Copyright © 2001 by the Genetics Society of America

Intrahaplotype Polymorphism at the Brassica *S* Locus

Christine Miege,* Véronique Ruffio-Châble,† Mikkel H. Schierup,‡ Didier Cabrillac,* Christian Dumas,* Thierry Gaude* and J. Mark Cock*

**Reproduction et Développement des Plantes, UMR 5667 CNRS-INRA-ENSL, Ecole Normale Supérieure de Lyon, 69364 Lyon, France, †UMR Amélioration des Plantes et Biotechnologies Végétales, Institut National de la Recherche Agronomique, 35650 Le Rheu, France and ‡Department of Ecology and Genetics, University of Aarhus, DK-8000 Aarhus C, Denmark*

Manuscript received May 3, 2001
Accepted for publication July 16, 2001

ABSTRACT

The S locus receptor kinase and the S locus glycoproteins are encoded by genes located at the *S* locus, which controls the self-incompatibility response in Brassica. In class II self-incompatibility haplotypes, S locus glycoproteins can be encoded by two different genes, *SLGA* and *SLGB*. In this study, we analyzed the sequences of these genes in several independently isolated plants, all of which carry the same *S* haplotype (S_2). Two groups of S_2 haplotypes could be distinguished depending on whether *SRK* was associated with *SLGA* or *SLGB*. Surprisingly, *SRK* alleles from the two groups could be distinguished at the sequence level, suggesting that recombination rarely occurs between haplotypes of the two groups. An analysis of the distribution of polymorphisms along the S domain of SRK showed that hypervariable domains I and II tend to be conserved within haplotypes but to be highly variable between haplotypes. This is consistent with these domains playing a role in the determination of haplotype specificity.

MOST flowering plant species are hermaphrodite and possess flowers in which male (stamen) and female (pistil) reproductive organs are in close proximity. However, outcrossing is favored in the majority of these species due to the presence of self-incompatibility (SI) systems that allow the pistil to recognize and reject self-pollen. In Brassica, SI is controlled by a single genetic locus (the *S* locus) and self-pollen is rejected on the stigma surface (reviewed in Cock 2000; Nasrallah 2000).

One of the most striking features of the *S* locus is its extremely high level of polymorphism. For example, >80 different *S* locus alleles (or haplotypes) have been identified in *Brassica oleracea* (Ruffio-Châble and Gaude 2001). This high level of polymorphism is thought to be the result of balancing selection, which maintains large numbers of *S* haplotypes in Brassica populations over long periods of time (Uyenoyama 1995). Another consequence of this unusual mode of evolution is that *S* haplotypes exhibit *trans*-specific evolution; they are often more similar to an *S* haplotype from another species in the Brassicaceae than to other *S* haplotypes from the same species (Dwyer *et al.* 1991; Sakamoto *et al.* 1998).

Physical mapping and sequence analysis of different *S* haplotypes has revealed large-scale rearrangements that can extend to several hundred kilobases in this region of the genome (Boyes and Nasrallah 1993; Boyes *et al.* 1997; Cui *et al.* 1999; Suzuki *et al.* 1999). Several expressed genes are located within this highly polymorphic region and two of these genes have recently been shown to play key roles in the SI response. Takasaki *et al.* (2000) used a transgenic approach to show that *SRK* encodes the female component of the SI response. *SRK* encodes the *S* locus receptor kinase, a membrane-anchored stigma glycoprotein that shares structural similarity with animal receptor protein kinases (Stein *et al.* 1991; Delorme *et al.* 1995; Stein *et al.* 1996). *SCR* (or *SP11*) was shown to encode the male component of the self-incompatibility response (Schopfer *et al.* 1999; Takayama *et al.* 2000). *SCR* encodes the *S* locus cysteine-rich protein, a small secreted protein that is thought to be located in the pollen coating (Stephenson *et al.* 1997; Schopfer *et al.* 1999; Takayama *et al.* 2000).

While the functions of *SRK* and *SCR* seem to be relatively clear, that of a third *S* locus gene, *SLG* (*S locus glycoprotein*), is less well understood. *SLG* encodes a secreted glycoprotein that closely resembles the extracellular (or S) domain of SRK. The similarity between these two genes is often particularly marked within haplotypes, indicating concerted evolution (Stein *et al.* 1991; Kusaba *et al.* 1997; Nishio and Kusaba 2000). Initially, several pieces of evidence suggested that SLG functioned in the SI response (Toriyama *et al.* 1991; Shiba *et al.* 1995). However, more recent evidence argues against a role for this protein in the recognition of self-pollen (Gaude *et al.* 1995; Cabrillac *et al.* 1999; Okazaki *et al.* 1999; Nishio and Kusaba 2000) although SLG may play an auxiliary role by mediating maturation of the SRK protein (Dixit *et al.* 2000). If this is the case,

Corresponding author: J. Mark Cock, Reproduction et Développement des Plantes, UMR 5667 CNRS-INRA-ENSL, Ecole Normale Supérieure de Lyon, 46 allée d'Italie, 69364 Lyon Cedex 07, France.
E-mail: mark.cock@ens-lyon.fr

SLG is probably not the only protein that can carry out this function as self-incompatible plants that lack SLG have been identified (Okazaki *et al.* 1999; Nishio and Kusaba 2000). It has been suggested that a truncated form of SRK, eSRK (Giranton *et al.* 1995), may substitute for SLG in these plants (Dixit *et al.* 2000). Takasaki *et al.* (2000) showed that coexpression of SLG and SRK in transgenic plants resulted in a stronger SI response than expression of SRK alone, again supporting the hypothesis that SLG plays an auxiliary role. Another function that has been proposed for SLG is as one of the components on the stigma surface that mediate pollen adhesion (Luu *et al.* 1999). In the latter case, SLG would function independently of the SI system as pollen adhesion is not influenced by the SI response (Luu *et al.* 1997).

Both *SLG* and *SRK* are highly polymorphic. The alleles of these genes can be grouped into two distinct classes, class I and class II, and there is a correlation between these classes and observed SI phenotypes. SI in Brassica is sporophytic and a nonlinear hierarchy of dominance relationships has been shown to exist between *S* haplotypes (Thompson and Taylor 1966). Class I haplotypes tend to be dominant and to confer a strong SI response whereas class II haplotypes tend to be recessive to the class I haplotypes and to confer a weaker SI response. Two different *SLG* genes, *SLGA* and *SLGB*, have recently been shown to be associated with class II *S* haplotypes (Cabrillac *et al.* 1999). Both *SLGA* and *SLGB* contain a single intron but *SLGA* differs from *SLGB* in that its second exon encodes a membrane-spanning domain, allowing the production of alternative transcripts that encode both secreted and membrane-anchored proteins (Tantikanjana *et al.* 1993; Cabrillac *et al.* 1999).

Data from genetic crosses indicate that each *S* haplotype encodes an independent recognition system. Self-pollen rejection is initiated as a result of the interaction of male and female components encoded by the same *S* haplotype. In principle, therefore, comparison of different haplotypes should permit the identification of sequence motifs that determine haplotype specificity. This approach has been used profitably, for example, as an aid to functional analysis of polymorphic disease resistance gene loci (Ellis *et al.* 1999). However, a drawback with applying this approach to self-incompatibility systems is the very high level of polymorphism at the *S* locus. For example, searches for regions that potentially determine haplotype specificity in the S domain of SRK have identified several extremely polymorphic hypervariable (HV) regions (Nasrallah *et al.* 1987; Kusaba *et al.* 1997). However, the level of polymorphism throughout the rest of the S domain is also very high and it is possible that the HV regions are simply under a low level of selective constraint and that this is sufficient to explain the higher level of polymorphism in these regions (Nasrallah 1997). The approach that we have taken here, to circumvent this problem, is to compare not only *S* haplotypes that confer different SI specificities ("between haplotype" comparisons) but also several independently isolated *S* haplotypes that confer the same SI specificity ("within haplotype" comparisons). Regions under relaxed selective constraint are expected to be more divergent (compared to the rest of the protein) in within-haplotype comparisons whereas regions involved in determining haplotype specificity should be conserved in genes that confer the same SI specificity. Sequence heterogeneity at *S* locus gene loci has been previously reported for the S_2 haplotype (Kusaba *et al.* 2000). Here we analyze an additional 10 S_2 haplotypes. Two groups of S_2 haplotype were defined, depending on the presence of either only *SLGA* or only *SLGB*. We show that, at the sequence level, hypervariable domains HV1 and HV2 are conserved in independent SRK_2 alleles but are divergent between haplotypes, indicating a role for these domains in determining *S* haplotype specificity.

MATERIALS AND METHODS

Plant material and pollination analysis: The seven S_2 *B. oleracea* plants analyzed in this study were from diverse origins and include varieties of kale, broccoli, and Brussels sprouts (Table 1). Incompatibility phenotypes were determined by self- and cross-pollinations using previously described procedures (Ruffio-Châble and Gaude 2001). Homozygous S_6 and S_8 tester lines were from the collection at the Genetic Resources Unit, Horticulture Research International (Wellesbourne, United Kingdom), which has recently been duplicated at INRA, Le Rheu (Rennes, France).

Immunological detection of SLG proteins in stigma extracts: Protein extraction, separation of proteins by isoelectric focusing, electrotransfer onto nitrocellulose membranes, and immunodetection of antigen with antibodies were performed as described previously (Gaude *et al.* 1991, 1993). Polyclonal rabbit serum specific for class I SLG proteins and monoclonal antibody MAb 85-36-71, which recognizes class II SLG proteins, have been described previously (Gaude *et al.* 1991, 1993; Giranton *et al.* 1995).

DNA blot analysis: DNA blot analysis was carried out as described (Cabrillac *et al.* 1999). Probes corresponding to intron sequences of *SLGA* and *SLGB* (Cabrillac *et al.* 1999) and to *SLA* (Pastuglia *et al.* 1997b) have been described. The *SRK* probe was derived from the third intron of SRK_2 (J. M. Cock, unpublished data) by PCR amplification with SK42 (5′-GTATAAATAATGAAGGAATCACTATGAAAT-3′) and SK40 (5′-GATCACTTATACAAAACCAACAGAGCAG-3′).

PCR and PCR-restriction fragment length polymorphism segregation analysis: *SLGA* and *SLGB* alleles were amplified from genomic DNA, using the PCR and oligonucleotides PS3 (see below) and SG17 (5′-TGTTCCGTCTGTCAAGTCCCA CTGCTGCGG-3′) for the progeny of plant i/j and PS4CM (5′-CGGAATATGGTATAAGAAAGTCTCCCA-3′) and SG39 (5′-CTTTGCGTTTCAACACGTTGATTCA-3′) for the progeny of plant k/l. *SRK* alleles were similarily amplified using the oligonucleotides PS5CM (5′-GGAATATGGTATAAAAAAGCCCC CTG-3′) and SG2 (see below). PCR amplification conditions were 94° for 3 min followed by 30 cycles of denaturation at 94° for 40 sec, annealing at 51° for 40 sec, and extension at 72° for 1 min in a GeneAmp PCR system 9600 cycler (Perkin-Elmer, Shelton, CT). PCR products were digested with either

TABLE 1

Origins of the seven S_2 plants analyzed in this study

Plant	Cultivar	Description	Origin
c	*B. oleracea* var. *albogabra*	Rapid cycling (DJ2024)	HRI, Wellesbourne, United Kingdom
d	*B. oleracea* var. *albogabra*	Rapid cycling	T. Hodgkin collection, United Kingdom
e	*B. oleracea* var. *italica*	Broccoli ("Ramoso Calabrese Tardivo," Scaravatti, Italy)	INRA, Rennes-Le Rheu, France
f	*B. oleracea* var. *acephala*	Kale (CFO 56 34)	INRA, Rennes-Le Rheu, France
g/h	*B. oleracea* var. *gemmifera* × *B. oleracea* var. *acephala*	Brussels sprouts × curly kale (DJ6033)	HRI, Wellesbourne, United Kingdom
i/j	*B. oleracea* var. *acephala*	Kale (D. Ockendon collection)	HRI Genetic Resources Unit, Wellesbourne, United Kingdom
k/l	*B. oleracea* var. *gemmifera*	Brussels sprouts (DJ8134)	HRI, Wellesbourne, United Kingdom

*Taq*I or *Nde*I restriction enzymes (GIBCO BRL, Gaithersburg, MD) to detect a restriction fragment length polymorphism (RFLP).

cDNA cloning and DNA sequencing: Total RNA (1 μg) from stigmas of each of the seven S_2 plants was reverse transcribed with Superscript II reverse transcriptase (GIBCO BRL), using oligonucleotide RA1 (Frohman *et al.* 1988) as a primer. *SRK* and *SLG* sequences were PCR amplified from the resulting cDNA or from genomic DNA extracted from the same plants, using different combinations of the following oligonucleotides:

5′ oligonucleotides:
PS3, 3′-ATGAAAGGGGTACAGAACAT-5′ (Nishio *et al.* 1996)
SDOM1, 3′-GWTGGTAYCTCGGRATRTGGTA-5′
3′ oligonucleotides:
SK30, 3′-TTCTCGCCCTCATAAACACAACAG-5′
SK38, 3′-CTCCAACCTATGATTTTTCCAGT-5′
SK66DC, 3′-CTCCTCCAAAAGCAGAACACGATAACACTC-5′
SG2, 3′-GGCCTGCAGCAGCATTCAATCTGAC-5′

PCR products were amplified under the same conditions described above and were cloned into pGEM-T easy (Promega, Madison, WI) and sequenced by the dideoxynucleotide chain termination method of Sanger *et al.* (1977) on an automatic sequencer (Applied Biosystems, Foster City, CA). Sequence data were analyzed with Lasergene sequence analysis software (DNAstar, London).

Sequence analysis: Multiple sequence alignments were constructed using the Megalign program (Lasergene; DNAstar). Neighbor-joining trees were constructed from these alignments using ClustalW software on the Pasteur Institute website (http://bioweb.pasteur.fr/#log; Thompson *et al.* 1994). Nonsynonymous substitutions were identified using DnaSP, version 3 (Rozas and Rozas 1999).

The nucleotide sequence data presented in this article have been submitted to the EMBL, GenBank, and DDBJ nucleotide sequence databases under accession nos. AJ306573–AJ306591.

RESULTS

Identification of Brassica plants carrying S_2 haplotypes: Diallel crosses were carried out to test the self-incompatibility phenotypes of seven putative S_2 *B. oleracea* plants from diverse origins (including kale, broccoli, and Brussels sprout lines; Table 1). Table 2 shows that a strong incompatible reaction was observed in the majority of cross-pollinations between the seven plants. In several crosses, however, either the SI reaction was weak or the cross was fully compatible. This was often the case, for example, when plant f was used as the female partner or when plant g/h was used as the male partner. This type of variability in the strength of the SI response has been observed previously, particularly with class II haplotypes. For example, Ruffio-Châble *et al.* (1997) described similar levels of variability in the strength of the SI response between different Brassica lines homozygous for the S_{15} haplotype. This phenomenon is thought to be due to the action of suppresser and modifier genes

TABLE 2

Self- and cross-incompatibility phenotypes of the seven S_2 plants in diallel crosses

	♂							
♀	c	d	e	f	g/h	i/j	k/l	S_8
c	1.5[a]	0.0	0.3	0.0	3.3	0.0	0.7	8.0
d	0.0	0.0	0.0	0.3	1.7	0.0	4.0	7.2
e	3.3	0.0	0.0	0.3	3.5	1.7	0.0	7.5
f	5.7	3.3	6.7	6.0	6.0	0.0	3.5	7.0
g/h	0.0	0.0	0.0	0.3	0.0	0.3	0.0	8.0
i/j	1.7	1.0	0.3	2.0	2.3	0.3	1.0	7.7
k/l	2.3	0.0	0.0	0.3	1.7	1.7	0.0	7.2
S_6	6.7	6.3	1.7	6.3	7.3	7.3	7.0	7.0

The male and female partner in each cross was as indicated. A homozygous S_8 plant and a homozygous S_6 plant were used as controls to verify pistil and pollen fertility, respectively.

[a] Mean pollen tube scores were determined by aniline blue staining of pistils 24 hr after pollination. Three pistils were scored for each cross. Scores were attributed as follows: 0, 0 tubes; 1, 1 or 2 tubes; 2, 3–5 tubes; 3, 6–9 tubes; 4, 10–14 tubes; 5, 15–25 tubes; 6, 26–50 tubes; 7, 51–100 tubes; 8, >100 tubes. A mean score was determined from three pistils tested for each cross.

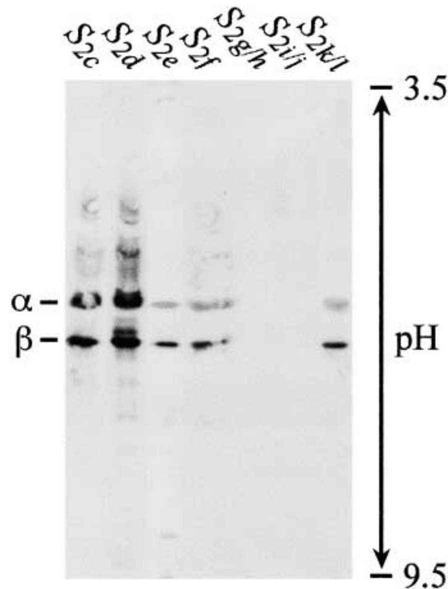


FIGURE 1.—Immunodetection of *S*-locus-encoded glycoproteins in stigma extracts from the seven S_2 plants. Soluble stigma proteins were separated by isoelectric focusing gel electrophoresis. Immunodetection was performed with monoclonal antibody mAb 85-36-7. The α- and β-polypeptides detected by the antibody are indicated.

that are unlinked to the *S* locus but that modulate the SI response (HINATA *et al.* 1994). In general, however, each of the seven lines exhibited a strong SI response with several of the other members of the group, indicating that they all carried a common *S* specificity. All of the plants tested, with the exception of plant f, were self-incompatible.

Immunodetection of SLG in stigma extracts of the S_2 plants: Immunological detection of SLG in stigma extracts separated by isoelectric focusing is a highly discriminative method for typing *S* haplotypes (RUFFIO-CHÂBLE *et al.* 1997, 1999; OKAZAKI *et al.* 1999). Figure 1 shows that two bands, α and β, characteristic of the S_2 haplotype (GAUDE *et al.* 1995), were specifically recognized by the monoclonal antibody (mAb) 85-36-71 in all of the S_2 plants except g/h and i/j. No additional *S*-locus-specific proteins were detected with either mAb 85-36-71 or with a polyclonal antibody that specifically recognizes class I SLG proteins (DELORME *et al.* 1995; GAUDE *et al.* 1995) in any of the seven S_2 plants.

Variation in the complement of *SLG* genes in the various S_2 plants: The immunodetection experiments indicated heterogeneity among the S_2 plants. To investigate this further, we probed genomic DNA blots with probes that specifically detect four *S* locus genes, *SRK, SLGA, SLGB,* and *SLA* (Figure 2). Plants g/h and i/j showed markedly different patterns to the five other S_2 plants. Two bands (2.00 and 1.25 kbp) hybridized to the *SRK* probe in DNA of both of these two plants and neither of these bands corresponded to that detected in the other five S_2 plants with the same probe (a single fragment of 1.27 kbp). A 0.85-kbp fragment that hybridized to the *SLGB* probe was detected in DNA of plants g/h and i/j but no hybridization was detected to the *SLGA* probe. Conversely, a 1.10-kbp *SLGA* fragment was detected in DNA of the five other S_2 plants but, in most cases, no *SLGB* sequence was detected. The only exception was DNA from plant k/l, which contained both the 1.10-kbp *SLGA* and the 0.85-kbp *SLGB* bands, indicating the presence of both genes. The *SLA* probe hybridized to the same samples as the *SLGA* probe, as expected (*SLA* has been shown to be very closely linked to *SLGA*; BOYES and NASRALLAH 1995; PASTUGLIA *et al.* 1997b). These results were confirmed by probing genomic DNA digested with *Bam*HI (data not shown).

The presence of two bands that hybridized to the *SRK* probe in DNA from plants g/h and i/j indicated that the plants analyzed might have been heterozygous at the *S* locus. To investigate this further we analyzed 18 progeny of the plant i/j. Figure 3a (top) shows that the two sequences detected by the *SRK* probe segregated with a 4:10:4 ratio (homozygous i:heterozygous i/j:homozygous j) that closely approximates a Mendelian 1:2:1 ratio. When *Bam*HI-digested DNA was probed with the *SLGB* probe, the pattern obtained was consistent with the presence of two polymorphic allelic sequences. However, an additional, nonpolymorphic fragment was also detected in this experiment (Figure 3a, middle). An alternative approach, based on restriction digestion of PCR products, was therefore used to specifically amplify the *SLGB* alleles and to detect a polymorphism that was predicted from the sequences of the two *SLGB* alleles (Figure 3a, bottom). These analyses demonstrated that plant i/j was heterozygous at both the *SRK* and *SLGB* loci and allowed the identification of the alleles associated with each haplotype.

DNA gel blots indicated that plant k/l carried alleles of *SRK, SLGA,* and *SLGB,* suggesting that it may have been heterozygous at the *S* locus (Figure 2). To investigate this further, we screened genomic DNA from 10 progeny of plant k/l for the presence of *SLGA* and *SLGB,* using PCR and gene-specific oligonucleotide pairs. *SLGA* and *SLGB* segregated independently in this population, confirming that plant k was heterozygous (Figure 3b).

The conclusions that can be drawn from the data presented in Figures 2 and 3 are summarized in Figure 3c. Figure 3c also includes the two S_2 haplotypes characterized by TANTIKANJANA *et al.* (1993) and KUSABA *et al.* (2000), which we have designated as S_{2a} and S_{2b}, respectively. The figure shows that two classes of S_2 haplotype could be clearly distinguished. In the first, *SRK* is associated with *SLGA* (haplotypes a, c, d, e, f, and k) and in the second SRK is associated with *SLGB* (haplotypes b, g, h, i, j, and l). None of the plants analyzed carried an S_2 haplotype consisting of *SRK, SLGA,* and *SLGB* as has been described in the S_{15} haplotype.

Cloning and sequence analysis of the *S* locus genes

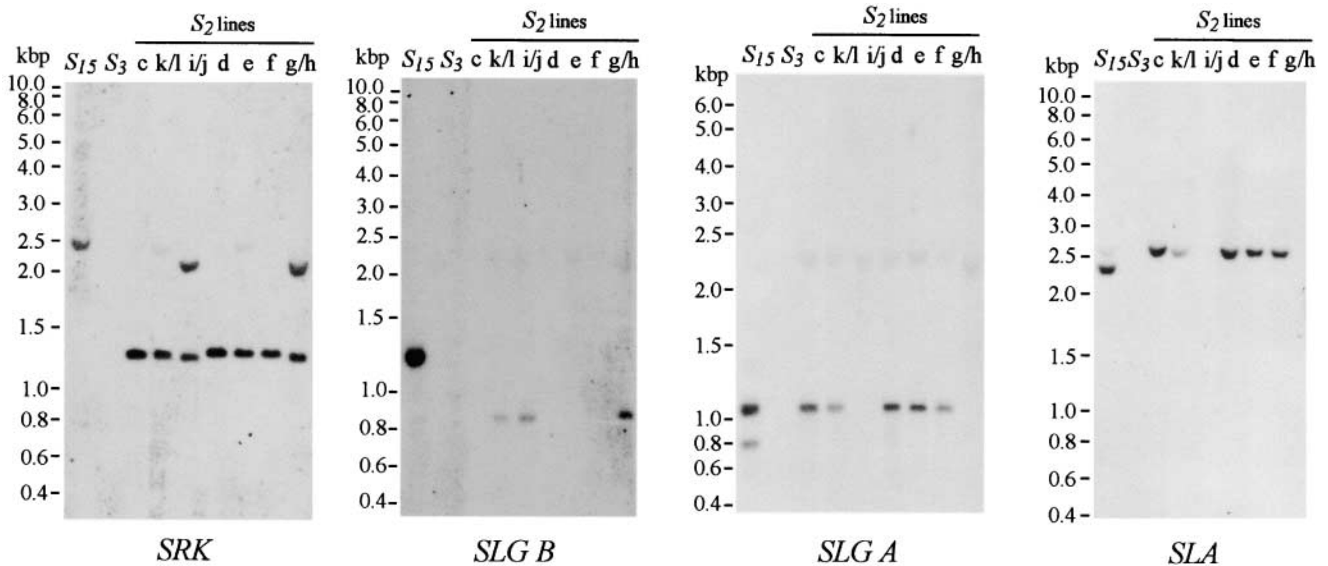


Figure 2.—Detection of *SRK*, *SLGA*, *SLGB*, and *SLA* alleles in *B. oleracea* plants carrying S_2, S_3, and S_{15} haplotypes. A gel blot of genomic DNA digested with *Ear*I was sequentially hybridized with the probes indicated beneath each panel. The positions of molecular length markers are shown at left in kilobases (kbp).

present in the *B. oleracea* S_2 plants: *SRK*, *SLGA*, and *SLGB* sequences were amplified from the seven S_2 plants by reverse transcription (RT)-PCR or by PCR amplification of genomic sequences, using different combinations of six oligonucleotides designed to amplify class II *S* locus genes. We isolated clones corresponding to all of the *SRK*, *SLGA*, and *SLGB* alleles detected by DNA gel blotting except SRK_{2g} and SRK_{2h}. However, note that, based on comparison of *SLGB* sequences (see below), these two *SRK* alleles are likely to be highly similar to SRK_{2i} and SRK_{2j}. The cloned fragments shared a region corresponding to almost the entire S domain (represented by a shaded bar in Figure 5a). To demonstrate that the sequences that we designated *SRK* were not derived from *SLGA* or *SLGB* genes, we amplified, by RT-PCR, a fragment that included both S domain and kinase domain sequences (data not shown).

To estimate the level of sequence polymorphism at each *S* locus gene locus we compared, at both the nucleotide and deduced amino acid sequence levels, the S domains of the *SRK*, *SLGA*, and *SLGB* alleles from the seven S_2 plants and from the two S_2 haplotypes that have been described previously (Tantikanjana *et al.* 1993; Kusaba *et al.* 2000). Figure 4a shows that polymorphisms were detected in *SRK*, *SLGA*, and *SLGB* although *SLGA* was significantly less polymorphic than *SLGB* and *SRK*. Despite this polymorphism, alleles of the three genes could be clearly distinguished at the sequence level. $SLGA_2$ was less closely related to SRK_2 and $SLGB_2$ than these latter genes were to each other.

Figure 4b shows a phylogenetic tree that was constructed from an alignment of the deduced amino acid sequences of the S domains of the S_2 sequences with those of other class II *S* locus genes. The allelic forms of SRK_2 clustered in one part of the tree together with the two previously reported sequences, SRK_{2a} and SRK_{2b}. SRK sequences corresponding to other class II alleles fell outside this cluster. Taken together with the data from genetic crosses (Table 2), this confirmed that the seven plants analyzed in this study carried the S_2 haplotype.

Surprisingly, two phylogenetically distinct groups of SRK_2 sequence could be distinguished (Figure 4b). These two groups correlated exactly with the presence of either *SLGA* or *SLGB* in each haplotype (compare Figure 4b with Figure 3c).

Sequence analysis of the S_2 plants also provided a possible explanation as to why no SLG proteins were detected in stigma extracts of plants g/h and i/j (Figure 1). The epitope recognized by the mAb 85-36-71 (IYVNTLSSSE) is conserved at the predicted N terminus of the mature SLGA protein, but a variant sequence (TYVNTMSSSE) is predicted to occur in the SLGB protein encoded by the i, j, g, and h haplotypes, and no *SLGA* gene was detected in these plants. From a practical point of view, heterogeneity concerning the presence or absence of *SLG* genes in haplotypes that confer the same SI phenotype could present problems for *S* haplotype typing methods that are based on detection of *SLG* at either the DNA or protein level (Nishio *et al.* 1996; Ruffio-Châble *et al.* 1997; Sakamoto *et al.* 1998; Okazaki *et al.* 1999; Ruffio-Châble *et al.* 1999). In this respect, methods based on determining the *SRK* genotype are likely to be more reliable (Nishio *et al.* 1997).

Sequence polymorphism in the HV domains of *SRK*: Recognition of self-pollen in self-incompatible Brassica is thought to involve haplotype-specific interaction between SCR on the male side and SRK on the female side.

It has, however, been difficult to identify the regions of these two proteins that determine haplotype specificity because of their extremely polymorphic nature. For example, although hypervariable regions have been identified in the extracellular domain of SRK [HV1, HV2, HV3, and a deletable region (DR); Nasrallah *et al.* 1987; Kusaba *et al.* 1997], it is not known whether these domains determine haplotype specificity or whether they are merely regions that are under a low level of selective constraint and therefore are free to diverge to a greater degree. One way of addressing this question is to carry out comparisons both between alleles of the same *S* haplotype (where domains determining *S* haplotype specificity should be conserved) and between alleles from different *S* haplotypes (where domains determining *S* haplotype specificity should be polymorphic).

When the 10 SRK_2 alleles were compared and polymorphic residues were positioned with respect to the S domain, we noted that none of the substitutions affected residues of the DR, HV1, or HV2 (Figure 5a, top). In contrast, when SRK sequences from different class II haplotypes were compared, polymorphic residues appeared to be clustered in these regions (Figure 5a, bottom). To further investigate this observation, we calculated the percentage of substitutions observed in each region on a per residue basis so that substitution rates could be compared despite the differences in length of each region (Figure 5b). This analysis suggested that the pattern of substitutions observed when the different SRK_2 sequences were compared differed from that observed when SRK sequences were compared between haplotypes. On average, when SRK sequences of the three class II haplotypes were compared, a substitution was more than three times more likely to be found at a residue within the DR, HV1, or HV2 than at a residue elsewhere in the S domain. In contrast, when the SRK_2 sequences were compared, no substitutions were observed in these regions. We used a modified version of the Hudson-Kreitman-Aguadé (HKA) test (Hudson *et al.* 1987) to determine whether there was a significant difference in the distribution of nonsynonymous substitutions between the DR/HV1/HV2 and the rest of the region analyzed in a comparison of SRK_2 alleles *vs.* a comparison between the SRK_2 alleles and SRK_{15} (*i.e.*, comparing within-haplotype and between-haplotype comparisons). We did not include HV3 with the other HV regions in this analysis because it is divergent between some SRK_2 alleles (Figure 5b) and yet conserved in several *SRK* alleles corresponding to different *S* haplotypes (see below), suggesting that it has not evolved in the

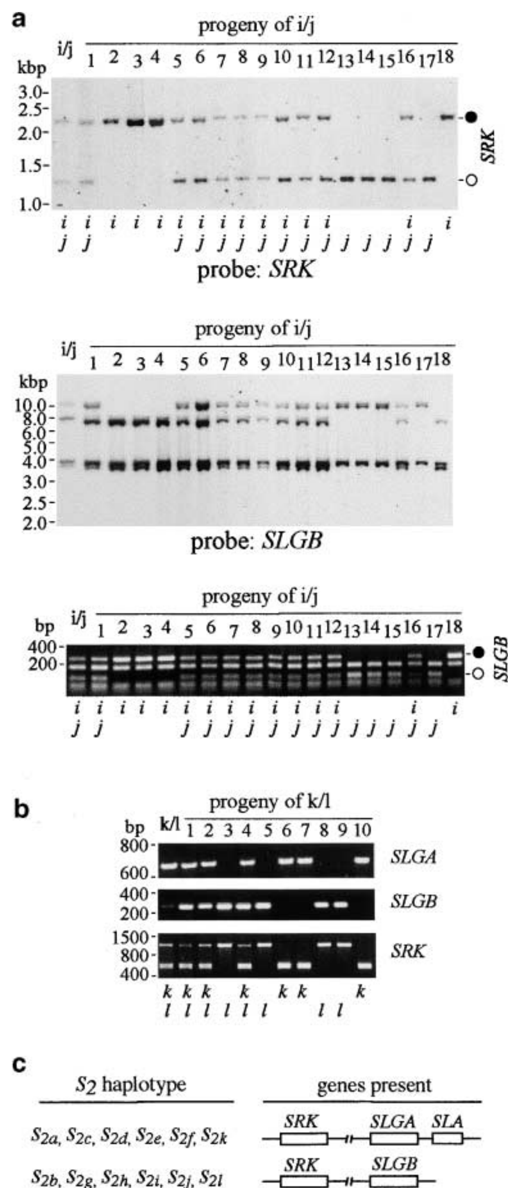
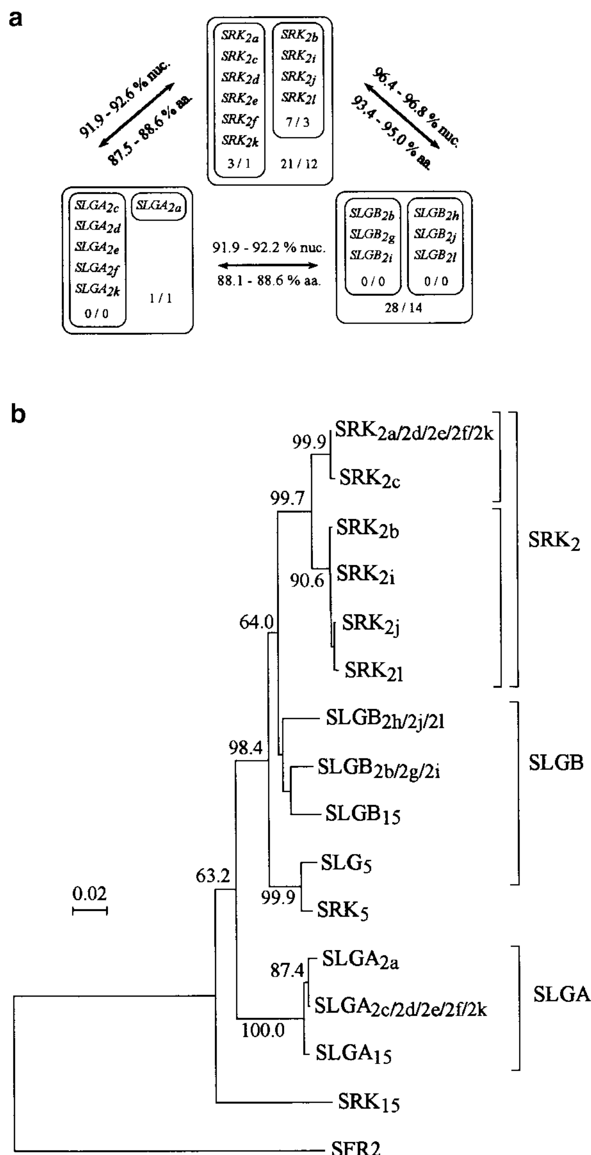


Figure 3.—Linkage analysis of the *SLG* and *SRK* alleles in plants i/j and k/l. (a) Segregation of *SRK* and *SLGB* alleles in self-progeny of plant i/j. DNA gel blot analysis was performed on 5 μg/lane of *Ear*I- (top) or *Bam*HI-digested (middle) genomic DNA of plant i/j and 18 self-progeny using *SRK* (top) and *SLGB* (middle) probes. The bottom shows a PCR-RFLP analysis of the segregation of the two *SLGB* alleles. *SLGB* sequences were specifically amplified with oligonucleotides PS3 and SG17 and digested with *Taq*I to reveal a restriction fragment length polymorphism predicted from sequence analysis of the two *SLGB* alleles. Bands corresponding to alleles of either *SRK* or *SLGB* from the S_{2i} and S_{2j} haplotypes are indicated by either a solid circle or an open circle, respectively. The S_2 haplotype(s) (S_{2i} or S_{2j}) carried by each of the progeny is (are) indicated below each lane. The positions of molecular length markers are shown at left in kilobases (kbp) or base pairs (bp). (b) Analysis of the segregation of *SLGA*, *SLGB*, and *SRK* in self-progeny of plant k/l. *SLGA*, *SLGB*, and *SRK* sequences were specifically amplified with oligonucleotides PS3 and SG38, PS4CM and SG39, or PS5CM and SG2, respectively. The S_2 haplotype(s) (S_{2k} and/or S_{2l}) carried by each of the progeny is (are) indicated below each lane. The positions of molecular length markers are shown at left in base pairs (bp). (c) Schematic representation showing the complement of genes present in S_{2a}, in S_{2b}, and in the 10 S_2 haplotypes carried by the seven plants analyzed in this study (including three plants that were heterozygous at the *S* locus).

same manner as the other HV domains. Table 3 shows that the HKA test gave a result that was very close to being significant ($P = 0.07$). This is interesting because the HKA test provides a conservative estimate of significance, particularly in this case, where the number of observed nonsynonymous substitutions is small.

Although the total number of polymorphic sites identified when $SLGB_2$ sequences were compared was similar to that observed for the SRK_2 sequences, a much higher proportion was located in the DR, HV1, and HV2 (7 out of 14 substitutions, Figure 5a). A Fisher's exact test showed that this difference in the distribution of polymorphisms in SLGB and SRK was highly significant ($P = 0.006$; Table 4), indicating that these domains are not under the same functional constraints in SLGB as in SRK. These data are consistent with several reports that indicate that SLG does not play a role in haplotype-specific recognition of self-pollen (Gaude *et al.* 1995; Cabrillac *et al.* 1999; Nishio and Kusaba 2000).



The difference in the distribution of polymorphisms in SRK and SLGB also supports the hypothesis that the DR, HV1, and HV2 domains have been conserved in the SRK_2 alleles. To investigate this difference further, we carried out a second HKA test, again to compare distributions of nonsynonymous substitutions between the DR/HV1/HV2 and the rest of the region analyzed, but this time comparing the "within SRK_2" data to a comparison between SRK_2 and $SLGB_2$. Table 5 shows that a significant difference was detected in this comparison ($P = 0.03$), suggesting that there has been either selection against changes in the DR, HV1, and HV2 of SRK_2 or that there has been positive selection for changes in these regions since the divergence of *SRK* and *SLGB*. It will, however, be important to carry out similar analyses on other *S* haplotypes to ensure that the observed difference is not due to gene conversion having, by chance, preferentially homogenized the non-HV regions in the S_2 haplotype.

To date, intrahaplotype polymorphism has been described in only one other Brassica haplotype, S_{13} (Kusaba *et al.* 2000). Surprisingly, and in contrast to our findings with the S_2 haplotype, when the S domains of SRK_{13} and $SRK_{13\text{-b}}$ were compared, one of the three amino acid substitutions was located in HV2. However, note that, if HV1 and HV2 are involved in determining *S* haplotype specificity, it is possible, and even likely, that only a subset of residues within the HV domains is involved in this function. Six of the seven differences between the predicted mature SLG_{13} and $SLG_{13\text{-b}}$ proteins were found

Figure 4.—Sequence comparison and phylogenetic relationships of the S domains of class II *SRK*, *SLGA*, and *SLGB* alleles. (a) Similarity between alleles of *SRK*, *SLGA*, and *SLGB* from 12 S_2 haplotypes. Alleles corresponding to each gene are grouped separately and the range of similarity between genes at the nucleotide (nuc.) and amino acid (aa.) levels is indicated. For each gene, alleles were subgrouped on the basis of similarity and the numbers of synonymous/nonsynonymous substitutions observed within each group and subgroup are indicated. All comparisons were carried out using a region of 1299 bp corresponding to the majority of the S domain (see Figure 5a where this region is indicated by a shaded bar). (b) Neighbor-joining tree based on class II SLG and SRK S domain sequences. SRK_{2a} and SLG_{2a} (Tantikanjana *et al.* 1993), SRK_{2b} and SLG_{2b} (Kusaba *et al.* 2000), SLG_5 (Scutt and Croy 1992), SRK_5, SRK_{15}, $SLGA_{15}$, and $SLGB_{15}$ (Cabrillac *et al.* 1999) have been described previously. SFR2 (Pastuglia *et al.* 1997a) was included as an outgroup. Based on sequence similarity and gene structure, SLG_5 is an allele of *SLGB*. The tree was constructed using the PAM250 amino acid substitution matrix. The numbers next to the branches are bootstrap values expressed as percentage confidence level and based on 2000 repeats. The scale indicates amino acid substitutions/site. Note that this tree is presented as a means of representing the relationships between the sequences shown and probably does not correctly describe the evolutionary events that generated these sequences. The phylogeny of these sequences has almost certainly been obscured in many cases by events that tend to result in concerted evolution within haplotypes.

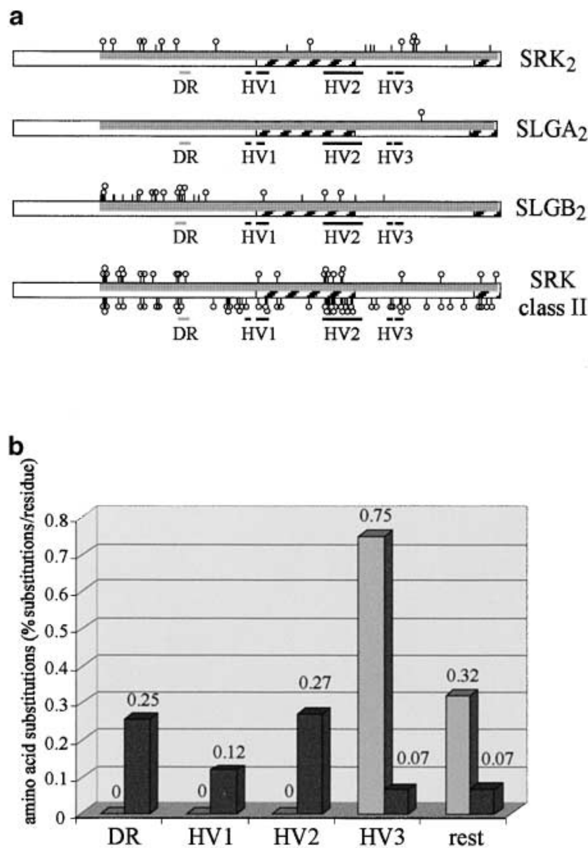


FIGURE 5.—Distribution of nucleotide and amino acid polymorphisms in *SRK*, *SLGA*, and *SLGB*. (a) Schematic representations showing the positions of observed nucleotide and amino acid substitutions along the S domains of *SRK*, *SLGA*, and *SLGB* following intra- and interhaplotype comparisons. Comparisons were carried out for the 10 SRK_2 alleles (SRK_2), the 6 $SLGA_2$ alleles ($SLGA_2$), the 6 $SLGB_2$ alleles ($SLGB_2$), and between SRK_2 and either SRK_5 or SRK_{15} (SRK class II). Vertical bars and vertical bars with open circles indicate the positions of nucleotide and amino acid residues that exhibited polymorphism. In the bottom only amino acid substitutions are shown, either between SRK_2 and SRK_5 (above the rectangle) or between SRK_2 and SRK_{15} (below the rectangle). Open rectangles represent signal peptides plus S domains; the region that was compared is indicated by a shaded bar. The positions of the DR (deletable region) and the three hypervariable regions (HV1, HV2, and HV3) are indicated beneath the S domains by solid bars. The hatched portions of the S domain represent the two variable regions defined by NASRALLAH *et al.* (1987). (b) Frequency of observed amino acid substitutions in different regions of the SRK S domain in either inter- or intrahaplotype comparisons. The mean frequencies of observed amino acid substitutions, expressed as percentage substitutions per amino acid residue, are indicated by lightly shaded and darkly shaded bars for comparisons within the 10 SRK_2 sequences and for comparisons between different class II SRK sequences (SRK_2/SRK_5, SRK_2/SRK_{15}, and SRK_2/SRK_{29}), respectively. SRK_2, SRK_5, and SRK_{15} are from *B. oleracea*, and SRK_{29} is from *B. rapa*. DR, deletable region; HV1, HV2, and HV3, hypervariable regions; rest, rest of the region indicated by the shaded bar in Figure 5a, excluding the DR, HV1, HV2, and HV3 regions.

TABLE 3

HKA test for difference in distribution of nonsynonymous substitutions between DR/HV1/HV2 and the rest of the S domain in a comparison between *SRK_2* alleles *vs.* a comparison between the *SRK_2* alleles and *SRK_{15}*

	Rest of the S domain[a]	DR, HV1, HV2
Within-haplotype polymorphism data		
Segregating sites (obs)	12	0
Segregating sites (exp)	8.65	3.35
Total no. of sites	1112	183
Between-haplotype divergence		
No. differences (obs)	37	19
No. differences (exp)	40.35	15.95
Total no. of sites	1113	187
	$\chi^2 = 3.25$	$P = 0.07$

Sample size was eight. obs, observed; exp, expected.

[a] The regions analyzed corresponded to the regions indicated by the shaded bars in Figure 5a.

within HV1, supporting our observation that substitutions are clustered in the HV domains in SLG proteins from the same haplotype.

To further investigate the role of the HV domains in SRK function, we compared the HV domains of all the SRK sequences available in the databases (Figure 6). The sequence of the HV3 region was not conserved in the SRK_2 sequences and, moreover, the HV3 domains of four of the SRK_2 sequences were identical to those of SRK_5 and SRK_{45} from *B. oleracea* and *B. campestris*, respectively. In contrast, HV1 and HV2 were both 100% conserved in the 10 SRK_2 sequences but no two SRK sequences from different haplotypes exhibited the same sequence in these regions. Taken together with the analyses described above, these data argue against DR, HV1, and HV2 merely being regions of the SRK protein that are under a low level of functional constraint. Rather, they indicate that substitutions in these regions are associated with the acquisition of new SI specificities.

The DR was conserved among the SRK_2 sequences but two pairs of class I SRKs shared identical DR domains ($BrSRK_8/BoSRK_{23}$ and $BoSRK_3/BnSRK_{A10}$, where Bo, Br, and Bn refer to the species of origin, *B. oleracea*, *B. rapa*, and *B. napus,* respectively; Figure 6). Therefore, this domain does not determine *S* haplotype specificity alone although it is possible that it may contribute to specificity by acting in combination with another polymorphic region of the protein such as HV1 or HV2.

DISCUSSION

***SLG* and *S* locus structure:** In a previous study, we reported the existence of two different *SLG* genes (*SLGA* and *SLGB*) in class II *S* haplotypes (CABRILLAC *et al.*

TABLE 4

Fisher's exact test to compare distribution of nonsynonymous substitutions between DR/HV1/HV2 and the rest of the S domain in SRK_2 and $SLGB_2$

	Rest of the S domain[a]	DR, HV1, HV2
$SLGB_2$	7	7
SRK_2	12	0
	$P = 0.006$	

[a] The regions analyzed corresponded to the regions indicated by the shaded bars in Figure 5a.

TABLE 5

HKA test for difference in distribution of nonsynonymous substitutions between DR/HV1/HV2 and the rest of the S domain in a comparison between SRK_2 alleles *vs.* a comparison between the SRK_2 alleles and $SLGB_2$

	Rest of the S domain[a]	DR, HV1, HV2
Within-haplotype polymorphism data		
Segregating sites (obs)	12	0
Segregating sites (exp)	8.77	3.23
Total no. of sites	1112	183
Between-gene divergence		
No. differences (obs)	7	7
No. differences (exp)	10.23	3.77
Total no. of sites	1106	189
	$\chi^2 = 4.715$	$P = 0.0299$

Sample size was eight.

[a] The regions analyzed corresponded to the regions indicated by the shaded bars in Figure 5a.

1999). Both *SLG* genes were present in an S_{15} haplotype but either only *SLGA* or only *SLGB* was detected in an S_5 and in an S_2 haplotype, respectively. More recently, Kusaba *et al.* (2000) described a novel form of the S_2 haplotype, S_{2b}, which exhibited both structural and sequence differences compared to the S_2 haplotype originally described by Tantikanjana *et al.* (1993). In this study, we analyzed the sequences of *SLG* and *SRK* genes from 10 additional S_2 haplotypes from diverse genetic backgrounds. On the basis of sequence similarities, the 10 S_2 haplotypes could be divided into two groups that resembled S_{2a} on one hand and S_{2b} on the other. In the first group, *SRK* was associated with *SLGA*, and *SLGB* was absent, while in the second group *SRK* was associated with *SLGB*, and *SLGA* was absent.

Taken together, these results indicate that *SRK*, *SLGA*, and *SLGB* were already associated at the *S* locus in an ancestor of the class II haplotypes. *SLGA* or *SLGB* would then have been lost from the *S* locus in some haplotypes, these events occurring, at least in part, after the divergence of the different class II *S* haplotype specificities. An alternative scenario has been proposed by Kusaba *et al.* (2000). They suggest that *SLGA* was originally acquired by the S_2 haplotype and then subsequently transmitted to the S_{15} haplotype. This latter model is based on the fact that *SLGB* is commonly found in class II haplotypes of both *B. oleracea* and *B. rapa* (Scutt and Croy 1992; Hatakeyama *et al.* 1998) whereas *SLGA* has been found only in the S_2 and S_{15} haplotypes so far. Note, however, that SRK_{15} is more similar to $SLGA_{15}$ than to $SLGB_{15}$ whereas the opposite is true for SRK_2, suggesting that, if *SLGA* was acquired after divergence of the class II haplotypes, then this probably occurred in the S_{15} haplotype. A more detailed analysis of independent isolates of additional class II haplotypes such as S_5 should help to distinguish between these different models.

The *S* locus region has been characterized in detail in several class I *S* haplotypes and there is currently no evidence for the presence of a second *SLG* gene in this class of haplotype (Boyes *et al.* 1997; Cui *et al.* 1999; Suzuki *et al.* 1999). It seems likely, therefore, that the acquisition of a second *SLG* gene by the class II haplotypes occurred after their divergence from class I haplotypes.

It is difficult to determine precisely how *SRK*, *SLGA*, and *SLGB* originated due to the highly polymorphic nature of these three genes. Comparison of their introns indicates a complex series of recombination and mutation events that tend to obscure their evolutionary history (Cabrillac *et al.* 1999). It seems probable, however, that *SLGA* arose as a result of a duplication of *SRK* (Tantikanjana *et al.* 1993). *SLGB* may have originated in a similar manner, although it is also possible that it was recruited to the *S* locus from another part of the genome. Analogous translocation events of this type have been observed between distant disease resistance gene clusters in tomato (Parniske and Jones 1999).

Comparison of the S_2 haplotypes indicated the existence of two distinct phylogenetic groups of SRK_2 alleles and these two groups correlated with the presence of either $SLGA_2$ or $SLGB_2$, respectively. The fact that *SRK* alleles from each group can be clearly distinguished at the sequence level suggests that recombination between these two types of S_2 haplotype has been suppressed or has occurred only rarely. An alternative possibility is that these two groups of S_2 alleles evolved independently in geographically isolated populations and have been brought together only recently. Note, however, that we found no correlation between the presence of *SLGA* or *SLGB* and the origin of the plants that were analyzed.

Hypervariable domains and haplotype specificity: In both sporophytic and gametophytic multiallelic SI systems, genes that encode female components of the SI system have been shown to possess regions of extreme sequence polymorphism, the HV domains. In the Solanaceae, S-RNAse HV domain swapping experiments

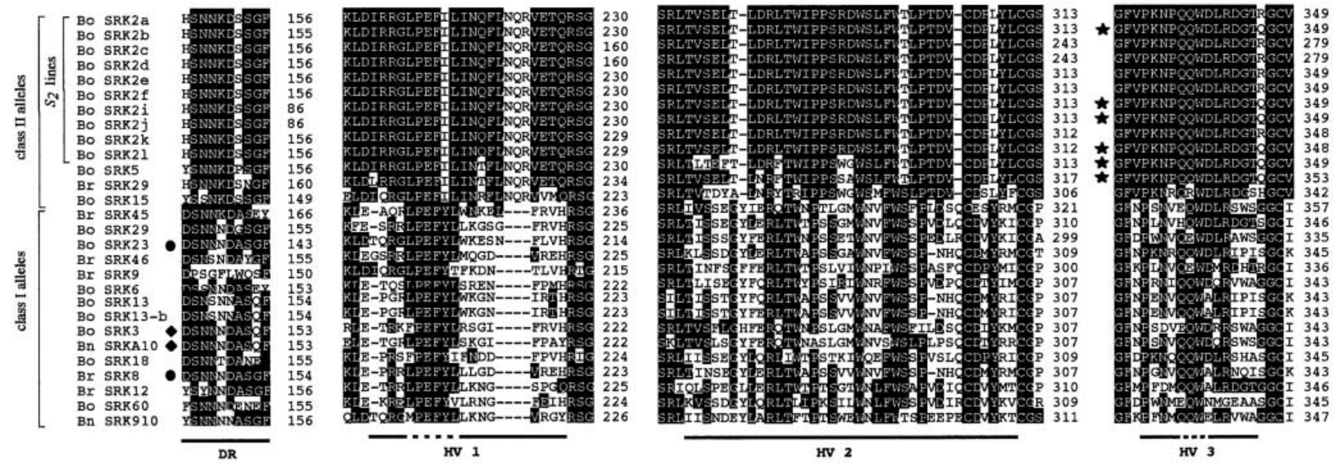


Figure 6.—Sequence comparison of the DR and HV regions from the SRK_2 alleles with other class II and class I *SRK* alleles. Solid bars indicate the positions of each region. Numbers indicate the position of the last residue of each block in relation to the N-terminal end of the predicted unprocessed polypeptide. Residues conserved in a majority of sequences are boxed. *, SRK sequences with identical HV3 regions; ● and ◆, two pairs of identical DR sequences; Bo, *B. oleracea*; Br, *B. rapa*; Bn, *B. napus*; DR, deletable region; HV1, HV2, and HV3, hypervariable regions.

have provided evidence that these domains are responsible for the difference in specificity between closely related alleles (Matton *et al.* 1997). However, other regions of these proteins also seem to be important because exchanges of HV domains between more distantly related alleles often result in loss of activity rather than acquisition of new allelic specificities (Kao and McCubbin 1996; Zurek *et al.* 1997).

By analogy to the SI systems of the Solanaceae, and bearing in mind the unusually high number of functional *S* haplotypes in Brassica, it seems likely that the highly polymorphic HV regions of SRK are also involved in determining *S* haplotype specificity. However, in the absence of a direct experimental demonstration of the role of these domains, it has been difficult to distinguish between this possibility and a second hypothesis in which the HV regions are under a low level of selective constraint and therefore free to diverge to a greater degree (Nasrallah 1997). In this study we have addressed this problem by comparing the level of polymorphism in these regions either within or between haplotypes. These comparisons show that HV1 and HV2 tend to be conserved in SRK molecules that confer the same SI specificity but are diverged in SRK molecules that confer different SI specificities, indicating a role for HV1 and HV2 in determining haplotype specificity.

In the Solanaceae, comparisons of the rates of nonsynonymous and synonymous substitution in the hypervariable region of S-RNases indicated that these regions have been subjected to positive selection (Clark and Kao 1991). We carried out a similar analysis of HV1 and HV2 in the class II *SRK* genes but did not find any direct evidence of positive selection acting in these regions (data not shown). Similar results were obtained by Awadalla and Charlesworth (1999) when they analyzed the HV regions of class I *SRK* genes. However, these authors noted that, while there was no strong evidence of positive selection acting on the HV regions (K_a/K_s ratios were close to 1), K_a/K_s ratios for these regions were significantly higher than for the rest of the protein. They suggested that evidence of positive selection may be obscured by the high level of substitution in these regions, particularly if adaptive substitutions can occur at only a small subset of the amino acid positions.

The role of *SLG* in *S* locus function and evolution: There is accumulating evidence that SLG is not involved in the haplotype-specific recognition of self-pollen (Gaude *et al.* 1995; Cabrillac *et al.* 1999; Okazaki *et al.* 1999; Nishio and Kusaba 2000). If this is the case, then it is necessary to explain why *SLG* and *SRK* alleles are often highly similar within the same *S* haplotype (Stein *et al.* 1991; Goring *et al.* 1993; Delorme *et al.* 1995; Kusaba *et al.* 1997). We have suggested that this concerted evolution of *SLG* and *SRK* is due to their close proximity in the genome, which tends to favor gene conversion events (Cabrillac *et al.* 1999). There is convincing evidence that gene conversion events occur between *SLG* and *SRK*, particularly within haplotypes (Goring *et al.* 1993; Watanabe *et al.* 1994; Suzuki *et al.* 1997; Cabrillac *et al.* 1999). The DR, HV1, and HV2 regions were highly conserved in the SRK_2 sequences analyzed in this study but this was not the case for $SLGB_2$, where there was a tendency for substitutions to accumulate in these regions. We suggest that the capacity of SLG to accumulate substitutions in its HV domains may be an important factor in the generation of new *S* haplotype specificities. Gene conversion events would provide a means of transferring novel HV domain sequences from *SLG* to *SRK* and this could accelerate the evolution of novel SI speci-

ficities on the female side. Two other factors may be important for the functionality of such a system. The first is the overall similarity between *SLG* and the S domain of *SRK*, particularly within haplotypes, which is presumably the result of (reciprocal) sequence homogenization. The second is the maintenance of SLG as a functional protein due to its role in pollen adhesion and/or SRK maturation. In this respect, it is interesting that SLG binds PCPA1, a component of the pollen coat that is related to SCR at the sequence level (Doughty *et al.* 1998).

We thank Fabienne Deguerry and Anne-Marie Thierry for technical assistance, and Rosemary MacClenaghan, Graham King, Angela Pinnegar, and Dave Astley from HRI Wellesbourne for providing Brassica lines. This work was funded by the Institut National de la Recherche Agronomique, the Centre National de la Recherche Scientifique, and the Ecole Normale Supérieure de Lyon.

LITERATURE CITED

Awadalla, P., and D. Charlesworth, 1999 Recombination and selection at *Brassica* self-incompatibility loci. Genetics **152:** 413–425.

Boyes, D. C., and J. B. Nasrallah, 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 J. B. Nasrallah, 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., M. E. Nasrallah, J. Vrebalov and J. B. Nasrallah, 1997 The self-incompatibility (*S*) haplotypes of *Brassica* contain highly divergent and rearranged sequences of ancient origin. Plant Cell **9:** 237–247.

Cabrillac, D., V. Delorme, J. Garin, V. Ruffio-Châble, J. L. Giranton *et al.*, 1999 The S_{15} self-incompatibility haplotype in *Brassica oleracea* includes three S gene family members expressed in stigmas. Plant Cell **11:** 971–986.

Clark, A. G., and T.-h. Kao, 1991 Excess nonsynonymous substitution of shared polymorphic sites among self-incompatibility alleles of Solanaceae. Proc. Natl. Acad. Sci. USA **88:** 9823–9827.

Cock, J. M., 2000 A receptor kinase and the self-incompatibility response in *Brassica*, pp. 270–298 in *Advances in Botanical Research, Thematic Vol. 32: Plant Protein Kinases*, edited by M. Kreis and J. C. Walker. Academic Press, London.

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

Delorme, V., J. L. Giranton, Y. Hatzfeld, A. Friry, P. Heizmann *et al.*, 1995 Characterization of the *S* locus genes, *SLG* and *SRK*, of the *Brassica* S_3 haplotype: identification of a membrane-localized protein encoded by the S locus receptor kinase gene. Plant J. **7:** 429–440.

Dixit, R., M. E. Nasrallah and J. B. Nasrallah, 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.

Doughty, J., S. Dixon, S. J. Hiscock, A. C. Willis, I. A. Parkin *et al.*, 1998 PCP-A1, a defensin-like *Brassica* pollen coat protein that binds the *S* locus glycoprotein, is the product of gametophytic gene expression. Plant Cell **10:** 1333–1347.

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

Ellis, J. G., G. J. Lawrence, J. E. Luck and P. N. Dodds, 1999 Identification of regions in alleles of the flax rust resistance gene L that determine differences in gene-for-gene specificity. Plant Cell **11:** 495–506.

Frohman, M. A., M. K. Dush and G. R. Martin, 1988 Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA **85:** 8998–9002.

Gaude, T., L. Denoroy and C. Dumas, 1991 Use of a fast protein electrophoretic purification procedure for N-terminal sequence analysis to identify *S*-locus related proteins in stigmas of *Brassica oleracea*. Electrophoresis **12:** 646–653.

Gaude, T., A. Friry, P. Heizmann, C. Mariac, M. Rougier *et al.*, 1993 Expression of a self-incompatibility gene in a self-compatible line of *Brassica oleracea*. Plant Cell **5:** 75–86.

Gaude, T., M. Rougier, P. Heizmann, D. J. Ockendon and C. Dumas, 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., M. J. Ariza, C. Dumas, J. M. Cock and T. Gaude, 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.

Goring, D. R., T. L. Glavin, U. Schafer and S. J. Rothstein, 1993 An S receptor kinase gene in self-compatible *Brassica napus* has a 1-bp deletion. Plant Cell **5:** 531–539.

Hatakeyama, K., T. Takasaki, M. Watanabe and K. Hinata, 1998 Molecular characterization of *S* locus genes, *SLG* and *SRK*, in a pollen-recessive self-incompatibility haplotype of *Brassica rapa* L. Genetics **149:** 1587–1597.

Hinata, K., A. Isogai and K. Isuzugawa, 1994 *Manipulation of Sporophytic Self-Incompatibility in Plant Breeding*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Hudson, R. R., M. Kreitman and M. Aguade, 1987 A test of neutral molecular evolution based on nucleotide data. Genetics **116:** 153–159.

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

Kusaba, M., T. Nishio, Y. Satta, K. Hinata and D. Ockendon, 1997 Striking sequence similarity in inter- and intra-specific comparisons of class I *SLG* alleles from *Brassica oleracea* and *Brassica campestris*: implications for the evolution and recognition mechanism. Proc. Natl. Acad. Sci. USA **94:** 7673–7678.

Kusaba, M., M. Matsushita, K. Okazaki, Y. Satta and T. Nishio, 2000 Sequence and structural diversity of the *S* locus genes from different lines with the same self-recognition specificities in *Brassica oleracea*. Genetics **154:** 413–420.

Luu, D.-T., P. Heizmann and C. Dumas, 1997 Pollen-stigma adhesion in kale is not dependant on the self(in)compatibility genotype. Plant Physiol. **115:** 1221–1230.

Luu, D. T., D. Marty-Mazars, M. Trick, C. Dumas and P. Heizmann, 1999 Pollen-stigma adhesion in *Brassica spp* involves SLG and SLR1 glycoproteins. Plant Cell **11:** 251–262.

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

Nasrallah, J. B., 1997 Evolution of the *Brassica* self-incompatibility locus: a look into *S*-locus gene polymorphisms. Proc. Natl. Acad. Sci. USA **94:** 7673–7678.

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

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

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

Nishio, T., M. Kusaba, M. Watanabe and K. Hinata, 1996 Registration of *S* alleles in *Brassica campestris* L by the restriction fragment sizes of *SLGs*. Theor. Appl. Genet. **92:** 388–394.

Nishio, T., M. Kusaba, K. Sakamota and D. J. Ockendon, 1997 Polymorphism of the kinase domain of the *S*-locus receptor kinase gene (*SRK*) in *Brassica oleracea* L. Theor. Appl. Genet. **95:** 335–342.

Okazaki, K., M. Kusaba, D. J. Ockendon and T. Nishio, 1999 Characterization of *S* tester lines in *Brassica oleracea*: polymorphism of restriction length of *SLG* homologues and isolelectric points of *S*-locus proteins. Theor. Appl. Genet. **98:** 1329–1334.

Parniske, M., and J. D. Jones, 1999 Recombination between diverged clusters of the tomato Cf-9 plant disease resistance gene family. Proc. Natl. Acad. Sci. USA **96:** 5850–5855.

Pastuglia, M., D. Roby, C. Dumas and J. M. Cock, 1997a Rapid induction by wounding and bacterial infection of an S gene family receptor-like kinase gene in *Brassica oleracea*. Plant Cell **9:** 49–60.

Pastuglia, M., V. Ruffio-Châble, V. Delorme, T. Gaude, C. Dumas *et al.*, 1997b A functional S locus anther gene is not required for the self-incompatibility response in *Brassica oleracea*. Plant Cell **9:** 2065–2076.

Rozas, J., and R. Rozas, 1999 DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics **15:** 174–175.

Ruffio-Châble, V., and T. Gaude, 2001 *S*-haplotype polymorphism in *Brassica oleracea*. Acta Hortic. **546:** 257–261.

Ruffio-Châble, V., Y. Hervé, C. Dumas and T. Gaude, 1997 Distribution of *S*-haplotypes and its relationship with self-incompatibility in *Brassica oleracea* L. in inbred lines of cauliflower (*B. oleracea* var '*botrytis*'). Theor. Appl. Genet. **94:** 338–346.

Ruffio-Châble, V., J. P. Le Saint and T. Gaude, 1999 Distribution of *S*-haplotypes and relationship with self-incompatibility in *Brassica oleracea*. 2. In varieties of broccoli and romanesco. Theor. Appl. Genet. **98:** 541–550.

Sakamoto, K., M. Kusaba and T. Nishio, 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.

Sanger, F., S. Nicklen and A. R. Coulson, 1977 DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:** 5463–5467.

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

Scutt, C. P., and R. R. Croy, 1992 An S_5 self-incompatibility allele-specific cDNA sequence from *Brassica oleracea* shows high homology to the *SLR2* gene. Mol. Gen. Genet. **232:** 240–246.

Shiba, H., K. Hinata, A. Suzuki and A. Isogai, 1995 Breakdown of the self-incompatibility in *Brassica* by the antisense RNA of the *SLG* gene. Proc. Jpn. Acad. **71:** 81–83.

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

Stein, J. C., R. Dixit, M. E. Nasrallah and J. B. Nasrallah, 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.

Stephenson, A. G., J. Doughty, S. Dixon, C. Elleman, S. Hiscock *et al.*, 1997 The male determinant of self-incompatibility in *Brassica oleracea* is located in the pollen coating. Plant J. **12:** 1351–1359.

Suzuki, G., M. Watanabe, A. Isogai and K. Hinata, 1997 Highly conserved 5′-flanking regions of two self-incompatibility genes, SLG^9 and SRK^9. Gene **191:** 123–126.

Suzuki, G., N. Kai, T. Hirose, K. Fukui, T. Nishio *et al.*, 1999 Genomic organization of the *S* locus: identification and characterization of genes in the *SLG/SRK* region of the S^9 haplotype of *Brassica campestris* (syn. rapa). Genetics **153:** 391–400.

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

Takayama, S., H. Shiba, M. Iwano, H. Shimosato, F. S. Che *et al.*, 2000 The pollen determinant of self-incompatibility in *Brassica campestris*. Proc. Natl. Acad. Sci. USA **97:** 1920–1925.

Tantikanjana, T., M. E. Nasrallah, J. C. Stein, C. H. Chen and J. B. Nasrallah, 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.

Thompson, J. D., D. G. Higgins and T. J. Gibson, 1994 CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:** 4673–4680.

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

Toriyama, K., M. K. Thorsness, J. B. Nasrallah and M. E. Nasrallah, 1991 A *Brassica 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.

Watanabe, M., T. Takasaki, K. Toriyama, S. Yamakawa, A. Isogai *et al.*, 1994 A high degree of homology exists between the protein encoded by *SLG* and the S receptor domain encoded by *SRK* in self-incompatible *Brassica campestris* L. Plant Cell Physiol. **35:** 1221–1229.

Zurek, D. M., B. Mou, B. Beecher and B. McClure, 1997 Exchanging sequence domains between *S*-RNases from *Nicotiana alata* disrupts pollen recognition. Plant J. **11:** 797–808.

Communicating editor: D. Charlesworth