Plant Cell Rep (2006) 25: 466–474
DOI 10.1007/s00299-005-0077-2

GENETICS AND GENOMICS

Jong-In Park · Hirokazu Hakozaki · Makoto Endo · Yoshinobu Takada · Hitoshi Ito · Masanori Uchida · Tomihiro Okabe · Masao Watanabe

Molecular characterization of mature pollen-specific genes encoding novel small cysteine-rich proteins in rice (*Oryza sativa* L.)

Received: 15 July 2005 / Revised: 30 September 2005 / Accepted: 2 October 2005 / Published online: 6 January 2006

Abstract In our previous cDNA microarray analysis, we identified 53 mature anther-specific genes, whose function was unknown, in rice. We reanalyzed these genes from the viewpoint of the specific amino acid motif. Out of 53 genes, three genes, *Os-26*, *Os-32*, and *Os-169* (renamed as *OsSCP1*, *OsSCP2*, and *OsSCP3*), encoded cysteine-rich motif (Cys-X_3-Cys-X_{13}-Cys-X_3-Cys), indicating that they were novel small cysteine-rich proteins. From the search of specific elements in promoter regions, several pollen-specific elements were found. In order to determine whether three promoters were functional in pollen or not, the gene constructs with promoter regions fused to the β-glucuronidase gene were transformed into tobacco. Histochemical analysis showed that these promoters were active in the mature pollen grains and pollen tubes. Furthermore, *OsSCP1* and *OsSCP3* formed a multigene family tandemly in the rice genome. From the results, *OsSCP*s might have important roles in mature pollen development and pollen tube growth.

Keywords Gene family · Mature pollen-specific genes · *Oryza sativa* L. · Promoter analysis · Small cysteine-rich protein

Abbreviations GUS: β-glucuronidase · OsSCPs: *Oryza sativa* small cysteine-rich proteins · RT-PCR: reverse transcription – polymerase chain reaction

Communicated by K. Toriyama

J.-I. Park · H. Hakozaki · M. Endo · H. Ito · M. Uchida · T. Okabe
Faculty of Agriculture, Iwate University,
Morioka 020-8550, Japan

H. Hakozaki · M. Endo · M. Watanabe (✉)
Laboratory of Plant Reproductive Genetics, Graduate School of Life Sciences, Tohoku University,
2-1-1 Katahira, Aoba-ku, Sendai 980-8577, Japan
e-mail: nabe@ige.tohoku.ac.jp
Tel.: +81-22-217-5681
Fax: +81-22-217-5683

Y. Takada
Division of Natural Science, Osaka Kyoiku University,
Kashiwara 582-8582, Japan

M. Watanabe
The 21st Century Center of Excellence Program,
Iwate University,
Morioka 020-8550, Japan

Present address:
H. Ito
Fukushima Meisei High School,
Fukushima 960-1192, Japan

M. Uchida
Fukushima Prefecture Agricultural College,
Yabuki 969-0292, Japan

T. Okabe
Higashi Shirakawa Agricultural High School,
Tanakura 963-6131, Japan

Introduction

The male gametophyte, pollen, develops within the anther compartment of the stamen and requires cooperative functional interactions between gametophytic and sporophytic tissues. Pollen development consists of several distinct stages. A microspore mother cell undergoes meiosis to give rise to a tetrad of four microspores. After release upon the dissolution of the callose wall, each uninucleate microspore undergoes an asymmetric mitotic division to give rise to two cells with distinct fates: the vegetative cell and the generative cell. The larger vegetative cell is transcriptionally active (Bedinger 1992; McCormick 2004). Recently, a number of anther-specific cDNA clones have been identified and isolated using DNA microarray and macroarray techniques from several model plants, *Lotus japonicus*, *Arabidopsis thaliana*, and *Oryza sativa* (Endo et al. 2002; Amagai et al. 2003; Honys and Twell 2003; Endo et al. 2004). Although pollen-specific genes have been studied extensively in most cases, it remains unclear what roles these genes play.

The genes encoding small cysteine-rich proteins form a gene family in several plant species. The small cysteine-rich

proteins are highly differentiated in their biological function and were related to several physiological phenomena, e.g. storage, protection, lipid transfer, cell–cell communication, etc. (Jose-Estanyol et al. 2004). In the case of cell–cell communication, *SP11*/*SCR*, encoding a male *S* determinant in *Brassica* self-incompatibility (SI), is a small cysteine-rich protein, and functioned as a ligand molecule for female *S* determinant, *SRK*, *S* receptor kinase (reviewed in Watanabe et al. 2003; Takayama and Isogai 2005). In another case, lipid transfer protein, LTP, also encoding a different type of small cysteine-rich protein, has an important role in transferring lipids in several tissues (Kader 1996). The pollen-specific gene, *LAT52*, also encodes a small cysteine-rich protein, which could interact with pollen-specific receptor protein kinase, LePRK2 (Twell et al. 1989; Tang et al. 2002). LAT52 is required for the pollen to germinate in vitro and achieve fertilization in vivo (Muschietti et al. 1994).

In the previous experiments, we identified over 100 anther-specific genes in rice (Endo et al. 2004). In this current study, we identified three novel pollen-specific clones, *OsSCP1*, *OsSCP2*, and *OsSCP3*. The deduced amino acid sequences of OsSCPs contained an open reading frame (ORF) encoding a small cysteine-rich protein. In order to determine whether the promoter regions of *OsSCP*s were functional in pollen or not, transgenic tobacco plants having the β-glucuronidase (*GUS*) gene regulated by *OsSCP*s promoter were made. A blue *GUS* signal was specifically observed in mature pollen grains and pollen tubes. Furthermore, *OsSCP1* and *OsSCP3* formed a multigene family in the rice genome. We discuss the function of the pollen-specific *OsSCP* genes.

Materials and methods

Plant material and mRNA isolation

Plants of *O. sativa* cv. Koshihikari were grown in a green house. Anther and pistil tissues in various stages (A1: anther of uninucleate stage; A2: anther of binucleate stage; A3: anther of trinucleate stage; P1: pistil of uninucleate stage; P2: pistil of binucleate stage; and P3: pistil of trinucleate stage) were collected for isolation mRNA as described previously (Endo et al. 2004). Leaf sheath (Ls) and leaf blade (Lb) were harvested at the heading stage. Furthermore, root (R) and shoot (S) of seedling were also collected for RNA experiment. Lemma and palea (L/P) were also harvested for mRNA isolation. Isolation of poly $(A)^+$ RNA from each tissue was performed using a FastTrack 2.0 mRNA isolation kit (Invitrogen, San Diego, CA) as described in Watanabe et al. (2000).

RT-PCR

RT-PCR was performed according to Endo et al. (2002). Briefly, poly $(A)^+$ RNA was reverse-transcribed to synthesize the first strand cDNA by using the First-Strand cDNA synthesis kit (Amersham-Pharmacia, Uppsala, Sweden). Then, cDNA was used as a template for PCR amplification with a set consisting of primers specific to each gene: OsSCP1RT-F(5′-ATGGCCCAGAACAAGACCAT-3′),OsSCP1RT-R(5′-GCAACTGTCTACGCACTTCTTGTTGGTG-3′),OSCP2 RT-F(5′-ATGGCCCAAAACAAGACCAT-3′),OsSCP2 RT-R(5′-TCCTTGGCTTCCTTAGACATG-3′),OsSCP3R T-F(5′-CAGAACAAGACCATTGCGGT-3′), and OsSCP3 RT-R(5′-CGTTGTTTGAATATAAGCGACGAC-3′). PCR was performed with Ex*Taq* DNA polymerase (TaKaRa Shuzo, Shiga, Japan) for 25 or 35 cycles of denaturation for 30 s at 94°C, annealing for 30 s at 58°C and extension for 1 min at 72°C, followed by a final extension for 5 min. Actin primers as a control were 5′-TCCATCTTGGCATCTCTCAG-3′ (forward) and 5′-GTACCCGCATCAGGCATCTG-3′ (reverse).

5′-RACE and sequence analysis

Full-length cDNA clones of *OsSCP*s from the cultivar, Koshihikari, were obtained by the 5′-RACE procedure according to Watanabe et al. (1999). cDNA inserts were sequenced by the dideoxy chain-terminator method using a model 310 DNA sequencer (Applied Biosystems, Foster City, CA). The DNA sequence was analyzed with Genetyx software (Software Development, Tokyo, Japan). A homology search was performed using BLAST (Altschul et al. 1997). The motif search was performed with the TAIR web site program (http://www.arabidopsis.org/cgi-bin/patmatch/nph-patmatch.pl). The presence of a signal peptide was predicted using the PSORT program (http://psort.nibb.ac.jp).

Construction of promoter–GUS fusion

The promoter regions of *OsSCP1*, *OsSCP2*, and *OsSCP3* were identified from a DNA database of genomic sequences of rice. These promoter regions were amplified by PCR using each of the following primers: OsSCP1pro-F (5′-ACTAGTTCTGTTGTTCTACCCAATATTGTTAC-3′), OsSCP1pro-R (5′-GGATCCTAGCACCGTTTCTATTGAGGGTAACTGGTGTGGAT-3′), OsSCP2pro-F (5′-ACTAGTACTGGGTCGACTTGGGCCGGGGAGGAGAGAGAAA-3′), OsSCP2pro-R (5′-GGATCCTAGCACCGTTTATGCTGAGCAGAATTGGT-3′), OsSCP3pro-F (5′-ACTAGTCTGCCACTGTCTACAGCGGGAAGGTTAACAAT-3′), and OsSCP3pro-R (5′-GGATCCTAGCGCTGTTTTTCTTGAGTGTAATTGGTGTGGA-3′). All PCR fragments were subcloned into the pCR2.1 plasmid vector (Invitrogen). The nucleotide sequence of these promoter fragments was confirmed by sequencing of the plasmid using the DNA sequencer (ABI 310; PE Biosystems, Foster City, CA). The binary vector, pBI–GSH, was used. This vector was derived from pBI–BG vector (Okada et al. 2000) by replacement of the *Bra r* 1 promoter region to multi-cloning sites. The promoter region of each *OsSCP* was taken as a *Spe*I–*Bam*HI fragment from each plasmid

and inserted into the *Spe*I–*Bam*HI site of pBI–GSH. As a result, each promoter region was placed in front of the reporter gene, *GUS*. The constructs, each formed by promoter–GUS fusion, were named *OsSCP1–GUS*, *OsSCP2–GUS*, and *OsSCP3–GUS*, respectively.

Transformation of tobacco plants

The constructed *OsSCP1–GUS*, *OsSCP2–GUS*, and *OsSCP3–GUS* were transferred to *Agrobacterium tumefaciens* strain EHA101 using the freeze-thaw method (An et al. 1988). Leaf disks of *Nicotiana tabacum* cv. Petit Havana SR1 were transformed according to the procedure of Horsch et al. (1988). Transgenic tobacco plants were selected on the basis of hygromycin resistance (100 μg/ml). These plants were maintained under greenhouse conditions until maturity. The vegetative and floral tissues were used in GUS assay.

Histochemical GUS assay

Vegetative tissues, and anther and pollen of different developmental stages, were obtained from primary transgenic plants and wild-type plants as a control. GUS staining was performed according to Tsuchiya et al. (1994), except for addition of 20% methanol in the staining solution. Briefly, for staining, several tissues were incubated in 200 mM $NaPO_4$, pH 7.0, 20% methanol, 0.3% Triton X-100, 12.5 mM $K_3Fe(CN)_6$, 12.5 mM $K_4Fe(CN)_6$, and 38.3 mM X-Gluc (5-bromo-4-chloro-3-indoyl-β-D-glucuronide). Plant tissues were vacuum infiltrated briefly, then incubated at 37°C overnight. After staining, chlorophyll was cleared from the sample by 75% ethanol treatment. Photos were taken using a Nikon E800 microscope system. The developmental stages of microspores and pollen grains were determined by staining nuclei with DAPI (Watanabe et al. 1991). Pollen germination was performed according to Lush et al. (1998).

Results

Identification and characterization of *OsSCP* genes

In our previous study, we isolated 89 mature anther-specific genes (Cluster RA3 genes) in rice by using cDNA microarray analysis (Endo et al. 2004). However, the function of 60% of the Cluster RA3 genes (53 genes) was unknown at the time we published the array data. Because a large number of ESTs (expressed sequence tags) and full-length cDNA clones were deposited in the public database after publishing our data (Yazaki et al. 2004), we reanalyzed these unknown function genes and their encoding proteins from the viewpoint of the specific amino acid motifs. Within 53 genes, we found that three genes, *Os-26*, *Os-32*, and *Os-169* (renamed as *OsSCP1*, *OsSCP2*, and *OsSCP3*), had a small ORF, which was constructed from 127, 129, 129 amino acid residues, respectively, and encoded specific cysteine-rich motif (Cys–X_3–Cys–X_{13}–Cys–X_3–Cys; Fig. 1). Based on the homology search of *OsSCP*s against the rice full-length cDNA database with the BLAST program, *OsSCP1*, *OsSCP2*, and *OsSCP3* were found to be identical to full-length cDNA clones, AK064756 (chromosome 2), AK071781 (chromosome 8), and AK120999 (chromosome 2), respectively. Among three OsSCPs, the sequence identity was 35–47% at the amino acid level, and eight cysteine residues were completely conserved. Interestingly, the nucleotide difference between our cDNA clone and the full-length cDNA clone in each gene was not observed in both ORF and UTR (5′ and 3′) regions in spite of the difference of material cultivar, Koshihikari (our cDNA clone) and Nipponbare (full-length cDNA clone). Based on the comparison between cDNA and the genomic clone, the *OsSCP*s had no intron. By using the full-length cDNA information of *OsSCP*s, a BLASTP search of the OsSCPs was performed against the public database. However, OsSCPs showed no significant similarity to other known proteins. As described above, a number of small cysteine-rich proteins have been identified and characterized (Jose-Estanyol et al. 2004). When the motif search against OsSCPs was performed to find proteins, which have the same cysteine-rich motif as the OsSCPs on the web site, eight clones encoding OsSCP-like proteins having a specific cysteine motif were found only in the rice genome, not in other plant genomes (data not shown). Furthermore, we performed a signal peptide search with PSORT program. All three OsSCP proteins had N-terminal signal peptide, like other pollen-specific cysteine-rich proteins, SLR1-BP, PEC-3, and PCP-A1 (Stanchev et al. 1996; Toriyama et al. 1998; Takayama et al. 2000a). When compared the nucleotide sequences of these eight cDNA clones each other, eight cDNA clones were divided into three groups. In the first groups, the nucleotide sequence of three cDNA clones

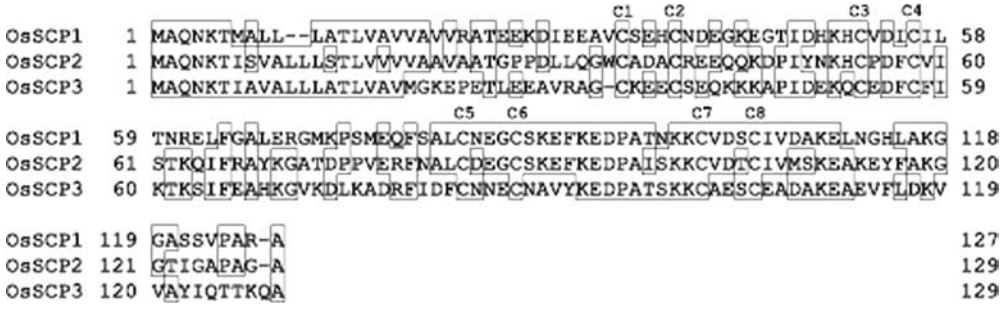

Fig. 1 Multiple alignment of amino acid sequences of OsSCP1, OsSCP2, and OsSCP3. Gaps were introduced to optimize the alignment. Conserved amino acid residues are indicated by *boxes*. The conserved eight cysteine residues in the mature proteins are shown as C1–C8

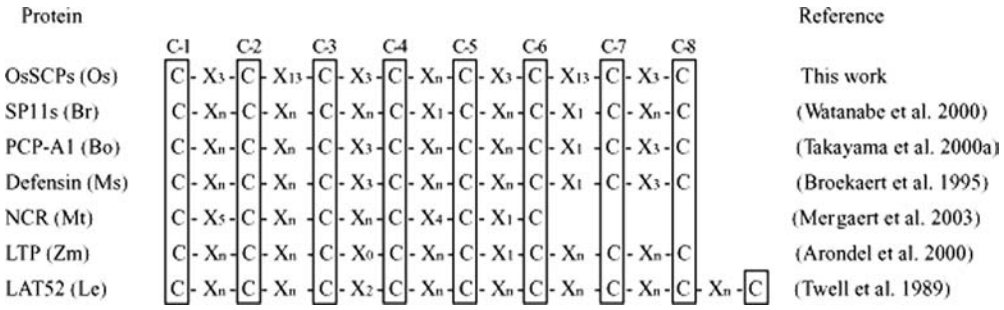

Fig. 2 Multiple alignment of the cysteine motifs of OsSCPs with the cysteine motifs of other proteins. Os, *Oryza sativa*; Br, *Brassica rapa*; Bo, *Brassica oleracea*; Ms, *Medicago sativa*; Mt, *Medicago truncatula*; Zm, *Zea mays*; Le, *Lycopersicon esculentum*

was identical to that of *OsSCP1*. In the second group, the nucleotide sequence of two cDNA clones was identical to that of *OsSCP2*. In the third group, the nucleotide sequence of three cDNA clones was identical to that of *OsSCP3*. Furthermore, the *OsSCP*s formed a small gene family in the rice genome. From these results, OsSCPs were found to be a novel type of cysteine-rich protein (Fig. 2). Thus, we carefully analyzed the *OsSCP*s.

In order to confirm the expression pattern of *OsSCP*s, RT-PCR was performed with several rice tissues. In each gene, specific amplification was observed only in the anther at the trinucleate stage containing mature pollen grain. In other anther at the developmental stages and other tissues, no band was detected (Fig. 3). This expression pattern was coincident to that observed by cDNA microarray analysis (Endo et al. 2004).

In order to determine the promoter region in each *OsSCP*, we compared the nucleotide sequence between full-length cDNA clone and genomic DNA clone of each *OsSCP* gene. We defined the promoter region as the nucleotide sequence between *OsSCP* and the nearest upstream ORF (intergenic region). The length of the putative promoter region of each *OsSCP* was 2033 bp, 952 bp, and 1845 bp, respectively (Fig. 4). The putative TATA box sequence was found at −74 for all *OsSCP* promoter sequences when the start of initiation codon (ATG) was marked as +1. Furthermore, in order to determine the spatial and temporal expression pattern of the *OsSCP*s, the 2033 bp promoter region (nucleotides −2034 to −1) of *OsSCP1* (*OsSCP1-b*, see below),

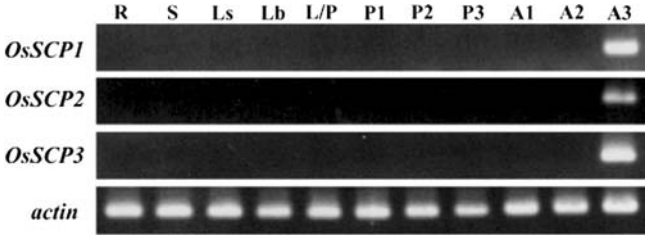

Fig. 3 RT-PCR analysis of anther-specific cDNA clones, *OsSCP1*, *OsSCP2*, and *OsSCP3*. The mRNA was isolated from root of seedling (R), shoot of seedling (S), leaf sheath at the heading stage (Ls), leaf blade at the heading stage (Lb), lemma and palea (L/P), pistil of uninucleate stage (P1), pistil of binucleate stage (P2), pistil of trinucleate stage (P3), anther of uninucleate stage (A1), anther of binucleate stage (A2), and anther of trinucleate stage (A3) in rice. The gene encoding actin was used as a positive control

the 952 bp promoter region (nucleotides −953 to −1) of *OsSCP2*, and the 1845 bp promoter region (nucleotides −1846 to −1) of *OsSCP3* (*OsSCP3-a*, see below) were fused with β-glucuronidase (*GUS*) genes (Fig. 4). For histochemical GUS staining, two to three independent tobacco transgenic plants were used in each *OsSCP* gene. All the transgenic lines of *OsSCP1* (*OsSCP1-b*)–*GUS*, *OsSCP2*–*GUS*, and *OsSCP3* (*OsSCP3-a*)–*GUS* exhibited GUS activity in pollens at a very late developmental stage, as described below (Fig. 5). No GUS activity was detected in other floral organs and vegetative organs (data not shown).

For the next step, we examined the GUS activity in anthers of transgenic plants according to different devel-

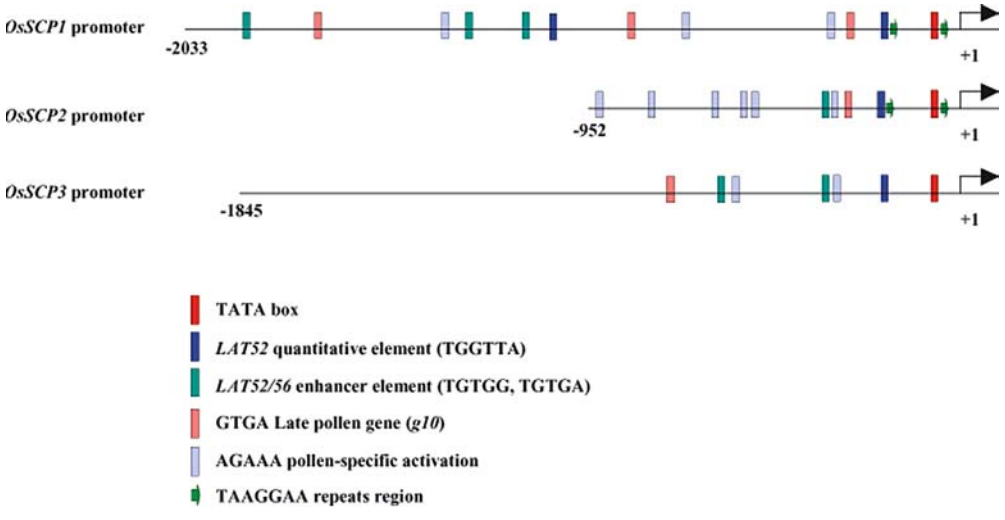

Fig. 4 Physical map of the promoter regions of *OsSCP1*, *OsSCP2*, and *OsSCP3*. Nucleotides are numbered with the first nucleotide of the translation initiation codon marked as +1

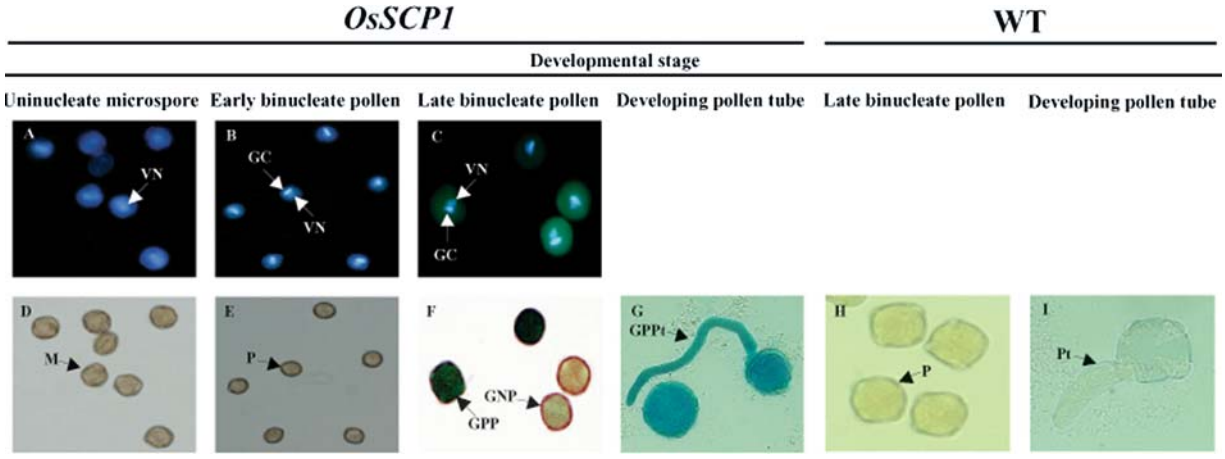

Fig. 5 Histochemical analysis of GUS activity conferred by *OsSCP1–GUS* and wild-type in transgenic tobacco. Developmental stage of microspore and pollen was determined by DAPI staining (**A**, **B**, **C**). Representative photographs of uninucleate microspore (**D**), early binucleate pollen (**E**), late binucleate pollen (**F**), and germinated pollen and pollen tube (**G**) in transgenic plant are shown. Blue staining indicates GUS activity. In wild-type plant, GUS activity was not observed in late binucleate pollen (**H**) and germinated pollen and pollen tube (**I**). WT, wild type; VN, vegetative nucleus; GC, generative cell; M, microspore; P, pollen; GPP, GUS-positive pollen; GNP, GUS-negative pollen; GPPt, GUS-positive pollen tube; Pt, pollen tube

opmental stages. Developmental stage of microspore and pollen was determined by staining of nuclei with DAPI (Fig. 5A–C). Because transgenic plants having each *OsSCP* construct showed the similar GUS expression pattern, we presented the *OsSCP1* construct as a representative. GUS activity was specifically observed in late binucleate pollen grain (Fig. 5F). In other words, GUS activity was not detected in uninucleate microspores or in early binucleate pollen grains (Fig. 5D and E). After growing the pollen tube in the pollen germination medium in vitro, GUS expression was also detected from the pollen tube of all transgenic plants (Fig. 5G). Segregation of GUS-positive and GUS-negative pollen grains was observed, as shown in Fig. 5C. The ratio of GUS-positive and GUS-negative pollen grains was 504:464 in transgenic plant having *OsSCP1–GUS* construct, indicating a segregation pattern of 1:1 (significant at the 5% level). In the non-transgenic wild-type plants, no GUS expression was detected in the pollen grain and the pollen tube (Fig. 5H and I).

Identification of pollen-specific *cis*-elements of the *OsSCP* genes

We surveyed the potential pollen-specific *cis*-elements, which have been already identified in other plant species, in each promoter region. The sequence motifs of the *LAT52*/*LAT56* (TGTGG and TGTGA) enhancer element, *LAT52* quantitative element (TGGTTA), *LAT52* pollen-specific activation element (AGAAA), *g10* late pollen gene element (GTGA), which were required for the specific expression of mature pollen (Twell et al. 1990; Eyal et al. 1995; Bate and Twell 1998; Rogers et al. 2001), were observed. In the case of *LAT52*/*LAT56*, the nucleotide sequences of the enhancer elements (TGTGG and TGTGA) were found at −1850, −1285, and −1130 for the *OsSCP1* (*OsSCP1-b*) promoter, at −345 for the *OsSCP2* promoter, and at −617 and −345 for the *OsSCP3* (*OsSCP3-a*) promoter (Fig. 4). The sequence motifs of the *LAT52* quantitative element were present at −1059 and −201 for the *OsSCP1* (*OsSCP1-b*) promoter, at −208 for the *OsSCP2* promoter, and at −201 for the *OsSCP3* (*OsSCP3-a*) promoter (Fig. 4). In the case of *LAT52*, the nucleotide sequences of the pollen-specific activation element (AGAAA) were found at −1322, −694, and −326 for the *OsSCP1* (*OsSCP1-b*) promoter, at −923, −790, −620, −548, −524, and −344 for the *OsSCP2* promoter, and at −577 and −326 for the *OsSCP3* (*OsSCP3-a*) promoter (Fig. 4). The sequence motifs of the *g10* late pollen gene element (GTGA) were present at –1657, −842, and −288 for the *OsSCP1* (*OsSCP1-b*) promoter, at −294 for the *OsSCP2* promoter, and at −756 for the *OsSCP3* (*OsSCP3-a*) promoter (Fig. 4). Furthermore, the repeated sequence (TAAGGAA) was present at position −195 and −53 for the *OsSCP1* (*OsSCP1-b*), and −202 and −53 for the *OsSCP2*, though the function of this repeated sequence was unknown (Fig. 4).

Genomic organization of the *OsSCP* genes

As a next step, we surveyed the genomic organization of each *OsSCP* gene. In the case of *OsSCP2* gene, we could not find the gene duplication in the rice genome, indicating that this *OsSCP2* gene was single copy in rice genome. As previously described, we found two cDNA clones of *OsSCP2* in the public database. Combining these two data, these two cDNA clones were derived from single *OsSCP2* gene.

In contrast, in the case of *OsSCP1* and *OsSCP3* genes, we identified gene duplication in the rice genome. In *OsSCP1*, seven *OsSCP1* homologous genomic clones (*OsSCP1-a* to *OsSCP1-g*) were tandemly duplicated within the 100-kb genomic region in chromosome 2 of the rice genome. In

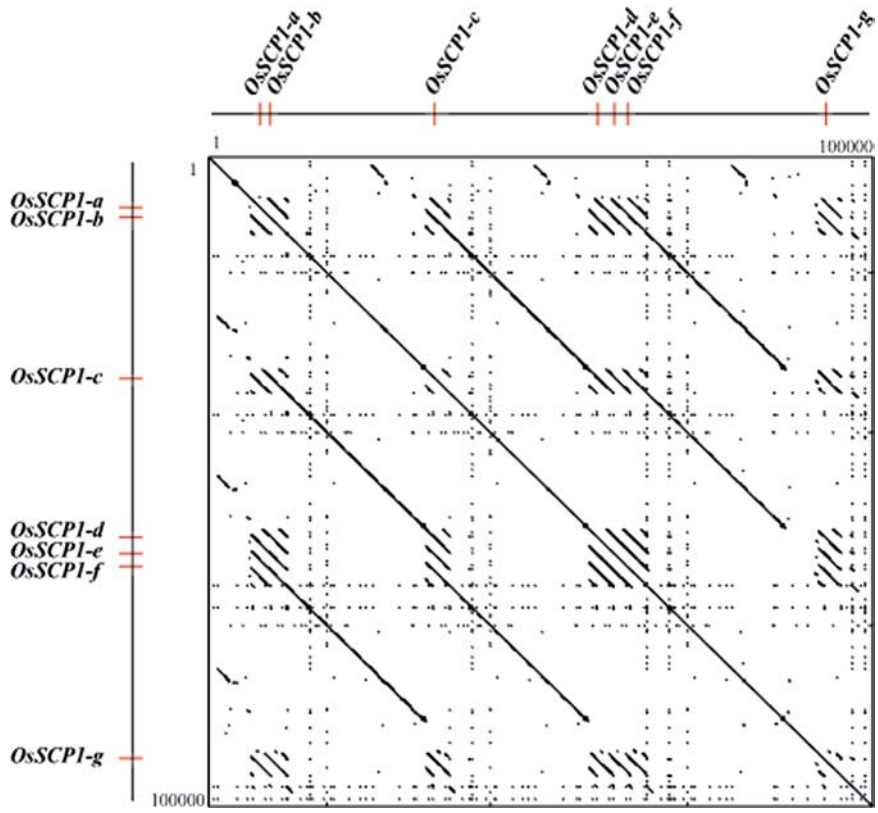

Fig. 6 Duplication in the *OsSCP1* region. Sequence similarities within 100 kb containing seven *OsSCP1* homologous sequences are shown by Huge plot, which was constructed with Genetyx software. Within the plot, each dot represents a match of successive six nucleotides

Fig. 7 Schematic structure of promoter regions in *OsSCP1* homologs (**A**) and *OsSCP3* homologs (**B**)

our experiment, we used popular type of the promoter of the genomic clone, *OsSCP1-b*, for promoter analysis, as previously described. From the Huge plot analysis, three long duplications were also detected (Fig. 6). Interestingly, the direction of the transcription of seven duplicated genomic clones was the same. Furthermore, when we compared the promoter regions (−2033 to +1) of seven homologous sequences, they were divided into four groups (Fig. 7A). The first group was *OsSCP1-a*. The second group contained three genomic clones, *OsSCP-b*, *-e*, *-f*. In these three genomic clones, the nucleotide sequence difference up to −2033 bp was not observed. The third group contained two genomic clones, *OsSCP-c*, *-d*. In these two genomic clones, the nucleotide sequence of promoter region up to −2033 bp was completely identical. The fourth group was *OsSCP1-g*. In the case of the *OsSCP1-a* promoter sequence, the region of the high sequence similarity to other groups (99%) was restricted up to −483 bp (Fig. 7A). Furthermore, several *cis*-elements, which were required for pollen-specific expression as described above, were also found in each genomic clones, though the location of the elements was different among the four groups (data not shown).

In the case of *OsSCP3*, one duplicated homologous clone (*OsSCP3-b*) was located at 3.7-kb downstream of *OsSCP3* (*OsSCP3-a*), tandemly. Although a low sequence similarity in the promoter region (−1845 to +1) was observed between *OsSCP3-a* and *OsSCP3-b*, *cis*-elements were found in both *OsSCP3-a* and *OsSCP3-b* (Fig. 7B).

Discussion

In this study, we identified and characterized novel pollen-specific genes encoding small cysteine-rich proteins

(OsSCPs). Small cysteine-rich proteins having eight conserved cysteine residues are widely distributed inside and outside the plant kingdom (Jose-Estanyol et al. 2004). These types of proteins contain the male *S* determinant of *Brassica* self-incompatibility (SP11; Takayama et al. 2000b; Watanabe et al. 2000), pollen coat protein, which can interact with SLR1 stigma-specific protein (PCP-A1, Stanchev et al. 1996; SLR1-BP, Takayama et al. 2000a), defensin peptides, which have anti-bacterial activity (Broekaert et al. 1995; Kanzaki et al. 2002; Park et al. 2002), nodule-specific cysteine-rich protein, whose functions are related to the nodule development and defense system (NCR; Mergaert et al. 2003), lipid transfer protein, which functions to transfer several different phospholipids (Zachowski et al. 1998), and pollen-specific ligand molecule, LAT52, which can interact with pollen-specific receptor kinase, LePRK2 (Twell et al. 1989; Tang et al. 2002). When aligned with these proteins at the amino acid level, the sequence and cysteine motif are highly diverted among these cysteine-rich proteins (Fig. 2), indicating that conserved cysteine motif would be important for their specific functions in each protein.

To date, several genes encoding the pollen-expressed cysteine-rich proteins (SP11, SLR1-BP, PCP-A1, PEC-3, and LAT52) have been characterized (Takayama et al. 2000a, b; Watanabe et al. 2000; Stanchev et al. 1996; Toriyama et al. 1998; Twell et al. 1989). All of these proteins had N-terminal signal peptide, indicating that these proteins should be extracellular proteins in pollen surface. As discussed above, a part of functions has been determined, and was important for pollen development and/or pollen–stigma interaction. From these data, OsSCPs would be important for reproductive process, like other pollen-expressed cysteine-rich proteins.

Gene duplication was observed in several genes in plants (Blanc et al. 2000). In the case of the plant disease resistance gene (*R* gene), genetically linked multigene families were clustered into a single locus. This *R* gene cluster should have contributed to the generation of the new *R* gene, because the nucleotide sequence diversity was observed among the duplicated homologous genes (Song et al. 1997; Dixon et al. 1998). The *Arabidopsis* genome contained gene families homologous to the pollen coat proteins (LCRs) and male *S* determinant of *Brassica* SI (SCRLs). Both gene families encoded different types of small cysteine-rich protein, and a part of the gene family formed a gene cluster (Vanoosthuyse et al. 2001). In the gene cluster, each clone was highly diverted. Although there was no experimental data showing that these clustered genes were inherited together as a unit (haplotype), these genes might have had different functions, respectively. In the *Brassica* SI, the duplication of the *SP11* was also observed. In this duplication, the nucleotide sequence of two *SP11* genes was completely identical. However, the deletion of the promoter region was found in one *SP11* gene, indicating that this clone should have been a pseudogene (Shiba et al. 2004). In this study, in *OsSCP1* and *OsSCP3*, gene duplication was observed (Fig. 6). Interestingly, the nucleotide sequences of ORF in the duplicated clones were completely conserved except for one clone in *OsSCP1*. Furthermore, pollen-specific *cis*-elements were also observed in each clone, indicating that these six duplicated clones of *OsSCP1* would have been redundant, though the precise expression pattern was determined in only the representative clone, *OsSCP1-b*. From the RT-PCR analysis, *OsSCP*s were specifically expressed in the mature anther containing trinucleate pollen grains (Fig. 3). In addition to these data, the nucleotide sequence of *OsSCP*s in two different varieties, Koshihikari and Nipponbare was completely conserved, though several nucleotide variations in other genomic regions were observed in these two varieties (Shirasawa et al. 2004). Combining the results of *cis*-elements and RT-PCR of *OsSCP*s, *OsSCP*s, which were identified and characterized in this experiment, would have been mature anther-specific genes. This high redundancy and high sequence conservation of *OsSCP*s indicate the importance of this gene in pollen maturation, pollen germination, and pollen tube growth.

In our previous study, 111 mature anther-specific genes were identified in model legume, *Lotus japonicus* (Endo et al. 2002). Out of 111 mature anther-specific genes, one interesting gene encoding receptor-like protein kinase was found. On the basis of our preliminary data, this gene was specifically expressed in mature pollen grain in *L. japonicus* (H. Masuko, M. Endo, and M. Watanabe, unpublished data). To date, two types of the pollen-specific receptor kinases (PRK1 and LePRK2) have been isolated and characterized (Mu et al. 1994; Muschietti et al. 1998). The receptor domain of LePRK2 could interact with pollen-specific small cysteine-rich protein, LAT52 (Tang et al. 2002). This pollen-specific ligand and receptor should be important for pollen germination and pollen tube elongation (Tang et al. 2002). In the rice genome, 2210 candidate genes encoding receptor-like kinases have been identified (Shiu et al. 2004). However, the functions of the receptor-like kinases have been discovered in only a few genes (Song et al. 1995; Nonomura et al. 2003; Suzaki et al. 2004). If the OsSCPs which were identified in this study were functioned as a ligand, the discovery of the orphan receptor would be important for understanding the pollination and fertilization in future.

Acknowledgements This work was supported in part by a grant from the Ministry of Agriculture, Forestry and Fisheries of Japan (Rice Genome Project MA-2211), and Grants-in-Aid for the 21 Century Center of Excellence Program from the Japan Society for Promotion of Science (JSPS) to MW. The authors are grateful to Ayako Chiba, Yukiko Ohyama, and Hiroyuki Ishikawa (Iwate University) for technical assistance. The authors also thank Ryo Tsuwamoto and Tomohiro Kakizaki (Iwate University) for providing pBI-GSH binary vector and for their helpful discussions. ME, HH, and YT are the recipients of a Research Fellowship of the JSPS for Young Scientists.

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl Acids Res 25:3389–3402

Amagai M, Ariizumi T, Endo M, Hatakeyama K, Kuwata C, Shibata D, Toriyama K, Watanabe M (2003) Identification of reproductive organ-specific genes of cruciferous model plant, *Arabidopsis thaliana*, by using a combination of *Arabidopsis* macroarray and mRNA derived from *Brassica olearacea*. Sex Plant Reprod 15:213–220

An G, Ebert PR, Mitra A, Ha SB (1988) Binary vectors. In: Gelvin SB, Schilpeoort RA (eds) Plant molecular biology manual. Kluwer, Dordrecht, pp 1–19

Bate N, Twell D (1998) Functional architecture of a late pollen promoter: pollen-specific transcription is developmentally regulated by multiple stage-specific and co-dependent activator elements. Plant Mol Biol 37:859–869

Bedinger P (1992) The remarkable biology of pollen. Plant Cell 4:879–887

Blanc G, Barakat A, Guyot R, Cooke R, Delseny M (2000) Extensive duplication and reshuffling in the *Arabidopsis* genome. Plant Cell 12:1093–1101

Broekaert WF, Terras FR, Cammue BP, Osborn RW (1995) Plant defensins: novel antimicrobial peptides as components of the host defense system. Plant Physiol 104:1353–1358

Dixon MS, Hatzixanthis K, Jones DA, Harrison K, Jones JDG (1998) The tomato *Cf-5* disease resistance gene and six homologs show pronounced allelic variation in leucine-rich repeat copy number. Plant Cell 10:1915–1925

Endo M, Matsubara H, Kokubun T, Masuko H, Takahata Y, Tsuchiya T, Fukuda H, Demura T, Watanabe M (2002) The advantages of cDNA microarray as an effective tool for identification of reproductive organ-specific genes in a model legume, *Lotus japonicus*. FEBS Lett 514:229–237

Endo M, Tsuchiya T, Saito H, Matsubara H, Hakozaki H, Masuko H, Kamada M, Higashitani A, Takahashi H, Fukuda H, Demura T, Watanabe M (2004) Identification and molecular characterization of novel anther-specific genes in *Oryza sativa* L. by using cDNA microarray. Genes Genet Syst 79:213–226

Eyal Y, Curie C, McCormick S (1995) Pollen specificity elements reside in 30 bp of the proximal promoters of two pollen-expressed genes. Plant Cell 7:373–384

Honys D, Twell D (2003) Comparative analysis of the *Arabidopsis* pollen transcriptome. Plant Physiol 132:640–652

Horsch RB, Fry J, Hoffmann N, Neidermeyer J, Rogers SG, Fraley RT (1988) Leaf disc transformation. In: Gelvin SB, Schilperoort RA (eds) Plant molecular biology manual. Kluwer, Dordrecht, pp A5/1–A5/9

Jose-Estanyol M, Gomis-Ruth FX, Puigdomenech P (2004) The eight-cysteine motif, a versatile structure in plant proteins. Plant Physiol Biochem 42:355–365

Kader J-C (1996) Lipid-transfer proteins in plants. Annu Rev Plant Physiol Plant Mol Biol 47:627–654

Kanzaki H, Nirasawa S, Saito H, Ito M, Nishihara M, Terauchi R, Nakamura I (2002) Overexpression of the wasabi defensin gene confers enhanced resistance to blast fungus (*Magnaporthe grisea*) in transgenic rice. Theor Appl Genet 105:809–814

Lush WM, Grieser F, Wolters-Arts M (1998) Directional guidance of *Nicotiana alata* pollen tubes in vitro and on the stigma. Plant Physiol 118:733–741

McCormick S (2004) Control of male gametophyte development. Plant Cell 16:S142–S153

Mergaert P, Nikovics K, Kelemen Z, Maunoury N, Vaubert D, Kondorosi A, Kondorosi E (2003) A novel family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. Plant Physiol 132:161–173

Mu JH, Lee HS, Kao Th (1994) Characterization of a pollen-expressed receptor-like kinase gene of *Petunia inflata* and the activity of its encoded kinase. Plant Cell 6:709–721

Muschietti J, Dircks L, Vancanneyt G, McCormick S (1994) LAT52 protein is essential for tomato pollen development: pollen expressing antisense *LAT52* RNA hydrates and germinates abnormally and cannot achieve fertilization. Plant J 6:321–338

Muschietti J, Eyal Y, McCormick S (1998) Pollen tube localization implies a role in pollen–pistil interactions for the tomato receptor-like protein kinase LePRK1 and LePRK2. Plant Cell 10:319–330

Nonomura K, Miyoshi K, Eiguchi M, Suzuki T, Miyao A, Hirochika H, Kurata N (2003) The *MSP1* gene is necessary to restrict the number of cells entering into male and female sporogenesis and to initiate anther wall formation in rice. Plant Cell 15:1728–1739

Okada T, Sasaki Y, Ohta R, Onozuka N, Toriyama K (2000) Expression of *Bra r* 1 gene in transgenic tobacco and *Bra r* 1 promoter activity in pollen of various plant species. Plant Cell Physiol 41:757–766

Park HC, Kang YH, Chun HJ, Koo JC, Cheong YH, Kim CY, Kim MC, Chung WS, Kim JC, Yoo JH, Koo YD, Koo SC, Lim CO, Lee SY, Cho MJ (2002) Characterization of a stamen-specific cDNA encoding a novel plant defensin in Chinese cabbage. Plant Mol Biol 50:59–69

Rogers HJ, Bate N, Combe J, Sullivan J, Sweetman J, Swan C, Lonsdale DM, Twell D (2001) Functional analysis of *cis*-regulatory elements within the promoter of the tobacco late pollen gene *g10*. Plant Mol Biol 45:577–585

Shiba H, Park JI, Suzuki G, Matsushita M, Nou IS, Isogai A, Takayama S, Watanabe M (2004) Duplicated *SP11* genes produce alternative transcripts in the S^{15} haplotype of *Brassica oleracea*. Genes Genet Syst 79:87–93

Shirasawa K, Monna L, Kishitani S, Nishio T (2004) Single nucleotide polymorphisms in randomly selected genes among *japonica* rice (*Oryza satava* L.) varieties identified by PCR-RF-SSCP. DNA Res 11:275–283

Shiu SH, Karlowski WM, Pan R, Tzeng YH, Mayer KF, Li WH (2004) Comparative analysis of the receptor-like kinase family in Arabidopsis and Rice. Plant Cell 16:1220–1234

Song WY, Wang GL, Chen LL, Kim HS, Pi LY, Holsten T, Gardner J, Wang B, Zhai WX, Zhu LH, Fauquet C, Ronald P (1995) A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. Science 270:1804–1806

Song WY, Pi LY, Wang GL, Gardner J, Holsten T, Ronald PC (1997) Evolution of the rice *Xa21* disease resistance gene family. Plant Cell 9:1279–1287

Stanchev BS, Doughty J, Scutt CP, Dickinson H, Croy RR (1996) Cloning of *PCP1*, a member of a family of pollen coat protein (PCP) genes from *Brassica oleracea* encoding novel cysteine-rich proteins involved in pollen–stigma interactions. Plant J 10:303–313

Suzaki T, Sato M, Ashikari M, Miyoshi M, Nagato Y (2004) The gene *FLORAL ORGAN NUMBER1* regulates floral meristem size in rice and encodes a leucine-rich repeat receptor kinase orthologous to *Arabidopsis* CLAVATA1. Development 131:5649–5657

Takayama S, Isogai A (2005) Self-incompatibility in plants. Annu Rev Plant Biol 56:467–489

Takayama S, Shiba H, Iwano M, Asano K, Hara M, Che F-S, Watanabe M, Hinata K, Isogai A (2000a) Isolation and characterization of pollen coat proteins of *Brassica campestris* that interact with *S*-locus related glycoprotein 1 involved in pollen–stigma adhesion. Proc Natl Acad Sci USA 97:3765–3770

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

Tang W, Ezcurra I, Muschietti J, McCormick S (2002) A cysteine-rich extracellular protein, LAT52, interacts with the extracellular domain of the pollen receptor kinase LePRK2. Plant Cell 14:227–2287

Toriyama K, Hanaoka K, Okada T, 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

Tsuchiya T, Toriyama K, Ejiri S, Hinata K (1994) Molecular characterization of rice genes specifically expressed in the anther tapetum. Plant Mol Biol 26:1737–1746

Twell D, Wing R, Yamaguchi J, McCormick S (1989) Isolation and expression of an anther-specific gene from tomato. Mol Gen Genet 217:240–245

Twell D, Yamaguchi J, McCormick S (1990) Pollen-specific gene expression in transgenic plants: coordinate regulation of two different tomato gene promoters during microsporogenesis. Development 109:705–713

Vanoosthuyse V, Miege C, Dumas C, Cock JM (2001) Two large *Arabidopsis thaliana* gene families are homologous to the *Brassica* gene superfamily that encodes pollen coat proteins and the male component of the self-incompatibility response. Plant Mol Biol 16:17–34

Watanabe M, Shiozawa H, Isogai A, Suzuki A, Takeuchi T, Hinata K (1991) Existence of *S*-glycoprotein-like proteins in anthers of self-incompatible species of *Brassica*. Plant Cell Physiol 32:1039–1047

Watanabe M, Suzuki G, Toriyama K, Takayama S, Isogai A, Hinata K (1999) Two anther-specific 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*). Sex Plant Reprod 12:127–134

Watanabe M, Ito A, Takada Y, Ninomiya C, Kakizaki T, Takahata Y, Hatakeyama K, Hinata K, Suzuki G, Takasaki T, Satta Y, Shiba H, Takayama S, Isogai A (2000) Highly divergent sequences of the pollen self-incompatibility (*S*) gene in class-I *S* haplotypes of *Brassica campestris* (syn. *rapa*) L. FEBS Lett 473:139–144

Watanabe M, Takayama S, Isogai A, Hinata K (2003) Recent progresses on self-incompatibility research in *Brassica* species. Breed Sci 53:199–208

Yazaki J, Kojima K, Suzuki K, Kishimoto N, Kikuchi S (2004) The Rice PIPELINE: a unification tool for plant functional genomics. Nucl Acids Res 32:D383–D387

Zachowski A, Guerbette F, Grosbios M, Jolliot-Croquin A, Kader JC (1998) Characterization of acyl-binding by a plant lipid-transfer protein. Eur J Biochem 257:443–448