Characterization of Expressed Genes in the *SLL2* Region of Self-Compatible *Arabidopsis thaliana*

Yoshinobu Takada,[1] Akiko Ito,[1,†] Chie Ninomiya,[1] Tomohiro Kakizaki,[1] Yoshihito Takahata,[1] Go Suzuki,[2] Katsunori Hatakeyama,[3] Kokichi Hinata,[3] Hiroshi Shiba,[4] Seiji Takayama,[4] Akira Isogai,[4] and Masao Watanabe[1,*]

Faculty of Agriculture, Iwate University, Morioka 020-8550, Japan,[1] Division of Natural Science, Osaka Kyoiku University, Kashiwara 582-8582, Japan,[2] Research Institute of Seed Production Co. Ltd., Aoba-ku, Sendai 989-3204, Japan,[3] and Graduate School of Biological Sciences, Nara Institute of Science and Technology, Ikoma 630-0101, Japan[4]

(Received 21 May 2001; revised 17 August 2001)

Abstract

Self-incompatibility in *Brassica* species is regulated by a set of *S*-locus genes: *SLG*, *SRK*, and *SP11/SCR*. In the vicinity of the *S*-locus genes, several expressed genes, *SLL2* and *SP2/ClpP*, etc., were identified in *B. campestris*. *Arabidopsis thaliana* is a self-compatible *Brassica* relative, and its complete genome has been sequenced. From comparison of the genomic sequences between *B. campestris* and *A. thaliana*, microsynteny between gene clusters of *Arabidopsis* and *Brassica SLL2* regions was observed, though the *S*-locus genes, *SLG*, *SRK*, and *SP11/SCR* were not found in the region of *Arabidopsis*. Almost all genes predicted in this region of *Arabidopsis* were expressed in both vegetative and reproductive organs, suggesting that the genes in the *SLL2* region might not be related to self-incompatibility. Considering the recent speculation that the *S*-locus genes were translocated as a single unit between *Arabidopsis* and *Brassica*, the translocation might have occurred in the region between the *SLL2* and *SP7* genes.

Key words: *Arabidopsis thaliana*; expressed genes; genomic organization; *S* locus; synteny

Many flowering plants have mechanisms to avoid self-fertilization. One of such mechanisms is self-incompatibility (SI), which is defined as the inability of a fertile hermaphrodite plant to produce zygotes after self-pollination.[1] In *Brassica* species, the SI system is sporophytically controlled by a single *S* locus with multiple alleles.[2] At the *Brassica S* locus, three highly polymorphic genes, *SRK* (*S* receptor kinase), *SLG* (*S* locus glycoprotein) and *SP11/SCR* (*S* locus protein 11 or *S* locus cysteine-rich protein), regulate the recognition reaction of SI, and occur in tandem.[3–10]

In the flanking regions of the three SI genes, 12 expressed genes, including *SP2*, *SP4*, and *SLL2* (*S* locus-linked gene 2), were identified in the S^9 haplotype of *B. campestris*.[3] The genomic organization of *S* locus might be conserved in *Brassica* species.[5,11] In self-compatible cruciferous plants such as *Arabidopsis thaliana*, *SP2/ClpP*, and *SLL2* were identified on chromosome I, and the region was reported to be homeologous to the *S* locus region.[12] Because *SRK* and *SLG* were not in the *S* homeologous region of *A. thaliana*, the report was interesting as it provides evolutional evidence of deletion of self-recognition genes in the self-compatible *Brassica* relatives.

In the present report, by using information of genome sequence of *A. thaliana*, we analyzed the 170-kb contiguous DNA sequence that was expected to contain the *S*-locus homeologous region of *A. thaliana* in order to examine the gene organization and microsynteny of the *S*-homeologous regions at the sequence level. The results revealed that not only *SLG* and *SRK*, but also *SP11/SCR*, were absent in the *SLL2* region of *A. thaliana*. All of the genes expressed in this region were characterized in detail, and the evolutional relationship between the region and SI is discussed.

1. Genomic Organization of the *SLL2* Region in *A. thaliana*

According to Conner et al. (1998),[12] several genes whose orthologs were linked to the SI genes in *Brassica* were localized on the right arm of chromosome I in

Communicated by Satoshi Tabata

* To whom correspondence should be addressed. Laboratory of Plant Breeding, Faculty of Agriculture, Iwate University, 3-18-8, Ueda, Morioka 020-8550, Japan, Tel. +81-19-621-6152, Fax. +81-19-621-6177, E-mail: nabe@iwate-u.ac.jp

† Present address: Iwate Biotechnology Research Center, Kitakami 024-0003, Japan

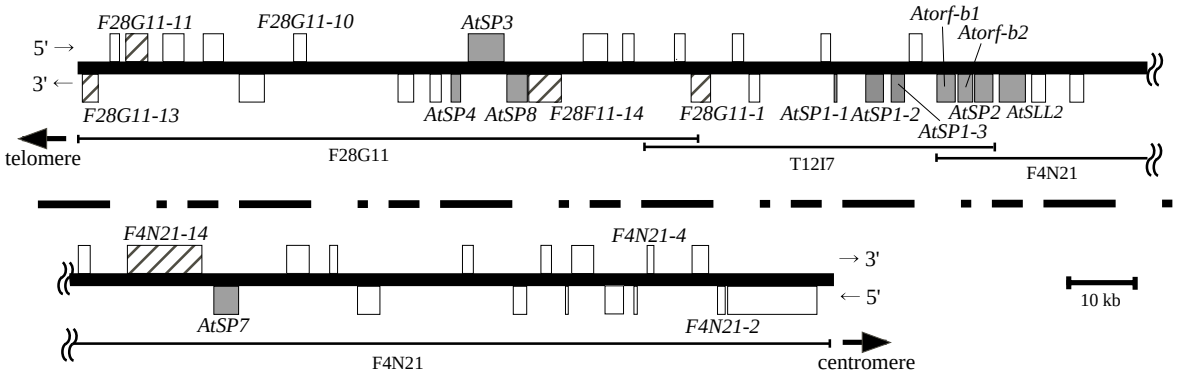

Figure 1. Gene map of the 170-kb contiguous genomic region of *A. thaliana SLL2* region. The region corresponding to the BAC F28G11, T12I7, F4N21 is indicated by the thin *bar* below the map. The boxes denote the location of the 47 predicted genes: the transcription of the genes is in the opposite direction above and below the physical map. Microsynteny observed between genes is indicated by filled boxes between *Arabidopsis* and *Brassica* (see text and Fig. 2 in detail). Striped boxes on the map represent the genes which were reported by Cui et al. (1999)[11] and Conner et al. (1998).[12] Open boxes indicate the predicted genes by AGI annotation. The possible existence of genes (exons) in the 170-kb contiguous DNA sequence was predicted by using GENSCAN program[18] and GenMark.[19]

A. thaliana. In Conner's map, *SLL2*[13] and *SP2/ClpP*[3,14] were tightly linked in *A. thaliana*. To determine whether or not other *Brassica S*-linked genes were linked to *SLL2* and *SP2/ClpP* in *Arabidopsis*, we searched the *Arabidopsis* BAC database at the TAIR (The Arabidopsis Information Resource) web site for BAC clones which contain *SLL2* and *SP2/ClpP*. We found a 180-kb BAC clone, named T12I7, containing both *SLL2* and *SP2/ClpP*, and speculated that T12I7 might include the *Brassica S*-locus counterpart region of *A. thaliana* reported in Conner et al. (1998).[12] When we found the BAC clone T12I7, the sequence data was merely draft data and contained many undetermined sequences. Therefore, we obtained BAC DNA of the T12I7 clones (kindly gifted from the Arabidopsis Biological Resource Center), for further analysis. PCR amplification with the T12I7 BAC DNA as a template confirmed the existence of the *SLL2*, *SP2/ClpP*, *SP3/CePP*, *SP4/SPA*, *SP7*, and *SP8/Fmt* (all of them were known to be linked to the *S*-locus genes in *Brassica*)[3,11,13–16] in the T12I7 clones. The nucleotide sequence homology of the *Brassica* and *Arabidopsis* genes ranged from 46% to 88%. Therefore, we designated them as *AtSLL2*, *AtSP2*, *AtSP3*, *AtSP4*, *AtSP7*, and *AtSP8*, respectively.

At the time the complete genome sequence of *A. thaliana* was published by AGI (Arabidopsis Genome Initiative),[17] the T12I7 region was covered by three new BAC clones (F28G11: AC074025, T12I7: AC079285, F4N21: AC013288). The contig of the three BAC clones (F28G11, T12I7, F4N21) was about 170-kb in length.

From the complete sequence of the 170-kb contig, 47 protein-coding regions were predicted by GENSCAN and GENMARK (Fig. 1); all of them were predicted as genes in the annotation of the AGI sequence.[17] The gene density of this region was high at one gene every 3.61 kb. *AtSLL2*, *AtSP2*, *AtSP3*, *AtSP4*, *AtSP7*, and *AtSP8* were located on these contigs, as shown in a gene map of *Arabidopsis SLL2* region (Fig. 1). *AtSP3*, *AtSP4*, and *AtSP8* were clustered, and *AtSP2* was located in the immediate vicinity of *SLL2*; these gene arrangements are similar to that observed in the S^9 haplotype of *B. campestris*.[3] Three annotated ORFs (T12I7-6, T12I7-7, and T12I7-8) homologous to *SP1* were tandemly located upstream of *AtSP2*, and were named as *AtSP1-1*, *AtSP1-2*, and *AtSP1-3*, respectively (Fig. 2). In the case of *orf-b*, one of the ORFs in the *Brassica* S^9 locus estimated by Suzuki et al. (1999),[3] two homologous ORFs (F4N21-20 and F4N21-21) were tandemly duplicated between *AtSP1-3* and *AtSP2*, and were designated as *Atorf-b1*, *Atorf-b2*, respectively The high degree of synteny between gene clusters of *Arabidopsis* and *Brassica SLL2* regions indicates that the two regions are homeologous (Fig. 2).

However, *SP5*, *SP6*, *SP10*, *orf-a*, *orf-c*, and *SAE1* that were reported to be linked to the *S* locus in the S^9 haplotype of *B. campestris*[3,15] were not identified in T12I7. The counterpart of *SP5* had been found in chromosome I, as described previously.[3] An ORF homologous to *SAE1* was mapped in chromosome V. ORFs homologous to both *SP10* and *orf-c* were widely scattered in the *Arabidopsis* genome. In contrast, no *orf-a* counterpart was found in *Arabidopsis*. Because of the short nucleotide sequence of *SP6*, we could not determine which gene was a counterpart of *Arabidopsis*.

In the 170-kb *SLL2* region, putative genes for F28G11-13 (calmodulin), F28G11-11 (fructokinase), F28G11-1 (sucrose proton sympoter), and F4N21-14 (DNA ligase), whose orthologs were mapped on the flanking region of the *S* locus in *Brassica* species,[11,12] were also identified, indicating that this *SLL2* region of *Arabidopsis*

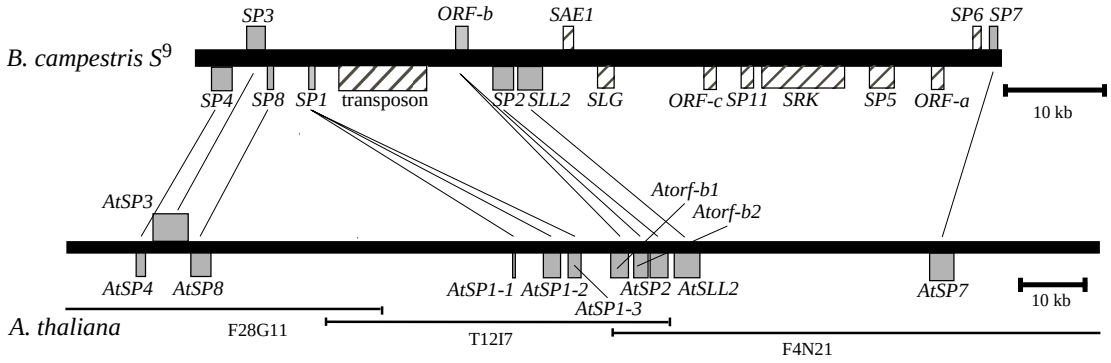

Figure 2. Comparative map of *SLL2* region between *B. campestris* S^9 haplotype and *A. thaliana*. Microsynteny was observed between genes indicated by filled boxes between *Arabidopsis* and *Brassica*. The sequences represented by striped boxes in *B. campestris* S^9 were not identified in the *Arabidopsis SLL2* region. Homology searches were performed using the BLAST program[20] and GENETYX-WIN Ver. 5.0 software (Software Development).

is homeologous to the *Brassica S* region (Fig. 1).

Sequences orthologous to *Brassica* SI genes, such as *SRK*, *SLG*, and *SP11/SCR*, were not found in the *Arabidopsis SLL2* region. Very recently, Kusaba et al. (2001)[21] reported that *SRK* and *SP11/SCR* orthologs were mapped to an *ARK3* region of chromosome IV of *A. thaliana* from the comparison with the *S* locus region of self-incompatible *A. lyrata*, a close relative of *A. thaliana*. The *SRK* and *SP11/SCR* orthologs of *A. thaliana* encodes nonfunctional protein, suggesting that *A. thaliana* is self-compatible due to inactivation of the SI genes by mutation. Thus, the *Brassica S* locus is located in a region that is syntenous with the *SLL2* region of *A. thaliana* chromosome I,[12] whereas the *A. lyrata S* locus is in a region corresponding to the *ARK3* region of *A. thaliana* chromosome IV.[21] Therefore, the *S* locus complex was translocated as a unit between these chromosomal positions, and only genes related to self-recognition, *SRK* and *SP11/SCR*, were included in the unit. Although the direction of the *S*-locus translocation events cannot be determined, either insertion of SI genes in *Brassica* or deletion of those in *Arabidopsis* occurred in the *SLL2* region, possibly between *SLL2* and *SP7* (see the gene map in Fig. 2), because the SI genes are located between *SLL2* and *SP7* in the S^9 haplotype of *B. campestris*.[3]

2. Expression of Predicted Genes in the *Arabidopsis SLL2* Region

In a wild species of tomato having gametophytic SI, genes regulating floral traits have been mapped near the *S* locus, indicating that genes related to reproductive functions formed a cluster.[22] Thus, it is important to determine the function of the genes, which are located at the flanking region of *S* locus in *Brassica* species.

Thus, we performed RT-PCR analysis followed by DNA gel blot analysis to determine the expression of the genes in the *Arabidopsis* 170-kb *SLL2* region. *AtSP2*, *AtSP3*, *AtSP4*, *AtSP7*, *AtSP8*, and *AtSLL2* were expressed in both leaves and flower buds (Fig. 3). RT-PCR products of *AtSP3* were more abundant in leaves than in flower buds. The signal intensity of the RT-PCR products of *AtSP2*, *AtSP4*, *AtSP7*, *AtSP8*, and *AtSLL2* was similar between samples of leaves and flower buds. Among other genes located in the *SLL2* region, we performed RT-PCR analysis for selected 22 genes. Representative results of RT-PCR of six genes are presented in Fig. 3. Expression of the *LTP* (Lipid transfer protein: F4N21-4) was identified in flower buds but not in leaves. The expression of serine/threonine-type protein kinase (F28G11-10) and β-1,3-glucanase-like protein (F4N21-2) was predominant in flower buds. Therefore, it is possible that the genes for LTP, serine/threonine-type protein kinase, and β-1,3-glucanase-like protein are involved in the biological reaction of the reproductive organs, for example, pollen-stigma interaction. In fact, a gene homologous to *LTP* (F4N21-4) was specifically expressed in the anther at the uninucleate microspore stage.[25] During microsporogenesis, a β-1,3-glucanase has important roles in the degradation of callose, a β-1,3-glucan.[26,27] However, other all genes we examined were expressed in both leaves and flower buds with the same intensity, as in the case of calmodulin (F28G11-13), zinc-finger protein (F4N21-6) and cytochrome P450 (F28G11-4) in Fig. 3. Thus, expression of most of the genes in the *SLL2* region was not specific to flower buds. It is likely that these genes in the *SLL2* region are related to housekeeping functions and/or vegetative development rather than pollen-stigma interaction. It appears that the *SLL2* region is not related to SI recognition reaction. This conclusion is consistent with the report that the true *S* homeologous locus of *A. thaliana* is the region containing the *SRK* and *SP11/SCR* orthologs near the *ARK3* gene in

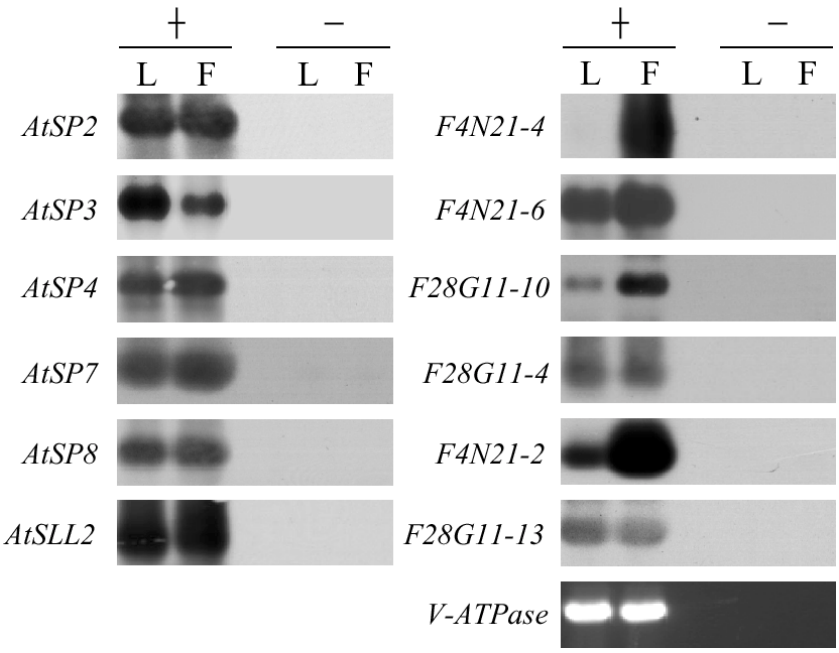

Figure 3. Expression of the predicted genes in *Arabidopsis SLL2* region. Poly(A)$^+$ RNA was extracted from leaves (L) and flower buds (F) of *A. thaliana* ecotype *Columbia* using the FastTrack mRNA isolation kit (Invitrogen). A reverse transcriptase reaction was carried out in the presence (+) or absence (−) of reverse transcriptase enzyme (First-Strand cDNA synthesis kit; Amersham-Pharmacia), and then cDNA was used as a template for PCR amplification with gene-specific primers. RT-PCR was performed according to the method of Watanabe et al. (2000).[6] The *V-ATPase* gene[23] was amplified as a positive control. The PCR products were size-fractionated by agarose electrophoresis and transferred to a nylon membrane (Roche). Hybridization and detection of the hybridized probe was carried out as described by Watanabe et al. (2000)[6] and Matsuda et al. (1996),[24] except that the membranes were washed twice in 0.1 × SSC, 0.1% SDS at 65°C for 20 min.

chromosome IV.[21]

Acknowledgements: This work was supported in part by Grant-in-Aid for Special Research on Priority Areas (B) (11238201) and for Scientific Research (B) (11460001) to M. W. from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

1. Nettarncourt, de, D. 1977, Incompatibility in Angiosperms, Springer-Verlag, Berlin, pp. 1–230.
2. Bateman, A. J. 1955, Self-incompatibility system in angiosperms. III. Cruciferae, *Heredity*, **9**, 52–68.
3. Suzuki, G., Kai, N., Hirose, T. et al. 1999, Genomic organization of the *S* locus: identification and characterization of genes in *SLG/SRK* region of S^9 haplotype of *Brassica campestris* (syn. *rapa*), *Genetics*, **153**, 391–400.
4. Takayama, S., Shiba, H., Iwano, M. et al. 2000, The pollen determinant of self-incompatibility in *Brassica campestris*, *Proc. Natl. Acad. Sci. USA*, **97**, 1920–1925.
5. Suzuki, G., Watanabe, M., and Nishio, T. 2000, Physical distances between *S-locus* genes in various *S* haplotypes of *Brassica rapa* and *B. oleracea*, *Theor. Appl. Genet.*, **101**, 80–85.
6. Watanabe, M., Ito, A., Takada, Y. et al. 2000, Highly divergent sequences of the pollen self-incompatibility (*S*) gene in class-I haplotypes of *Brassica campestris* (syn. *rapa*) L, *FEBS Lett.*, **473**, 139–144.
7. Takasaki, T., Hatakeyama, K., Suzuki, G., Watanabe, M., Isogai, A., and Hinata, K. 2000, The *S* receptor kinase determines self-incompatibility in *Brassica* stigma, *Nature*, **403**, 913–916.
8. Dixit, R., Nasrallah, M. E., and Nasrallah, J. B. 2000, Post-transcriptional maturation of the *S* receptor kinase of *Brassica* correlates with co-expression of the *S*-locus glycoprotein in the stigmas of two *Brassica* strains and in transgenic tobacco plants, *Plant Physiol.*, **124**, 297–312.
9. Schopfer, C. R., Nasrallah, M. E., and Nasrallah, J. B. 1999, The male determinant of self-incompatibility in *Brassica*, *Science*, **286**, 1697–1700.
10. Shiba, H., Takayama, S., Iwano, M. et al. 2001, A pollen coat protein, SP11/SCR, determines the pollen *S*-specificity in the self-incompatibility of *Brassica* species, *Plant Physiol.*, **125**, 2095–2103.
11. Cui, Y., Brugiere, N., Jackman, L., Bi, Y.-M., and Rothstein, S. J. 1999, Structural and transcriptional comparative analysis of the *S* locus regions in two self-incompatible *Brassica napus* lines, *Plant Cell*, **11**, 2217–2232.
12. Conner, J. A., Conner, P., Nasrallah, M. E., and Nasrallah, J. B. 1998, Comparative mapping of the *Brassica S* locus region and its homeolog in *Arabidopsis*: implications for the evolution of mating systems in the

Brassicaceae, *Plant Cell*, **10**, 801–812.
13. Yu, K., Schafer, U., Glavin, T. L., Goring, D. R., and Rothstein, S. J. 1996, Molecular characterization of the *S* locus in two self-incompatible *Brassica napus* lines, *Plant Cell*, **8**, 2369–2380.
14. Letham, D. L. L. and Nasrallah, J. B. 1998, A *ClpP* homolog linked to the *Brassica* self-incompatibility (*S*) locus, *Sex. Plant Reprod.*, **11**, 117–119.
15. Watanabe, M., Suzuki, G., Toriyama, K., Takayama, S., Isogai, A., and Hinata, K. 1999, Two anther-expressed genes downstream of SLG^9: identification of a novel *S*-linked gene specifically expressed in anthers at the uninucleate stage of *Brassica campestris* (syn. *rapa*) L, *Sex. Plant Reprod.*, **12**, 127–134.
16. Casselman, A. L., Nasrallah, M. E., and Nasrallah, J. B. 2000, Using *S*-locus deletions to evaluate self-incompatibility candidate genes: an example for a novel anther-expressed gene, *Sex. Plant Reprod.*, **12**, 227–231.
17. Arabidopsis Genome Initiative 2000, Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*, *Nature*, **408**, 796–815.
18. Burge, C. and Karlin, S. 1997, Prediction of complete gene structures in human genomic DNA, *J. Mol. Biol.*, **268**, 78–94.
19. Borodovsky, M. and McIninch, J. 1993, GeneMark: Parallel gene recognition for both DNA strands, *Computers Chem.*, **17**, 123–133.
20. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. 1990, Basic local alignment search tool, *J. Mol. Biol.*, **215**, 403–410.
21. Kusaba, M., Dwyer, K., Hendershot, J., Vrebalov, J., Nasrallah, J. B., and Nasrallah, M. E. 2001, Self-incompatibility in the genus *Arabidopsis*: characterization of the *S* locus in the outcrossing *A. lyrata* and its autogamous relative *A. thaliana*, *Plant Cell*, **13**, 627–643.
22. Bernacchi, D. and Tanksley, S. D. 1997, An interspecifc backcross of *Lycopersicon esculentum*×*L. hirsutum*: linkage analysis and QTL study of sexual compatibility factors and floral traits, *Genetics*, **147**, 861–877.
23. Manolson, M. F., Oucllette, B. F. F., Filion, M., and Poole, R. J. 1988, cDNA sequence and homologies of the "57-kDa" nucleotide-binding subunit of the vacuolar ATPase from *Arabidopsis*, *J. Biol. Chem.*, **263**, 17987–17994.
24. Matsuda, N., Tsuchiya, T., Kishitani, S., Tanaka, Y., and Toriyama, K. 1996, Partial male sterility in transgenic tobacco carrying antisense and sense *PAL* cDNA under the control of a tapetum-specific promoter, *Plant Cell Physiol.*, **37**, 215–222.
25. Toriyama, K., Hanaoka, K., Okada, T., and Watanabe, M. 1998, Molecular cloning of a cDNA encoding a pollen extracellular protein as a potential source of a pollen allergen in *Brassica rapa*, *FEBS Lett.*, **424**, 234–238.
26. Chasan, R. 1992, Breaching the callose wall, *Plant Cell*, **4**, 745–746.
27. Steiglitz, H. 1977, Role of β-1,3-glucanase in postmeiotic microspore release, *Dev. Biol.*, **57**, 87–97.