Comparative analysis of the S-intergenic region in class-II S haplotypes of self-incompatible \textit{Brassica rapa} (syn. \textit{campestris})

Tomohiro Kakizaki\(^1\), Yoshinobu Takada\(^1,2\), Tomoaki Fujioka\(^1\), Go Suzuki\(^2\), Yoko Satta\(^3\), Hiroshi Shiba\(^4\), Akira Isogai\(^4\), Seiji Takayama\(^4\) and Masao Watanabe\(^5,6\)*

\(^1\)Faculty of Agriculture, Iwate University, Morioka 020-8550, Japan
\(^2\)Division of Natural Science, Osaka Kyoiku University, Kashiwara 582-8582, Japan
\(^3\)Graduate School of Biosystem Science, Graduate University for Advanced Studies, Hayama 240-0193, Japan
\(^4\)Graduate School of Biological Sciences, Nara Institute of Science and Technology, Ikoma 630-0101, Japan
\(^5\)Graduate School of Life Sciences, Tohoku University, Sendai 980-8577, Japan.
\(^6\)The 21st Century Center of Excellence Program, Iwate University, Morioka 020-8550, Japan

(Received 15 December 2005, accepted 26 December 2005)

In the \textit{Brassica} self-incompatibility (SI) system, a pollen determinant, \textit{SP11}, is involved in dominance/recessive relationships in pollen SI phenotypes. In order to gain some insights into the genomic structure around the \textit{SP11} and the mechanisms that give dominance/recessive relationships, we characterized the genomic region containing \textit{SP11} and \textit{SRK} genes in three pollen recessive class-II \textit{S} haplotypes. The direction of transcription of \textit{S} genes was completely conserved among class-II \textit{S} haplotypes. However, the region between \textit{SP11} and \textit{SRK} (\textit{S-intergenic region}) was highly polymorphic without short repetitive sequences. In addition, we found a sequence similarity between the short repetitive sequence and 5'-upstream region of \textit{SP11}. This sequence similarity was found to be potentially related to the expression of dominance relationships through the change of chromatin structure.

**Key words:** \textit{Brassica rapa} (syn. \textit{campestris}), genome structure, self-incompatibility, \textit{S} locus

Self-incompatibility systems prevent self-pollination and promote outbreeding. In \textit{Brassica}, this system is sporophytically controlled by a single \textit{S} locus with multiple alleles, and recognition of pollen is controlled by \textit{S} haplotypes (designated \(S^1, S^2, \ldots, S^n\)) (Bateman 1955). The molecules controlling the self-incompatibility have been identified as \textit{SP11} (\textit{S}-locus protein 11, also known as \textit{SCR}; Suzuki et al. 1999; Schopfer et al. 1999; Takayama et al. 2000) and \textit{SRK} (\textit{S receptor kinase}; Takasaki et al. 2000). An allele-specific interaction between \textit{SP11} and \textit{SRK} has been proposed to cause activation of signaling cascade in the stigma (Kachroo et al. 2001; Takayama et al. 2001). Because the \textit{S} locus appears to be a multigene complex, “\textit{S} allele” is referred to as “\textit{S} haplotype” (Nasrallah and Nasrallah 1993).

Based on the sequence diversity of the \textit{SP11} and \textit{SRK} (\textit{S} genes), \textit{S} haplotypes are classified into two classes, class-I and class-II. The amino acid sequence identities among class-II \textit{SP11}s are 63.2 to 94.6%, rather high compared with those of class-I \textit{SP11}s which range from 19.5 to 76.1% (Shiba et al. 2002). Interestingly, class-I \textit{S} haplotypes are dominant over those of class-II on the pollen (Nasrallah and Nasrallah 1993; Hatakeyama et al. 1998b). Thus, the pollen grain derived from \textit{S} heterozygotes having the class-I and class-II \textit{S} haplotypes shows \textit{S} phenotype of class-I \textit{S} haplotype. \textit{SP11} alleles frequently show dominance/recessive relationships. This relationship has been shown to be due to reduced mRNA abundance of the recessive \textit{SP11} allele in \textit{B. rapa} (Shiba et al. 2002; Kakizaki et al. 2003) and in \textit{Arabidopsis lyrata} (Kusaba et al. 2002). However, the mechanism of monoallelic expression of the \textit{SP11} gene has not been defined. It is important to compare genomic organiza-
tion of dominant S haplotypes and recessive S haplotypes precisely, because dominance relationships are determined by the S locus itself. The detailed genomic organization of the S locus has already been reported in five class-I S haplotypes in B. rapa (Suzuki et al. 1999; Kimura et al. 2002; Shiba et al. 2003) and in B. napus (Cui et al. 1999). In the class-I S haplotypes, two S genes (SP11 and SRK) are located close to each other on the S locus, and the physical distance between genes is within 15-kb. However, their direction of transcription of S genes is diverted among S haplotypes (Takayama et al. 2000). On the other hand, the report of the S locus organization of the class-II S haplotypes is limited to S60 haplotype of B. rapa (Fukai et al. 2003). The S29, S40 and S44 haplotypes have been classified into class-II S haplotypes, and linear dominance hierarchies have been observed by reciprocal test pollinations (S44 > S40 > S60 > S29) (Hatakeyama et al. 1998b; Kakizaki et al. 2003). Moreover, in accordance with the linear dominance hierarchies observed among class-II S haplotypes, SP11 alleles can be dominant or recessive, with corresponding changes in expression, dependent on their allelic partner (Kakizaki et al. 2003). However, little is known about the genomic organization of the S-intergenic region in class-II S haplotypes without S60 haplotype. To clarify the relation between dominance relationships and genomic organization around SP11, we investigated the S-intergenic region of three additional class-II S haplotypes.

At first, to characterize the S-intergenic region of three class-II S haplotypes S29, S40 and S44, we screened the genomic libraries that were composed of lambda phage vector with each SP11 cDNA probe. We obtained the clones #29-4, #40-20, and #44-4, containing the 5'-upstream region of the SP11 gene of S29, S40, and S44 haplotypes, respectively. The length of the DNA inserts in #29-4, #40-20, and #44-4 were estimated to be approximately 18-kb, 15-kb, and 14-kb, respectively, by long-PCR. From the shot-gun sequencing analysis of the three clones, the #29-4 clone contained the SP11-29 gene and the 5'-genomic region of the SRK29 gene. The #29-4 clone contained the 5'-upstream region of SRK29, whose nucleotide sequence was completely overlapped to the SRK29 genomic clone determined by Hatakeyama et al. (1998a). In the case of the #40-20 clone, it contained the SP11-40 gene and a part of the regions encoding the S-domain and transmembrane domain of the SRK40 gene. The #44-2 clone contained SP11-44 and a part of the region encoding the S-domain of SRK44 (Fig. 1).

The full-length cDNA for SRK40 and SRK44 was obtained by RT-PCR using the information of the 5'-UTR sequences on the lambda clones. The nucleotide sequence of SRK40 and SRK44 has been deposited in GenBank, EMBL, and DDBJ databases (accession nos. AB211197 and AB211198, respectively). It was confirmed that the nucleotide sequences of cDNA clones for SRK40 and SRK44 completely corresponded to those of each genomic clone, respectively. From the comparison of the genomic sequence and cDNA sequence, sites of intron/exon junction were estimated (Fig. 1).

The promoter sequence elements that are required for sporophytic expression have already been identified in the class-I S0 haplotype (Shiba et al. 2001). In situ hybridization analysis revealed that class-II SP11 was expressed predominantly in anther tapetum (Shiba et al. 2002). We compared precisely the nucleotide sequences of the promoter region of four class-II S haplotypes of SP11 from the S29, S40, S44 and S60 haplotypes of B. rapa (designated SP11-29, SP11-40, SP11-44 and SP11-60 respectively) (Fig. 2). These sequences are highly conserved (72.2 to 85.8% identity) in the SP11 promoter from -250 to -1 bp. However, there is little similarity in the region upstream beyond -250 bp. Therefore, it is considered that the region of about -250 bp is important for the expression of SP11. However, it is difficult to identify the sequence similarity between class-I and class-II promoter regions (data not shown).

As a next step, the sequence similarity of the S-intergenic region was surveyed by Harr plot analysis between

---

Fig. 1. Comparative map of the S-intergenic region among class-II S haplotypes. The genomic organization of the S-intergenic regions of the S29, S40, S44 and S60 haplotypes is represented. S60 haplotype was referred from Fukai et al. (2003; accession nos. AB097116). Arrows indicate the orientation of transcription. The restriction sites indicated are as follows: B, BamHI; E, EcoRI; S, SacI.
Comparative analysis of the S-intergenic region

Four S haplotypes. The representative results between S^{29} and S^{40} haplotypes are shown in Fig. 3. The plots shows that the short repetitive sequences were conserved within S haplotypes at 1-kb upstream of the SP11 gene (REP29, REP40, REP44 and REP60; Fig. 3). The unit of the 64-bp sequence was contained three times in this
This repetitive sequence had 63.8 to 91.0% similarity among four class-II S haplotypes. There was another direct repeat over 1 kb in S40 and S44 haplotypes, however, this repetitive sequence did not exist in the other two S haplotypes (S29 and S60; Fig. 1). The distance between S genes (SP11 to SRK) was 18.4-kb, 8.9-kb and 11.5-kb in S29, S40 and S44 haplotypes, respectively. In the case of the S60 haplotype, the distance has been determined as 6.6-kb (Fukai et al. 2003). From these data, the physical distance between S genes was relatively conserved among class-II S haplotypes. On the other hand, the distance between S genes was intensely diverted among class-I S haplotypes (Takayama et al. 2000). Except for the two S genes (SP11 and SRK) and the short repetitive sequences, we could not identify the sequence similarity among class-II S haplotypes. The sequence diversity throughout the SP11 and SRK region should contribute to the suppression of recombination within the S-intergenic region.

In the S-intergenic region of the S60 haplotype, anther-specific non-coding RNA, called SAN1 (S locus Anther-expressed Non-coding RNA like-1), was predominantly expressed in anther (Fukai et al. 2003). In several cases, the expression of the non-coding RNA from one allele correlates with the repression of the flanking genes on the same allele (Sleutels et al. 2002), it is considered that SAN1 might be related to the dominance relationship on the pollen (Fukai et al. 2003). Therefore, we precisely surveyed the S-intergenic region of three additional S haplotypes to identify the homologous sequence to SAN1. However, we could not identify the orthologous gene of SAN1 in other class-II S haplotypes. Thus, this SAN1 might not be related to the dominance relationships in the pollen S gene.

Interestingly, the high sequence similarity was observed between the short repetitive sequence and 5'-upstream region of SP11 (Fig. 2). Especially in the case of
S<sup>9</sup> haplotype, a part of REP-C shares 81.6% identity with
the 5'-upstream region (-136 to -102 bp) of SP11-29. It has
been known that such a repetitive element is modified
by DNA methylation, representing a transposable
element (Okamoto et al. 2001). Moreover, in one of the
cruciferous plants, Arabidopsis suecica, which was an
allotetraploid hybrid of A. thaliana and A. arenosa, the
rRNA genes inherited from A. thaliana were transcriptionally
silenced by selective DNA/histone methylation in spacer elements, whereas rRNA genes derived from A. arenosa were transcribed (Chen et al. 1998; Lawrence et
al. 2004). Because the REP-C that exists in the S-intergenic
region is similarly included in the SP11 promoter region,
such a repetitive cluster might be the target of recessive allele specific modification as is the case as well in rRNA repression. Therefore, it is considered that it is
important to analyze the state of modification of DNA
and/or histone in the recessive allele of the SP11 promotor
and repetitive sequences. From these analyses, the molecular mechanisms of the dominance relationship of SP11 in pollen will be discovered in the near future.

This work was supported in part by Grants-in-Aid for Special Research on priority Areas (Nos. 11238201 and 14360002) and Grants-in-Aid for the 21st Century Center of Excellence Program to Nara Institute of Science and Technology and Iwate University from the Ministry of Education, Culture, Sports, Science, and Technology of Japan and a Grant-in-Aid for Creative Scientific Research (No. 16303016) from Japan Society for Promotion of Science (JSPS). T. K. and Y. T. are the recipients of a Research Fellowship of the JSPS for Young Scientists. The
authors are grateful to Akayo Chiba, Yukiko Ohyama, and Hiroyuki Ishikawa (Iwate University) for technical assistance.

**REFERENCES**