (namely, in the segment defined by the B50 and B80 genes and in the segment defined by the *ARK3* and B160 genes). However, the segment between *ARK3* and B80, which contains the *SRK* and *SCR* genes (i.e., the *S* locus proper), differs markedly in the three regions analyzed, both in its size and organization. This segment is 32 kb in length in *A. thaliana*, ∼50 kb in the *A. lyrata Sa* haplotype, and ∼100 kb in the *A. lyrata Sb* haplotype. This segment also differs in the relative orientation and spacing of the *SRK* and *SCR* sequences, not only between the *A. lyrata Sa* and *Sb* haplotypes, as discussed above, but also between *A. thaliana* and the two *A. lyrata* regions. Interestingly, all three regions contain sequences with similarity to *pol* genes of retrotransposons: the *A. thaliana* and the *A. lyrata Sa* haplotype each contains at least one such sequence, whereas the *A. lyrata Sb* haplotype contains at least five such sequences.

Identification of *SRK* and *SCR* Orthologs in *A. thaliana*

The *SFR2*-like kinase T6K22.100 most likely represents the *A. thaliana* ortholog of *SRK*. This conclusion is based on several lines of evidence. First, the comparative mapping study of the largely syntenous *A. lyrata S* locus region and its homeolog in *A. thaliana* described above shows that the T6K22.100 sequence and the *A. lyrata SRK* genes are located in the same positions relative to flanking genes (Figure 5). Second, orthologs of *BRLK*-like and *ARK3* genes have been identified in *A. lyrata Sa* and *Sb* haplotypes and are distinct from the *SRK* gene. Third, *ARK3*, which is the *S* kinase gene that is most tightly linked to *SRK* in *A. lyrata*, is unlikely to function in SI because its promoter was found to be expressed at the root–hypocotyl junction and at the bases of lateral roots, axillary buds, and pedicels (Dwyer et al., 1994).

To identify sequences orthologous to *SCR*, we searched BAC T6K22 for the presence of sequences related to *A. lyrata SCR* genes. The search identified three *SCR*-related sequences (Figure 6A) downstream from the *SRK* ortholog T6K22.100, within a 15.7-kb region lacking predicted exons/ORFs (Figure 5). None of the three candidate *SCR* orthologs is predicted to encode full-length SCR proteins, and therefore they are designated Ψ*SCR1*, Ψ*SCR2*, and Ψ*SCR3* (Figure 6A). The Ψ*SCR1* sequence, located ∼700 bp downstream from T6K22.100 between positions 50,496 and 51,472 of the BAC T6K22 sequence, exhibits the exon–intron structure characteristic of *SCR* genes but contains a truncated ORF (Figure 6A). The Ψ*SCR2* and Ψ*SCR3* sequences are located ∼8.5 kb downstream from T6K22.100 at positions 43,623 to 43,664 and 43,724 to 43,687 of BAC T6K22, respectively, and share sequence similarity with the *SCRa* signal sequence (Figure 6A) as well as with the *SCRa* 5′ untranslated region (data not shown).

Analysis of the *A. thaliana SRK* Ortholog

The annotation of the *SFR2*-like kinase T6K22.100 predicts a functional gene that encodes a full-length receptor kinase

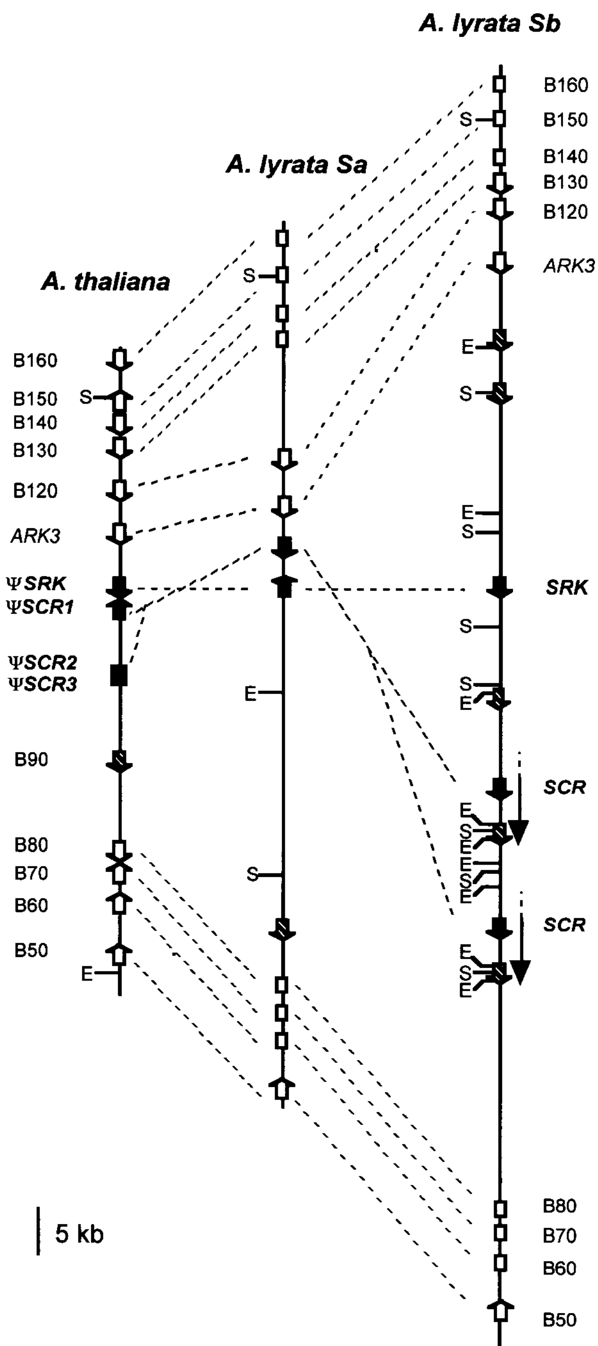


Figure 5. Comparative Map of the *A. lyrata Sa* and *Sb* Haplotypes and the Homeologous Region of *A. thaliana*.

The map of the *A. thaliana* region was drawn based on the sequence of BAC T6K22 (GenBank accession number AL031187). Numbers preceded by B represent the genes predicted from the *A. thaliana* T6K22 sequence (e.g., B160 is the predicted gene T6K22.160). The filled thick arrows indicate the positions and 5′→3′ orientation of the *SRK* and *SCR* genes of *A. lyrata* and the Ψ*SRK* and Ψ*SCR* sequences of *A. thaliana*. One filled box is used to show the location of the two short Ψ*SCR2* and Ψ*SCR3* sequences. The open thick arrows indicate the positions and 5′→3′ orientation of other sequences in the region. Open boxes on the *Sa* and *Sb* maps indicate sequences for which the direction of transcription has not been determined. The B130 and B140 sequences are very similar, and the relative order of the two sequences in *A. lyrata* is tentative. The hatched thick arrows indicate sequences with similarity to retrotransposon *pol* genes. The sizes of the arrows and boxes are not proportional to the sizes of the genes. The thin arrows to the right of the *Sb* map indicate the position and orientation of the direct repeat that includes the *SCRb* gene. Sequences shared by the three maps are connected by dashed lines. E, EagI; S, SmaI.

containing all of the features of receptor protein kinases in the *S* gene family, including 12 conserved cysteines in the extracellular *S* domain and the residues conserved in kinases. This annotation assumes the gene structure shown in Figure 6B, which is typical of *Brassica* and *A. lyrata SRK* as well as other receptor kinases in the *S* gene family.

To determine if this *SRK* ortholog produces transcripts that have the potential to encode a functional receptor kinase, we screened a cDNA library prepared from *A. thaliana* floral bud RNA with a fragment containing the *S* domain of the gene. We isolated three independent cDNA clones derived from the T6K22.100 gene. Sequence analysis of these cDNAs revealed that the gene is unlikely to produce a receptor kinase product. On the basis of sequence analysis of the three cDNA clones, the T6K22.100 gene is predicted to encode a 547–amino acid ORF rather than the 844–amino acid ORF predicted in the database (Figure 6B). This truncated ORF terminates within exon 4 due to the use, at the intron 3–exon 4 boundary, of a splice site different from that predicted in the database. As shown in Figure 6C, the use of this splice site introduces a stop codon six codons into exon 4. The T6K22.100 gene therefore is a nonfunctional ortholog of *SRK* and is designated Ψ*SRK*.

The isolation of cDNA clones for the Ψ*SRK* gene demonstrated that the gene was transcribed, and it was of interest to investigate its spatial and temporal expression pattern. To characterize the activity of the Ψ*SRK* promoter, we amplified a fragment spanning a region from 23 to ~1600 bp upstream of the initiating methionine codon and fused it to the *uidA* gene, which encodes the β-glucuronidase (GUS) reporter enzyme (see Methods). Transformation of *A. thaliana* with the resulting Ψ*SRK*::*uidA*::*nos* chimeric gene produced 25 independent transformants. In all transformants, GUS activity was restricted to the pistil. Twenty-three transformed families exhibited intense GUS staining in the stigma epidermis as well as in the transmitting tract of the style and ovary, as shown in Figure 7. In the remaining two families, GUS staining was confined to the stigma epidermis (data not shown). These results indicate that the Ψ*SRK* promoter exhibits a spatial activity pattern in pistils identical to that of

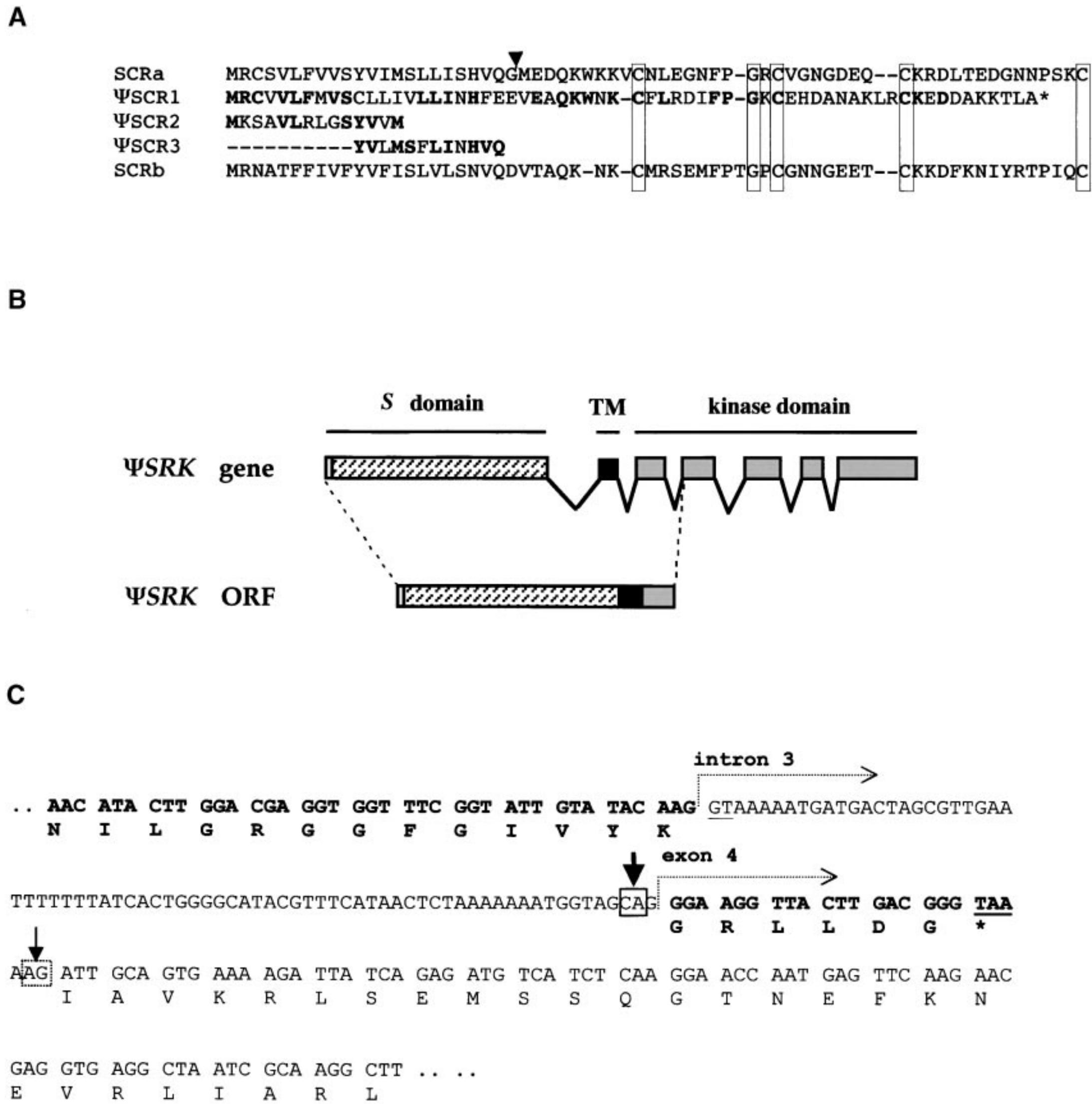


Figure 6. The Ψ*SRK* and Ψ*SCR* Sequences of *A. thaliana.*

(A) Alignment of the ΨSCR1, ΨSCR2, and ΨSCR3 deduced amino acid sequences with relevant regions of *A. lyrata* SCRa and SCRb. The arrowhead shows the position of the intron. The cysteine residues and glycine residue that are conserved in SCR peptides are marked by boxes. The residues in the ΨSCR1, ΨSCR2, and ΨSCR3 sequences that are shared with *A. lyrata* SCRa are shown in boldface. The first and second exons of ΨSCR1 are located, respectively, at positions 50,496 to 50,562 and 51,344 to 51,472 of the BAC T6K22 sequence. The Ψ*SCR2* and Ψ*SCR3* sequences are located at positions 43,623 to 43,664 and 43,724 to 43,687 of BAC T6K22, respectively.

(B) The truncated ORF encoded by the Ψ*SRK* gene. The exon–intron structure of the Ψ*SRK* gene as annotated in the sequence of BAC T6K22 is shown above the truncated ORF predicted from the sequence of cDNA clones. TM, transmembrane domain.

(C) Sequence of the Ψ*SRK* gene at the exon 3–intron 3–exon 4 junctions. The cDNA sequence and predicted ORF are shown in boldface. The 5′ splice site determined from the cDNA sequence is indicated by a box and thick arrow. The alternative 5′ splice site used in the annotation of the BAC T6K22 sequence is indicated by a dashed box and thin arrow. The stop codon that terminates the truncated ORF is underlined. Double dots indicate that the sequence continues beyond the point shown. (*) indicates the stop codon.

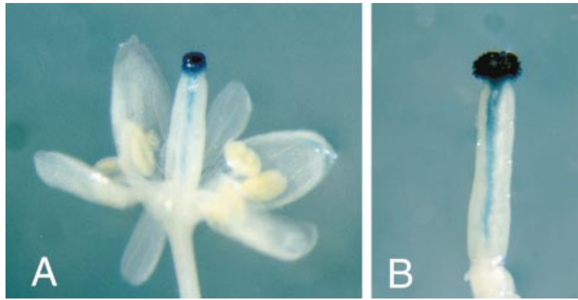


Figure 7. Activity of the Ψ*SRK* Gene Promoter.

(A) Whole flower showing pistil-specific GUS staining.
(B) Close-up view of a pistil showing GUS staining in the stigma epidermis and in the transmitting track of the style and ovary.
GUS activity was visualized by histochemical staining with 5-bromo-4-chloro-3-indolyl glucuronide.

the functional *SLG* and *SRK* genes of *Brassica* (Sato et al., 1991; Toriyama et al., 1991; Stein et al., 1996).

DISCUSSION

We have shown that the SI response of *A. lyrata* is similar to that of *Brassica* in its sporophytic control by a single specificity-determining locus, in its cytological manifestation by the inhibition of self-pollen at the surface of stigma epidermal cells, and in the molecular basis of self-recognition. These observations are consistent with a monophyletic origin of SI in the crucifer family.

Using a population derived by selfing a plant heterozygous for two *S* haplotypes that we designated *Sa* and *Sb*, we isolated from the *A. lyrata S* locus orthologs of the *Brassica SRK* and *SCR* genes that determine SI specificity in stigma and pollen. Our analysis of the two *S* haplotypes and the isolation of two alleles each for the *SRK* and *SCR* genes have demonstrated that these genes exhibit expression patterns that are largely similar to those of their *Brassica* orthologs. The two genes also occur in close physical proximity and exhibit extensive sequence polymorphism, as in *Brassica*. Interestingly, the ~37% sequence divergence exhibited by the *SRKa* and *SRKb* genes is comparable to the sequence divergence observed between class I and class II *SRK* alleles in *Brassica*, which implies an ancient divergence of the two alleles similar to that suggested for the two classes of *Brassica S* haplotypes. The *Sa* and *Sb* haplotypes might represent extremes in the spectrum of *S* haplotypic divergence for *A. lyrata*. The availability of *SRK* and *SCR* sequences from these *S* haplotypes should facilitate the isolation of other *SRK* and *SCR* alleles from this species and allow a study of allelic polymorphism and distribution in natural populations. Furthermore, the extensive sequence divergence of the *SRKa* and *SRKb* genes suggests that *A. lyrata S* haplotypes may fall into two distinct classes, as in *Brassica*. However, analysis of additional *A. lyrata S* haplotypes is needed to determine the range of *S* locus variability that exists in the species and to confirm the existence of two haplotypic classes. In any case, the observation that *SRKa* and *SRKb* sequences are equally diverged from the class I and class II *SRK* alleles of *Brassica* suggests that the *Sa* and *Sb* haplotypes of *A. lyrata* diverged from each other after the *Brassica–Arabidopsis* split. Alternatively, the phylogenetic relationships might reflect the occurrence of gene conversion or recombination events between *SRK* alleles within each lineage after speciation.

Even more divergent than the *SRK* alleles are the *SCRa* and *SCRb* alleles. The mature proteins predicted by *SCRa* and *SCRb* exhibit ~67% sequence divergence. Significantly, the *SCRa* and *SCRb* alleles also differ in their predicted signal peptide sequences. In this respect, they differ from the *Brassica SCR* alleles isolated to date, all of which are derived from class I *S* haplotypes and have highly conserved signal sequences (Schopfer et al., 1999; Nasrallah, 2000; Takayama et al., 2000; Watanabe et al., 2000). The signal sequence divergence in the *A. lyrata SCR* genes is consistent with an ancient divergence for these genes and with the conclusion that *Sa* and *Sb* belong to distinct classes of extant *S* haplotypes. Analysis of the *SCRa* and *SCRb* sequences thus adds to the body of evidence that the *SCR* gene is one of the most highly polymorphic genes in eukaryotes, exceeding the polymorphism of other recognition genes such as the pheromone genes that determine mating specificity in the ciliate *Euplotes raikovi* (Luporini et al., 1995) and the abalone sperm lysin gene that determines species-specific fertilization (Vacquier, 1998).

It is of interest that the *A. lyrata Sa* and *Sb* haplotypes appear to lack an *SLG*-like gene. *SLG*, which apparently was formed by a duplication of the *SRK* gene (Tantikanjana et al., 1993), is found in *Brassica* and *Raphanus* spp (Sakamoto et al., 1998). It occurs in tight genetic and physical linkage to the *SRK* and *SCR* genes and is expressed at high levels in plants homozygous for most *Brassica S* haplotypes, but it is absent (Okazaki et al., 1999) or expressed at low levels (Tantikanjana et al., 1993, 1996; Gaude et al., 1995; Cabrillac et al., 1999) in plants homozygous for a few *S* haplotypes. An *SRKb*-related sequence that may be genetically linked to the *S* locus was detected on DNA gel blots of *A. lyrata Sa* and *Sb* homozygotes. However, this sequence is unlikely to represent an *SLG* gene because it lacks a distinguishing feature of *SLG,* namely, that its alleles exhibit extensive *S* haplotype–associated polymorphisms that parallel those of the *SRK* gene. Nevertheless, our data do not rule out categorically the possibility that this *SRKb*-related sequence may be an *SLG* ortholog, albeit an atypical one.

The conclusion that the *A. lyrata Sa* and *Sb* haplotypes lack a "typical" *SLG* gene is supported by the observation that *A. lyrata* stigma RNA did not exhibit the high levels of *SRK*-homologous 1.6-kb transcripts that characterize most of the *Brassica* strains containing an *SLG* gene. The modest

amounts of 1.6-kb transcripts that are detected by *SRK S* domain probes in *A. lyrata* stigmas likely represent truncated *SRK* transcripts that correspond to the *S* domain of the gene. These short transcripts are produced by *SRK* genes (Stein et al., 1991; Giranton et al., 1995) and *SRK*-like genes of *Arabidopsis* (Tobias and Nasrallah, 1996). They are predicted to encode a secreted soluble form of the *SRK* ectodomain and have been shown to produce such a protein for at least one *Brassica SRK* allele (Giranton et al., 1995). The function of this class of protein is not known. However, the apparently general use of a polyadenylation signal within the first intron of *SRK* and *SRK*-like genes is suggestive of some functional significance for the soluble form of the receptor domain. It is possible that in *A. lyrata*, as well as in the self-incompatible *Brassica* plants that lack *SLG* or express low levels of SLG, the soluble form of SRK fulfills a role similar to that proposed for SLG and functions to enhance the SI response (Dixit et al., 2000; Takasaki et al., 2000).

Thus, analysis of the *A. lyrata S* locus has revealed both similarities and differences relative to the *S* locus of *Brassica*. The absence of an *SLG*-like gene in *A. lyrata* suggests that the ancestral *S* locus complex of crucifers consisted of the two SI specificity-determining genes *SRK* and *SCR*. The duplication event that gave rise to *SLG* must have occurred in a progenitor of the tribe Brassiceae/subtribe Brassicineae (which includes *Brassica* and *Raphanus*) and must have been lost subsequently in some members of these taxa. In this context, note that duplications, insertions, and deletions, all of which result in the restructuring and expansion or contraction of chromosomal regions, must have occurred frequently during the evolutionary history of the *S* locus. Clearly, these processes are limited neither to the *SLG* gene nor to *Brassica S* haplotypes. Indeed, in *A. lyrata*, the *Sb* haplotype differs from the *Sa* haplotype in the spacing and orientation of *SRK* and *SCR*, not only relative to each other but also relative to flanking markers.

The *Sb* haplotype also differs from the *Sa* haplotype by the presence of a tandem duplication of the *SCR* gene and its flanks, which contributes to but does not fully account for the approximately twofold expansion of the *SRK*/*SCR*–spanning region in the *Sb* haplotype relative to the *Sa* haplotype. The nature and origin of the remainder of the DNA in the expanded segments of the *S* haplotype remain to be determined, although some of this DNA consists of retroelement-like sequences, at least five of which were detected in the *Sb* haplotype. In any case, the structural heteromorphism of *S* haplotypes and the extreme divergence of SI specificity genes, both of which have been documented extensively in *Brassica*, appear to be common features of the *S* locus in the Brassicaceae. We conclude that the *S* locus of the Brassicaceae is a dynamic locus that has been, and perhaps continues to be, subjected to repeated rearrangements, deletions, and insertions. Consistent with this conclusion, the sequence identity of the *SCRb-1* and *SCRb-2* genes suggests that the direct repeat found in the *Sb* haplotype was duplicated quite recently. Note that rearranged gene order and fractured homology often are associated with complex loci and are thought to reduce the frequency of recombination within coadapted gene complexes (Ferris and Goodenough, 1994). Similarly, structural heteromorphism at the *S* locus might serve to maintain the tight association of the SI specificity genes over time (Boyes et al., 1997; Casselman et al., 2000).

An unexpected result of our study was the finding that the *S* locus occupies different chromosomal locations in *Brassica* and *Arabidopsis* spp. In *Brassica*, the *S* locus is located in a region that is syntenous with an *ETR1*-linked chromosomal segment of *A. thaliana* chromosome I (Conner et al., 1998), whereas in *A. lyrata*, the *S* locus maps to a region that corresponds to contig fragments 54 to 55 of *A. thaliana* chromosome IV. In both locations, the *S* locus is recognizable not only by the presence of its polymorphic SI specificity genes but also as an island exhibiting rearranged gene order and variable physical size that lies in the midst of an otherwise highly syntenous chromosomal region. These observations suggest that the *S* locus gene complex was translocated as a unit between two distant genomic locations.

When might this translocation have occurred? One possibility is that it occurred after the *Arabidopsis*–*Brassica* split. Translocation of the *S* locus presumably would have involved only one *S* haplotype and thus would represent a severe bottleneck on the number of *S* haplotypes in the derived lineage. Such an event, followed by diversification of the translocated locus and loss or inactivation of SI genes in the original chromosomal location, would explain the distinct locations of the *S* loci of *Brassica* and *Arabidopsis*. This scenario also is consistent with the phylogenetic analysis of *S* haplotypes, which, although limited, suggests a more recent divergence between *Brassica* class I and class II *S* haplotypes and between the *A. lyrata Sa* and *Sb* haplotypes than between *Brassica* and *A. lyrata S* haplotypes.

Further comparative mapping studies of the *S* locus in diverse taxa of the Brassicaceae are required to resolve the direction of *S* locus translocation events and the frequency with which they occur. Such studies also may elucidate the mechanisms that underlie *S* locus translocation in different lineages and those that underlie the restructuring of the *S* locus in a particular lineage. In this context, note that the *Sa* and *Sb* haplotypes of *A. lyrata*, the *S* locus region of *A. thaliana*, and several *Brassica S* haplotypes (Boyes and Nasrallah, 1995; Pastuglia et al., 1997b; Cui et al., 1999; Suzuki et al., 1999) all contain retroelement-like sequences. It is possible that retroelement activity, which is thought to be an effective agent of genome reorganization (Bennetzen, 1996), may have played a role in the translocation and restructuring of the *S* locus in the Brassicaceae.

The identification of the *ARK3*-linked region as the location of the *S* locus in *Arabidopsis* spp allowed us to identify orthologs of the *SRK* and *SCR* genes in *A. thaliana* and to understand the basis of self-compatibility in this species. The *A. thaliana SRK* ortholog is a nonfunctional gene that encodes a truncated ORF. We suspect that this Ψ*SRK* gene

has become inactivated relatively recently because it has remained transcriptionally active and its promoter has retained a pattern of activity identical to that described previously for the functional *Brassica SRK* genes in transgenic pistils (Stein et al., 1996). On the other hand, only partial remnants of the *SCR* gene were recognized in *A. thaliana*. Interestingly, the *A. thaliana* Ψ*SRK* and Ψ*SCR* sequences are more similar to the *A. lyrata SRKa* and *SCRa* genes than to *SRKb* and *SCRb,* and the relative orientation of Ψ*SRK* and the longest Ψ*SCR* sequence, Ψ*SCR1,* is similar to that of *SRKa* and *SCRa*. Although the position of ΨSRK and ΨSCR relative to the *ARK3* and *B80* flanking markers is different from that of the *SRKa* and *SCRa* genes, these relationships suggest that the nonfunctional *S* haplotype in the *A. thaliana* Columbia strain is derived from an *Sa*-like functional *S* haplotype.

The presence of the Ψ*SRK* and Ψ*SCR* sequences in *A. thaliana* indicates that the initial mutation that caused the shift to self-fertility in this species was not a complete deletion of the SI specificity genes but a more subtle mutation that led to the inactivation of one of these genes. The sequence of events that led to the inactivation of *SRK* and *SCR* cannot be surmised. However, because *SRK* and *SCR* function exclusively in SI, a mutation in either gene is expected to lead eventually to the decay of the other member of the SI recognition gene complex. Further analysis of different *A. thaliana* ecotypes will determine if all ecotypes contain the same nonfunctional *S* haplotype and the same lesions in their *SRK* and *SCR* orthologs as does the Columbia ecotype. Such a study might determine if the inactivation of SI genes occurred more than once in this species and might produce estimates of the number of *S* haplotypes that were included in the founding population of the *A. thaliana* lineage.

METHODS

Plant Material

Arabidopsis lyrata subsp *lyrata* is a herbaceous outcrossing perennial whose range in North America extends from Minnesota and Wisconsin south into Missouri, east into Georgia, north into Vermont, and west into Ontario (O'Kane and Al-Shehbaz, 1997). We used *A. lyrata* plants descended from accessions collected in Michigan (kindly provided by Charles Langley, University of California at Davis) and *A. thaliana* ecotype Columbia.

Pollination Analyses and Determination of the Self-Incompatibility Phenotype

Examination of pollen tube development at the stigma surface was performed using UV light fluorescence microscopy as described previously (Kho and Baer, 1968). To identify self-incompatibility (SI) genotype, we crossed several plants from a segregating population of 54 plants in all possible combinations. These crosses suggested a model of allelic interactions that was verified subsequently using selected homozygous and heterozygous plants as pollen parents in crosses to all plants in the segregating population to determine cross-compatibility relationships. Three groups of plants were identified as a result of this analysis (Table 1).

DNA Gel Blot Analysis

DNA was isolated from leaves according to Murray and Thompson (1980). The DNA was transferred to GeneScreen Plus nylon membranes (DuPont–New England Nuclear, Boston, MA), and the blots were prehybridized and hybridized at 65°C in 10% (w/v) dextran sulfate, 330 mM sodium phosphate, pH 7.0, 10 mM EDTA, and 5% (w/v) SDS. High-stringency washes were performed in a solution containing 0.2 × SSC (1 × SSC is 0.15 M NaCl and 0.015 M sodium citrate) and 0.1% (w/v) SDS at 65°C (in the analysis of the segregating population) or 68°C (in the restriction enzyme surveys).

RNA Isolation, Reverse Transcription–Polymerase Chain Reaction, and RNA Gel Blot Analysis

Pistils and anthers were dissected from flowers and floral buds at different stages of development. Total RNA was isolated by the Trizol method (Gibco BRL, Bethesda, MD). Poly(A)$^+$ RNA was purified using the FastTrack RNA isolation kit (Invitrogen, La Jolla, CA).

Reverse transcription–polymerase chain reaction (RT-PCR) was performed using the Advantage RT-for-PCR kit (Clontech, Palo Alto, CA). For amplification of *SRKa*, we used 10 ng of pistil poly(A)$^+$ RNA, the *S* domain primer 5′-CAAAGAATGGAACATG-3′, and the kinase domain primer 5′-TGATAGTCTTTTCACAGCAATCT-3′. Amplification of the 820-bp *SRKa* fragment was performed under the following conditions: 94°C for 5 min, 28 cycles of 94°C for 1 min, 55°C for 2 min, and 72°C for 2 min, and a final extension at 72°C for 5 min. The abundance of actin cDNA also was determined by PCR to verify that the PCRs contained equal amounts of cDNA. For amplification of actin, we used the following degenerate primers that were designed for the amplification of *A. thaliana* actin genes and kindly provided by M.K. Kandasamy (University of Georgia, Athens): 5′-TCYTTNCTNATRTCNACRTCRCAYTTCATDAT-3′ and 5′-GARAARATGACNCARATNATGTTYGARACNTT-3′. The 495-bp product was amplified under the following conditions: 94°C for 5 min, 32 cycles of 94°C for 1 min, 48°C for 2 min, and 72°C for 2 min, and a final extension at 72°C for 5 min.

For gel blot analysis, RNA was subjected to denaturing electrophoresis on 1% (w/v) agarose gels. After capillary transfer of the RNA to GeneScreen Plus membranes (DuPont–New England Nuclear), the blots were prehybridized and hybridized as described for DNA gel blot analysis. High-stringency washes were performed at 65°C in 0.2 × SSC and 0.1% (w/v) SDS (two washes of 10 min each), and low-stringency washes were performed at 65°C in 2 × SSC and 0.1% (w/v) SDS (two washes of 10 min each). The probes used in RNA gel blot analysis were either DNA probes generated by subcloning appropriate restriction fragments from genomic λ and bacterial artificial chromosome (BAC) clones into the pZero plasmid or probes derived from *SRK* and *SCR* cDNAs. The relative amounts of RNA loaded in different lanes were estimated by subsequent probing of the blots with a *Brassica* actin probe or by visualizing the intensity of rRNA staining with ethidium bromide. Quantitation of signal intensity was performed with a Molecular Dynamics PhosphorImager using the ImageQuant software package (Molecular Dynamics, Sunnyvale, CA).

Construction and Screening of cDNA Libraries

Three cDNA libraries were constructed in the bacteriophage vector λUniZAP: (1) a pistil cDNA library prepared with mRNA isolated from the pistils of an *SaSb* plant; (2) an anther cDNA library prepared with mRNA isolated from mixed stage anthers of an *SaSa* homozygote; and (3) an anther cDNA library prepared with mRNA isolated from mixed stage anthers of an *SbSb* homozygote.

The *SaSb* pistil library was initially screened at low stringency with a mixture of *Brassica S* domain probes derived from several class I and class II *SLG* genes and an *A. thaliana* probe derived from the *S* locus–related gene *AtS1* (Dwyer et al., 1992). Positive clones that also contained kinase sequences were identified by hybridization with a kinase domain probe derived from the *Brassica SRK6* gene. This screen resulted in the isolation of the *SRKa* sequence. Subsequent screening of the pistil library with a 550-bp XhoI fragment derived from an *Sb* BAC clone resulted in the isolation of the *SRKb* sequence.

The *SCRa* and *SCRb* cDNA clones were isolated by screening the *SaSa* and *SbSb* anther cDNA libraries, respectively, with *A. lyrata*–derived DNA probes that were generated by subcloning appropriate restriction fragments from genomic λ and BAC clones into the pZero plasmid (Invitrogen) or λDashII (Stratagene, La Jolla, CA), respectively, and subsequent purification of the insert.

Construction and Screening of *A. lyrata* Genomic λ and BAC Libraries

For the construction of a bacteriophage library, *A. lyrata SaSb* DNA was partially digested with Sau3A1, and a fraction containing fragments of 9 to 20 kb was cloned into the BamHI site of the λGEM11 vector (Promega, Madison, WI). The library was screened with an *S* domain probe derived from the *SRKa* cDNA to isolate *SRKa*-containing clones.

A large insert library of *A. lyrata SaSb* DNA was constructed in the pBeloBAC vector according to Woo et al. (1994). Because of the existence of a large family of genes related to *SRK* in crucifer genomes, the isolation of BAC clones for the *A. lyrata S* locus required the generation of single-copy molecular markers from the region. We searched the *A. thaliana* BAC T6K22 sequence for sequences that were apparently unique in the *A. thaliana* genome. One such sequence, T6K22.150, was amplified from BAC T6K22 by using the gene-specific primers 5′-TTGCTTATCGCCGTGGTTCC-3′ and 5′-AGCGACTCGATTTCCTTCTCTGTA-3′. Subsequent gel blot analysis of *A. lyrata* DNA demonstrated that the T6K22.150 probe produced a relatively strong hybridization signal and appeared to be a single-copy sequence in *A. lyrata* (data not shown). This probe was used to screen the *A. lyrata* BAC library. A total of 15 BAC clones were identified that, upon further analysis, were classified into two groups, one group of six clones derived from the *Sa* haplotype and another group of nine clones derived from the *Sb* haplotype.

Sequence Analysis and Database Searches

DNA sequencing was performed using an automated sequencer at the sequencing facility of the Cornell University Bioresource Center. Sequences were manipulated and aligned using DNASTAR Lasergene software (DNASTAR Inc., Madison, WI). BLAST searches were performed on the National Center for Biotechnology Information Internet site (http://www.ncbi.nlm.nih.gov). Sequences were aligned using the MegAlign feature of DNASTAR. Parsimony analyses were conducted on the aligned cDNAs or derived protein sequences by using the branch and bound algorithm of PAUP* 4.0b2 (Swofford, 1999).

Construction of Reporter Gene Fusions, Plant Transformation, and Analysis of Transgenic Plants

A DNA fragment spanning a region from 23 to ~1600 bp upstream of the putative initiating methionine codon of the T6K22.100 (Ψ*SRK*) sequence was generated by PCR amplification and two primers, 5′-GGATCCGAGATTAGGTTTGTACATG-3′ and 5′-GGAAAATGAACA-ACATATGAGGATCC-3′, which incorporated BamHI recognition sites to facilitate cloning. The amplified product was TA cloned into the pCRII vector (Invitrogen) and digested with BamHI. The promoter insert was excised from this vector by BamHI digestion and inserted into the BamHI site of the β-glucuronidase (GUS) expression vector pBI101 (Jefferson et al., 1987) 5′ of the *uidA* gene, which encodes the GUS protein. A plasmid containing the promoter fragment in the appropriate orientation was selected and introduced into *Agrobacterium tumefaciens* strain pCIB542/A136 (derived from helper plasmid EHA101; Hood et al., 1986). Transformation of *A. thaliana* strain C24 (obtained from M. Jacobs, Vrije Universiteit, Brussels, Belgium) with the Ψ*SRK*::*uida*::*nos* chimeric gene was performed according to Valvekens et al. (1988). Transgenic plants were selected on the basis of kanamycin resistance, and the stable integration of the transgene was confirmed by DNA gel blot analysis in primary transformants and in progeny plants generated by allowing the plants to self-pollinate.

The Ψ*SRK*::*uida*::*nos* primary transformants and their T2 progeny were analyzed for GUS activity by the histochemical method using the chromogenic substrate 5-bromo-4-chloro-3-indolyl glucuronide (Jefferson et al., 1987), as described previously (Toriyama et al., 1991).

ACKNOWLEDGMENTS

We thank Charles Langley (University of California, Davis) for providing the *A. lyrata* seed, Jeffrey Doyle (Cornell University, Ithaca, NY) for assistance with the phylogenetic analysis, M.K. Kandasamy (University of Georgia, Athens) for actin primers, and Bela Kudish and Michelle Battle for work related to the construction of the Ψ*SRK* reporter gene and plant transformation. We thank the Arabidopsis Biological Resources Center for providing BAC T6K22. This work was supported by Grant No. IBN-0077289 from the National Science Foundation.

Received October 11, 2000; accepted January 15, 2001.

REFERENCES

Al-Shehbaz, I.A. (1984). The tribes of the Cruciferae (Brassicaceae) in the southeastern United States. J. Arnold Arbor. Harv. Univ. **65,** 343–373.

Bateman, A.J. (1954). Self-incompatibility systems in angiosperms. II. *Iberis amara*. Heredity **8,** 305–332.

Bateman, A.J. (1955). Self-incompatibility systems in angiosperms. III. Cruciferae. Heredity **9,** 52–68.

Bennetzen, J.L. (1996). The contributions of retroelements to plant genome organization, function and evolution. Trends Microbiol. **4,** 347–353.

Boyes, D.C., and Nasrallah, J.B. (1993). Physical linkage of the *SLG* and *SRK* genes at the self-incompatibility locus of *Brassica oleracea*. Mol. Gen. Genet. **236,** 369–373.

Boyes, D.C., and Nasrallah, J.B. (1995). An anther-specific gene encoded by an *S* locus haplotype of *Brassica* produces complementary and differentially regulated transcripts. Plant Cell **7,** 1283–1294.

Boyes, D.C., Nasrallah, M.E., Vrebalov, J., and Nasrallah, J.B. (1997). The self-incompatibility (*S*) haplotypes of *Brassica* contain highly divergent and rearranged sequences of ancient origin. Plant Cell **9,** 1–12.

Cabrillac, D., Delorme, V., Garin, J., Ruffio-Chable, V., Giranton, J.-L., Dumas, C., Gaude, T., and Cock, M.J. (1999). The *S15* self-incompatibility haplotype in *Brassica oleracea* includes three *S* gene family members expressed in stigmas. Plant Cell **11,** 971–986.

Casselman, A.L., Vrebalov, J., Conner, J.A., Singhal, A., Giovannoni, J., Nasrallah, M.E., and Nasrallah, J.B. (2000). Determining the physical limits of the *Brassica S* locus by recombinational analysis. Plant Cell **12,** 23–33.

Chen, C.-H., and Nasrallah, J.B. (1990). A new class of *S*-sequences defined by a pollen recessive self-incompatibility allele of *Brassica oleracea*. Mol. Gen. Genet. **222,** 241–248.

Conner, J.A., Tantikanjana, T., Stein, J.C., Kandasamy, M.K., Nasrallah, J.B., and Nasrallah, M.E. (1997). Transgene-induced silencing of *S* locus genes and related genes in *Brassica*. Plant J. **11,** 809–823.

Conner, J.C., Conner, P., Nasrallah, M.E., and Nasrallah, J.B. (1998). Comparative mapping of the *Brassica S*-locus region and its homeologue in Arabidopsis: Implications for the evolution of mating systems in the Brassicaceae. Plant Cell **10,** 801–812.

Cui, Y.H., Brugiere, N., Jackman, L., Bi, Y.M., and Rothstein, S.J. (1999). Structural and transcriptional comparative analysis of the *S* locus regions in two self-incompatible *Brassica napus* lines. Plant Cell **11,** 2217–2231.

Cui, Y.H., Bi, Y.M., Brugiere, N., Arnoldo, M.A., and Rothstein, S.J. (2000). The S locus glycoprotein and the S receptor kinase are sufficient for self-pollen rejection in *Brassica*. Proc. Natl. Acad. Sci. USA **97,** 3713–3717.

de Nettancourt, D. (1977). Incompatibility in Angiosperms: Monographs on Theoretical and Applied Genetics 3. (Berlin: Springer-Verlag).

Dixit, R., Nasrallah, M.E., and Nasrallah, J.B. (2000). Post-transcriptional maturation of the S receptor kinase of *Brassica* correlates with co-expression of the S-locus glycoprotein in the stigmas of two *Brassica* strains and in transgenic tobacco plants. Plant Physiol. **124,** 297–312.

Dwyer, K.G., Balent, M.A., Nasrallah, J.B., and Nasrallah, M.E. (1991). DNA sequences of self-incompatibility genes from *Brassica campestris* and *B. oleracea*: Polymorphism predating speciation. Plant Mol. Biol. **16,** 481–486.

Dwyer, K.G., Lalonde, B.A., Nasrallah, J.B., and Nasrallah, M.E. (1992). Structure and expression of *AtS1*, an *Arabidopsis thaliana* gene homologous to the *S*-locus related genes of *Brassica*. Mol. Gen. Genet. **231,** 442–448.

Dwyer, K.G., Kandasamy, M.K., Mahosky, D.I., Acciai, J., Kudish, B.I., Miller, J.E., Nasrallah, M.E., and Nasrallah, J.B. (1994). A superfamily of *S* locus-related sequences in *Arabidopsis*: Diverse structures and expression patterns. Plant Cell **6,** 1829–1843.

Ferris, P.J., and Goodenough, U.W. (1994). The mating-type locus of *Chlamydomonas reinhardtii* contains highly rearranged DNA sequences. Cell **76,** 1135–1145.

Gaude, T., Rougier, M., Heizmann, P., Ockendeon, D.J., and Dumas, C. (1995). Expression level of the *SLG* gene is not correlated with the self-incompatibility phenotype in the class II *S* haplotypes of *Brassica oleracea*. Plant Mol. Biol. **27,** 1003–1014.

Giranton, J.L., Ariza, M.J., Dumas, C., Cock, J.M., and Gaude, T. (1995). The *S* locus receptor kinase gene encodes a soluble glycoprotein corresponding to the SRK extracellular domain in *Brassica oleracea*. Plant J. **8,** 827–834.

Hood, E.E., Helmer, G.L., Fraley, R.T., and Chilton, M.D. (1986). The hypervirulence of *Agrobacterium tumefaciens* A281 is encoded in a region of pTiBo542 outside of T-DNA. J. Bacteriol. **168,** 1291–1301.

Jefferson, R.A., Kavanaugh, T.A., and Bevan, M.W. (1987). GUS fusions: β-Glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. **6,** 3901–3907.

Kandasamy, M.K., Paolillo, D.J., Faraday, C.D., Nasrallah, J.B., and Nasrallah, M.E. (1989). The S-locus specific glycoproteins of *Brassica* accumulate in the cell wall of developing stigma papillae. Dev. Biol. **134,** 462–472.

Kho, Y.O., and Baer, J. (1968). Observing pollen tubes by means of fluorescence. Euphytica **17,** 298–302.

Koch, M., Bishop, J., and Mitchell-Olds, T. (1999). Molecular systematics and evolution of *Arabidopsis* and *Arabis*. Plant Biol. **1,** 529–537.

Kusaba, M., and Nishio, T. (1999). Comparative analysis of *S* haplotypes with very similar *SLG* alleles in *Brassica rapa* and *Brassica oleracea*. Plant J. **17,** 83–91.

Kusaba, M., Nishio, T., Satta, Y., Hinata, K., and Ockendon, D.J. (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.

Luporini, P., Vallesi, A., Miceli, C., and Bradshaw, R.A. (1995). Chemical signaling in ciliates. J. Eukaryot. Microbiol. **42,** 208–212.

Murray, M.G., and Thompson, W.F. (1980). Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. **8,** 4321–4325.

Nasrallah, J.B. (2000). Cell-cell signaling in the self-incompatibility response. Curr. Opin. Plant Biol. **3,** 368–373.

Nasrallah, J.B., Kao, T.-H., Chen, C.-H., Goldberg, M.L., and Nasrallah, M.E. (1987). Amino-acid sequence of glycoproteins encoded by three alleles of the *S* locus of *Brassica oleracea*. Nature **326,** 617–619.

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

O'Kane, S.L., and Al-Shehbaz, I.A. (1997). A synopsis of *Arabidopsis* (Brassicaceae). Novon **7,** 323–327.

Okazaki, K., Kusaba, M., Ockendon, D.J., and Nishio, T. (1999). Characterization of *S* tester lines in *Brassica oleracea*: Polymorphism

of restriction fragment length of *SLG* homologues and isoelectric points of S-locus glycoproteins. Theor. Appl. Genet. **98,** 1329–1334.

Pastuglia, M., Roby, D., Dumas, C., and Cock, J.M. (1997a). Rapid induction by wounding and bacterial infection of an *S* gene family receptor-like kinase in *Brassica oleracea*. Plant Cell **9,** 1–13.

Pastuglia, M., Ruffio-Chable, V., Delorme, V., Gaude, T., Dumas, C., and Cock, J.M. (1997b). A functional *S* locus anther gene is not required for the self-incompatibility response in *Brassica oleracea.* Plant Cell **9,** 2065–2076.

Price, R.A., Palmer, J.D., and Al-Shehbaz, I.A. (1994). Systematic relationships of *Arabidopsis*: A molecular and morphological perspective. In *Arabidopsis*, C.R. Somerville and E.M. Meyerowitz, eds (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 7–19.

Sakamoto, K., Kusaba, M., and Nishio, T. (1998). Polymorphism of the *S*-locus glycoprotein gene (*SLG*) and the *S*-locus related gene (*SLR1*) in *Raphanus sativus* L. and self-incompatible ornamental plants in the Brassicaceae. Mol. Gen. Genet. **258,** 397–403.

Sato, T., Thorsness, M.K., Kandasamy, M.K., Nishio, T., Hirai, M., Nasrallah, J.B., and Nasrallah, M.E. (1991). Activity of an *S*-locus gene promoter in pistils and anthers of transgenic *Brassica*. Plant Cell **3,** 867–876.

Schierup, M.H. (1998). The effect of enzyme heterozygosity on growth in a strictly outcrossing species, the self-incompatible *Arabis petraea* (Brassicaceae). Hereditas **128,** 21–31.

Schierup, M.H., Mable, B., Awadalla, P., and Charlesworth, D. (2001). Identification and characterization of a polymorphic receptor kinase gene linked to the self-incompatibility locus of *Arabidopsis lyrata*. Genetics (in press).

Schopfer, C.R., and Nasrallah, J.B. (2000). Self-incompatibility: Prospects for a novel putative peptide-signaling molecule. Plant Physiol. **124,** 935–940.

Schopfer, C.R., Nasrallah, M.E., and Nasrallah, J.B. (1999). The male determinant of self-incompatibility in *Brassica*. Science **286,** 1697–1700.

Schultz, O.E. (1936). Cruciferae. In Die Naturlichen Pflanzenfamilien, Vol. 17b, 2nd ed, A. Engler and H. Harms, eds (Leipzig, Germany: Engelmann), pp. 227–658.

Stein, J.C., Howlett, B., Boyes, D.C., Nasrallah, M.E., and Nasrallah, J.B. (1991). Molecular cloning of a putative receptor protein kinase gene encoded at the self-incompatibility locus of *Brassica oleracea*. Proc. Natl. Acad. Sci. USA **88,** 8816–8820.

Stein, J.C., Dixit, R., Nasrallah, M.E., and Nasrallah, J.B. (1996). SRK, the stigma-specific S locus receptor kinase of *Brassica*, is targeted to the plasma membrane in transgenic tobacco. Plant Cell **8,** 429–445.

Suzuki, G., Kai, N., Hirose, T., Fukui, K., Nishio, T., Takayama, S., Isogai, A., Watanabe, M., and Hinata, K. (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 4.0. (Sunderland, MA: Sinauer Associates).

Takahata, Y., and Hinata, K. (1980). A variation study of subtribe Brassicineae by principal component analysis. In Brassica Crops and Wild Allies: Biology and Breeding, S. Tsunoda, K. Hinata, and C. Gomez-Campo, eds (Tokyo: Japan Science Societies Press), pp. 33–49.

Takasaki, T., Hatakeyama, K., Suzuki, G., Watanabe, M., Isogai, A., and Hinata, K. (2000). The *S* receptor kinase determines self-incompatibility in *Brassica* stigma. Nature **403,** 913–916.

Takayama, S., Shiba, H., Iwano, M., Shimosato, H., Che, F.S., Kai, N., Watanabe, M., Suzuki, G., Hinata, K., and Isogai, A. (2000). The pollen determinant of self-incompatibility in *Brassica campestris*. Proc. Natl. Acad. Sci. USA **97,** 1920–1925.

Tantikanjana, T., Nasrallah, M.E., Stein, J.C., Chen, C.-H., and Nasrallah, J.B. (1993). An alternative transcript of the *S* locus glycoprotein gene in a class II pollen-recessive self-incompatibility haplotype of *Brassica oleracea* encodes a membrane-anchored protein. Plant Cell **5,** 657–666.

Tantikanjana, T., Nasrallah, M.E., and Nasrallah, J.B. (1996). The *Brassica S* gene family: Molecular characterization of the *SLR2* gene. Sex. Plant Reprod. **9,** 107–116.

Thompson, K.F., and Taylor, J.P. (1966). Non-linear dominance relationships between *S* alleles. Heredity **21,** 345–362.

Tobias, C.M., and Nasrallah, J.B. (1996). An *S*-locus-related gene in *Arabidopsis* encodes a functional kinase and produces two classes of transcripts. Plant J. **10,** 523–531.

Tobias, C.M., Howlett, B., and Nasrallah, J.B. (1992). An *Arabidopsis thaliana* gene with sequence similarity to the *S*-locus receptor kinase gene of *Brassica oleracea*. Plant Physiol. **99,** 284–290.

Toriyama, K., Thorsness, M.K., Nasrallah, J.B., and Nasrallah, M.E. (1991). An *S*-locus gene promoter directs sporophytic expression in the anther tapetum of transgenic *Arabidopsis*. Dev. Biol. **143,** 427–431.

Uyenoyama, M.K. (1995). A generalized least-squares estimate for the origin of sporophytic self-incompatibility. Genetics **139,** 975–992.

Uyenoyama, M.K. (2000). A prospectus for new developments in the evolutionary theory of self-incompatibility. Ann. Bot. **85** (suppl. A), 247–252.

Vacquier, V.D. (1998). Evolution of gamete recognition proteins. Science **281,** 1995–1998.

Valvekens, D., Van Montagu, M., and Van Lijsebettens, M. (1988). *Agrobacterium tumefaciens*–mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. Proc. Natl. Acad. Sci. USA **85,** 5536–5540.

van Treuren, R., Kuitteinen, H., Karkkainen, K., Baena-Gonzalez, E., and Savolainen, O. (1997). Evolution of microsatellites in *Arabis petraea* and *Arabis lyrata*, outcrossing relatives of *Arabidopsis thaliana*. Mol. Biol. Evol. **14,** 220–229.

Watanabe, M.A., et al. (2000). Highly divergent sequences of the pollen self-incompatibility (*S*) gene in class-I *S* haplotypes of *Brassica campestris* (syn. *rapa*) L. FEBS Lett. **473,** 139–144.

Woo, S.-S., Jiang, J., Gill, B.S., Patterson, A.H., and Wing, R.A. (1994). Construction and characterization of a bacterial artificial chromosome library of *Sorghum bicolor*. Nucleic Acids Res. **22,** 4922–4931.

Yang, Y.W., Lai, K.N., Tai, P.Y., Ma, D.P., and Li, W.H. (1999). Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. J. Mol. Evol. **48,** 597–604.