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 Brassicinae 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...
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., 1993; 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 (Takayama et al., 2000). This implies that the absence of SLG results in a high degree of sequence similarity to the extracellular domain or the SCR domain of *Arabidopsis* SRK (for S locus kinase) 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 Brassicaceae. 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 Brassicaceae 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, *Arabi-

**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

<table>
<thead>
<tr>
<th>Ovule Parent</th>
<th>Pollen Parent</th>
<th>SaSa (18)</th>
<th>SaSb (27)</th>
<th>SbSb (9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SaSa (18)</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>SaSb (27)</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>SbSb (9)</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

*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 SaSa 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
of our segregating population with the 550-bp Xhol probe confirmed the linkage of this gene to the S locus. Under high-stringency wash conditions, the 550-bp Xhol 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 using S domain probes from 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), Xhol (X), and EcoRI (E) and probed with the S domain of SRKa (C) and SRKb (D).

**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), Xhol (X), and EcoRI (E) and probed with the S domain of SRKa (C) and SRKb (D).

de restriction fragment in SacI and EcoRV digests. Two restriction fragments were detected in Xhol digests due to the presence of an internal Xhol 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 SRKs. 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 Xhol 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 Xhol 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* 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., 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 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 pl 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 pl 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
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 SRKα and SRKβ proteins. The SRKα 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).

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 loci 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](Image)

**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 3](Image)

**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.
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 ˈSCR1, ˈSCR2, and ˈSCR3 (Figure 6A). The ˈSCR1 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 ˈSCR2 and ˈSCR3 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 SCRα signal sequence (Figure 6A) as well as with the SCRα 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
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 CSRK.

The isolation of cDNA clones for the CSRK 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 CSRK 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 CSRK::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 CSRK promoter exhibits a spatial activity pattern in pistils identical to that of the two short ψSCR2 and ψSCR3 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.

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 CSRK gene of A. thaliana. One filled box is used to show the location of the SCRb gene. Sequences shared by the three maps are connected by dashed lines. E, EagI; S, SmaI.
Figure 6. The \( \Psi \text{SRK} \) and \( \Psi \text{SCR} \) Sequences of \( A. \) thaliana.

(A) Alignment of the \( \Psi \text{SCR1}, \Psi \text{SCR2}, \) and \( \Psi \text{SCR3} \) deduced amino acid sequences with relevant regions of \( A. \) lyrata \( \text{SCRa} \) and \( \text{SCRb} \). The arrowhead shows the position of the intron. The cysteine residues and glycine residue that are conserved in \( \text{SCR} \) peptides are marked by boxes. The residues in the \( \Psi \text{SCR1}, \Psi \text{SCR2}, \) and \( \Psi \text{SCR3} \) sequences that are shared with \( A. \) lyrata \( \text{SCRa} \) are shown in boldface. The first and second exons of \( \Psi \text{SCR1} \) are located, respectively, at positions 50,496 to 50,562 and 51,344 to 51,472 of the BAC T6K22 sequence. The \( \Psi \text{SCR2} \) and \( \Psi \text{SCR3} \) 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 \( \Psi \text{SRK} \) gene. The exon–intron structure of the \( \Psi \text{SRK} \) 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 \( \Psi \text{SRK} \) 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.
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* haploptic 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* 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 haploptic 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 *SCRa* alleles. The mature proteins predicted by *SCRa* and *SCRa* exhibit ~67% sequence divergence. Significantly, the *SCRa* and *SCRa* 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 *SCRa* 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 ptheromone 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* haploptic–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 transcripts are produced by SRK genes (Stein et al., 1991; Giranton et al., 1995) and SRK-like genes of Arabidopsis (Tobias and Nasrallah, 1995). 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 *ETRI*-linked chromosomal segment of *A. thaliana* chromosome I (Conner et al., 1995), 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 *vSRK* 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 \Psi SRK and \Psi SCR sequences are more similar to the A. lyrata SRKa and SCRa genes than to SRKb and SCRa, and the relative orientation of \Psi SRK and the longest \Psi SCR sequence, \Psi SCR1, is similar to that of SRKa and SCRa. Although the position of \Psi SRK and \Psi SCR 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 S-like functional S haplotype.

The presence of the \Psi SRK and \Psi 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) genotypes, 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)¹ 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)¹ RNA, the S domain primer 5’C-GAAAGTGGAAACATG-3’, and the kinase domain primer 5’-TGAATGCAGCAGAATC-3’. Amplification of the 820-bp SRKb 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 was also determined by PCR to verify that the PCR products contamed 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’-TCTTCTTATATRCONACRTORCAYTYCCATAT-3’ and 5’-GARAARATGACNCARATNATGYGARAGNTT-3’. The 495-bp product was amplified under the following conditions: 94°C for 5 min, 32 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.

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 nylon 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 AIS1 (Dwyer et al., 1992). Positive clones that also contained kinase sequences were identified by hybridization with a kinase domain probe derived from the Brassica SRKβ gene. This screen resulted in the isolation of the SRKα sequence. Subsequent screening of the pistil library with a 550-bp XhoI fragment derived from a SaSb BAC clone resulted in the isolation of the SRKb sequence.

The SCFa and SCFB 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.

Search the 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.

Construc...


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

Makoto Kusaba, Kathleen Dwyer, Jennifer Hendershot, Julia Vrebalov, June B. Nasrallah and Mikhail E. Nasrallah

*Plant Cell* 2001;13:627-643
DOI 10.1105/tpc.13.3.627

This information is current as of December 22, 2017

<table>
<thead>
<tr>
<th>References</th>
<th>This article cites 59 articles, 24 of which can be accessed free at: /content/13/3/627.full.html#ref-list-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>eTOCs</td>
<td>Sign up for eTOCs at: <a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a></td>
</tr>
<tr>
<td>CiteTrack Alerts</td>
<td>Sign up for CiteTrack Alerts at: <a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a></td>
</tr>
<tr>
<td>Subscription Information</td>
<td>Subscription Information for <em>The Plant Cell</em> and <em>Plant Physiology</em> is available at: <a href="http://www.aspb.org/publications/subscriptions.cfm">http://www.aspb.org/publications/subscriptions.cfm</a></td>
</tr>
</tbody>
</table>