

Analysis of expressed sequence tags from *Cryptomeria japonica* pollen reveals novel pollen-specific transcripts

NORIHIRO FUTAMURA,^{1,2} TOKUKO UJINO-IHARA,³ MITSURU NISHIGUCHI,¹ HIROYUKI KANAMORI,⁴ KENSUKE YOSHIMURA,³ MASAHIRO SAKAGUCHI⁵ and KENJI SHINOHARA¹

¹ Department of Molecular and Cell Biology, Forestry and Forest Products Research Institute, Tsukuba, Ibaraki 305-8687, Japan

² Corresponding author (futa@ffpri.affrc.go.jp)

³ Department of Forest Genetics, Forestry and Forest Products Research Institute, Tsukuba, Ibaraki 305-8687, Japan

⁴ Rice Genome Research Program, Institute of the Society for Techno-Innovation of Agriculture, Forestry and Fisheries, Tsukuba, Ibaraki 305-0854, Japan

⁵ Research Center for Allergy and Immunology, RIKEN Yokohama Institute, Yokohama, Kanagawa 230-0045, Japan

Received July 20, 2005; accepted March 1, 2006; published online September 1, 2006

Summary *Cryptomeria japonica* D. Don is one of the most important forest trees in Japan, but more than 10% of the Japanese population is allergic to its pollen. We constructed a cDNA library derived from pollen grains of *C. japonica* and performed an analysis of expressed sequence tags (ESTs). We obtained partial sequences from 1929 clones, which represented 1365 unique transcripts. Among the unique transcripts, 984 (72%) encoded proteins that were similar to *Arabidopsis* proteins with *E*-values of $< 10^{-5}$. Analysis of functional composition of the pollen ESTs revealed the overrepresentation of mRNAs for proteins involved in protein synthesis and post-translational modification. The most abundant transcripts were derived from novel genes (*CjMPI*-related genes) and encoded proteins that were not homologous to any proteins in current databases. The *CjMPI*-related genes formed a multi-gene family and were expressed specifically in the pollen grains of *C. japonica*. An analysis of homologies between ESTs from *C. japonica* pollen and proteins in the Structural Database of Allergenic Proteins revealed that products of 48 of the clones (2.5%) exhibited significant homology to known plant allergens. Our results provide new information about pollen-specific genes and potential allergens in *C. japonica* pollen.

Keywords: allergen, gene expression, sugi, transcriptome.

Introduction

Pollen has been studied not only because of its importance as the male partner in sexual reproduction, but also because it serves as a good model for investigations of cell growth and morphogenesis. A hybridization kinetics analysis of [³H]-cDNA with poly(A)+ RNA revealed that between 20,000 and 24,000 different mRNA sequences are present in pollen of *Tradescantia paludosa* E.S. Anderson & Woods. and *Zea mays* L. (Willing and Mascarenhas 1984, Willing et al.

1988). Of the estimated 20,000 genes that are expressed, approximately 10% are considered to be pollen-specific. In the 1990s, more than 150 pollen-expressed genes, belonging to more than 50 distinct classes, were isolated from over 28 species (Twell 2002). In 2003, two groups reported the results of comprehensive analyses of the pollen transcriptome using *Arabidopsis* GeneChip arrays (Affymetrix, Santa Clara, CA) prepared with oligonucleotide probes that represented approximately 8000 genes. One group identified 992 pollen-expressed mRNAs, nearly 40% (387 genes) of which were detected specifically in pollen (Honys and Twell 2003). The other group detected 1584 genes that were expressed in pollen and reported that 10% (162 genes) were selectively expressed in pollen (Becker et al. 2003). Recently, larger-scale transcriptome analyses were performed with the ATH1 Genome Array, which covers more than 80% of the *Arabidopsis* genome (Honys and Twell 2004, Pina et al. 2005). The cited authors identified 6587 and 7235 genes that were expressed in mature pollen grains, respectively. The global pattern of gene expression in *Arabidopsis* pollen was also characterized by the Serial Analysis of Gene Expression (Lee and Lee 2003) method. A total of 4211 unique tags were obtained from *Arabidopsis* pollen, 45% (1877 tags) of which matched proteins in the UniGene database. These analyses of the transcriptome were performed in angiosperms, particularly *Arabidopsis*, and information on gymnosperm pollen remains limited.

Pollen is known as an airborne allergen (Knox and Suphioglu 1996). Sugi (*Cryptomeria japonica* D. Don) is one of the most commercially important conifers in Japan because of its high utility and productivity. However, the allergic reactions caused by its pollen are a severe public-health problem in Japan. The existence of numerous allergens in *C. japonica* pollen has been suggested, but only a few antigens have been identified (Fujimura et al. 2004). The cDNAs of four allergens that have been characterized in *C. japonica* pollen, namely,

Cry j 1, Cry j 2, CJP-4 and CJP-6, have been cloned (Komiya et al. 1994, Namba et al. 1994, Sone et al. 1994, Kawamoto et al. 2002, Fujimura et al. 2005). Cry j 1 and Cry j 2 were shown to have pectate lyase and polymethylgalacturonase activity, respectively, and CJP-4 and CJP-6 are homologous to a Class IV chitinase and an isoflavone reductase, respectively (Ohtsuki et al. 1995, Taniguchi et al. 1995, Kawamoto et al. 2002, Fujimura et al. 2005). However, comprehensive characterization of other pollen allergens in *C. japonica* remains to be performed. To identify the pollen-expressed genes in *C. japonica* and to identify novel candidate genes for pollen allergens, we characterized 3655 expressed sequence tags (ESTs) from a total of 1929 cDNA clones that were randomly selected from a cDNA library of *C. japonica* pollen. We found that several amino acid sequences deduced from the pollen cDNAs were homologous to known sequences in plant allergens. In addition, we found pollen-abundant transcripts derived from a novel type of gene and characterized the patterns of expression and the distribution of these genes in the *C. japonica* genome.

Materials and methods

Plant materials and construction of the cDNA library

All plant materials and the pollen cDNA library were prepared as described previously (Futamura et al. 2002b). Total RNA was isolated from mature pollen grains as described by Sone et al. (1994). The cDNA library in λ ZAP II (Stratagene, La Jolla, CA) was excised in plasmid form [pBluescript SK(-); Stratagene] by coinfection with ExAssist helper phage according to the instructions from Stratagene.

Sequencing of ESTs and data processing

Plasmid DNAs were obtained from overnight cultures by the alkaline lysis method (Sambrook et al. 1989) and purified with MultiScreen NA and MultiScreen FB (Millipore, Bedford, MA). The resultant plasmids were sequenced by automatic sequencers according to the manufacturer's instructions (ABI 3700; Applied Biosystems, Foster City, CA). The upstream M13 reverse primer (5'-AGCGGATAACAATTTTCACACAGG-3') was used for forward sequencing. The clones with good-quality sequences were selected for reverse sequencing and the T7 primer (5'-TAATACGACTCACTATAGGG-3') was used for reverse sequencing. Clones from which high-quality sequences had not been determined because of the presence of long polyA tails were selected for sequencing with a mixture of dT-anchored primers (5'-TTTTTTTTTTTTTTT-TTTTTV-3', where V = A, C and G).

Raw sequence-trace files were processed by the Phred program (Ewing et al. 1998). Low-quality sequences, as well as vector and linker sequences, were removed by an in-house program. Any resultant short (< 60 bp) sequences were eliminated from subsequent analysis. Contaminating regions of the genomes of *Escherichia coli* and bacteriophage λ were identified with BLASTN, and those with a score > 100 or an *E*-value < 10^{-10} were excluded. The remaining sequences were used for further analysis.

Clustering was performed with the CAP3 program (Huang and Madan 1999) with an overlap-length cutoff of 40, an overlap percent-identity cutoff of 95 and other parameters set at default values.

Functional classification and annotation of ESTs

The ESTs were categorized on the basis of the putative functions of encoded products using the eukaryotic orthologous groups of proteins (COG) database (Tatusov et al. 2003). The results of BLASTX analysis of both 5'- and 3'-end sequences of each clone were compared, and the functional category with the higher score was adopted.

The assembled sequences were first annotated on the basis of a BLASTX homology search. As our databases, we used RefSeq of *Arabidopsis thaliana* (L.) Heynh from the National Center for Biotechnology Information (NCBI, Pruitt et al. 2005), the database of full-length cDNA clones from the knowledge-based Oryza Molecular Biological Encyclopedia (Kikuchi et al. 2003) and the non-redundant protein (NRP) database with entries from GenPept, Swissprot, PIR and PDB.

Similarities to ESTs from *Pinus*, *Picea* and *Populus* and to ESTs from the inner bark of *C. japonica* were determined with the TBLASTX program, using *Pinus* Gene Index Release 6.0, Spruce Gene Index Release 1.0, Poplar Gene Index Release 2.0 from the TIGR Gene Indices (Lee et al. 2005) and the ESTs collected from inner bark of *C. japonica* (Ujino-Ihara et al. 2000).

Similarities to known plant allergens were determined by BLASTX comparison with proteins in SDAP (Structural Database of Allergenic Proteins). To confirm the nature of abundant transcripts, we grouped assembled sequences with BLASTN scores of more than 250 and compared them with known amino acid sequences in a publicly accessible database (NRP database).

Reverse transcription and PCR (RT-PCR)

Before reverse transcription, we treated total RNA with RQ1 RNase-free DNase (Promega, Madison, WI) to remove contaminant DNA. First-strand cDNA was synthesized from 1 μ g of the RNA using the SuperScript III first-strand synthesis system (Invitrogen, Carlsbad, CA) with an oligo(dT) primer. The reaction mixture for PCR contained 1 μ l (1/21 of the original volume) of the solution of the first-strand cDNA, 12.5 μ l of GoTaq Green Master Mix (Promega) and 0.4 μ M of gene-specific primers (Table 1), in a total volume of 25 μ l. Amplification with all primers except for those for the cDNA for cucumisin-like serine protease and chlorophyll a/b-binding protein was performed with incubations as follows: 90 s at 95 °C; and then 35 cycles of 30 s at 95 °C, 30 s at 56 °C and 30 s at 72 °C. Conditions with the primers for the cDNA for cucumisin-like protease were as follows: 90 s at 95 °C; and then 35 cycles of 30 s at 95 °C, 30 s at 54 °C and 30 s at 72 °C. Conditions with the primers for the cDNA for chlorophyll a/b-binding protein were as follows: 90 s at 95 °C; and then 26 cycles of 30 s at 95 °C, 30 s at 56 °C and 30 s at 72 °C. Five μ l of PCR product solution were fractionated on a 2.0% agarose gel and cDNAs were stained with ethidium bromide.

Table 1. Primers used for analysis by RT-PCR.

Tentative designation	Forward primer	Reverse primer
Oxidoreductase-like protein	5'-AATGGATTACGCCACCAGAG-3'	5'-CCTGCTTTGGTCTTCAGCTC-3'
Peptidyl-prolyl <i>cis-trans</i> isomerase	5'-GCTCCAGTTCTTCATCTGC-3'	5'-TTGACCAAACCTAGGCAAC-3'
Anther-specific protein SF18 precursor	5'-ATTCACGAGATGGCCAAGAG-3'	5'-TTCTCCATTTGCATTTTCC-3'
Photoassimilate-responsive protein (putative)	5'-AATGGAATGGATTGAGAGCG-3'	5'-AGTCCAGCAAGGGAAAAGGT-3'
40S Ribosomal protein S2 (putative)	5'-GTGTAAGGTCACGGGCAAGT-3'	5'-AAGGCGATTTTGAGAAACGA-3'
60S Acidic ribosomal protein P2	5'-TTTGACAGTAGGGAGCAGCTT-3'	5'-CTCAAACGCGAACTGCATTA-3'
Nuclear RNA binding protein A (putative)	5'-AGCTGAAAGGGAAGAGAGGG-3'	5'-CCTACGGGCACCTGATTA-3'
Arginine decarboxylase	5'-CAAGGCTGTCGACTTGTTCA-3'	5'-AGAGCTCTGGTAATGGCGAA-3'
Cucumisin-like serine protease	5'-AGCGTACAAGGTTTGTGGG-3'	5'-TGACAAGCTAACCGCATCTG-3'
Polyubiquitin	5'-TGTTTGCCAGCAAAGATCAG-3'	5'-GTCTCAGGGGAGGTATGCAA-3'
Similar to desiccation-related protein	5'-CACTGGGAGCAGAAGGAAAG-3'	5'-TTTTCATCCGATTTTGACA-3'
Cysteine proteinase inhibitor	5'-CTGCAGGAGATTCTTTTGGC-3'	5'-TGCATGCATACAGTGCTTCA-3'
Nucleolin (putative)	5'-AGGCAATCATCTGGTCCAAC-3'	5'-CCATCACCACCTCCACTTCT-3'
60S Acidic ribosomal protein P3	5'-GAAGGATCGGGAGTTGTTCA-3'	5'-TTCTGAGTGAGGCCCATTTCT-3'
Oleosin	5'-GTGATGGTGACTCCCGTTCT-3'	5'-TGCACAGCTCCAGCATAATC-3'
RTN/Nogo gene	5'-CTCTTTTGGCTGCAGGATTC-3'	5'-ACTCTTTGGACACCAATGGC-3'
Delta-8 sphingolipid desaturase	5'-TCACGCTTCTGTTTGACTG-3'	5'-CAGAATGTGGCGCTCTTGTA-3'
Hexose transporter-like protein	5'-AGCGTAGAAACAGGCCTCAA-3'	5'-CATGCAAAGGCTGAGACAAA-3'
60S Acidic ribosomal protein P1a	5'-GAAGAGAAGCGTTGACGACC-3'	5'-GTACAGCAAAAACGCTGCAA-3'
Chlorophyll a/b-binding protein	5'-TCAGCCAGACCCAGAGACTT-3'	5'-GTTCTCCAGGGACCTTTTC-3'

DNA and RNA gel blot analysis

Total RNA was isolated from pollen grains as described by Sone et al. (1994) and from other plant organs as described by Shinohara and Murakami (1996). Genomic DNA was extracted from current needles as described by Wagner et al. (1987). DNA and RNA gel blot analyses were performed as described previously (Futamura et al. 2002a). The genomic DNA was digested separately with *EcoR* I, *EcoR* V, *Hind* III and *Bam*HI.

To generate the entire cDNA probe for the short type of *CjMPI*, we labeled the cDNA insert with [α -³²P]dCTP using the Megaprime DNA labeling system according to the manufacturer's instructions (Amersham Biosciences, Piscataway, NJ). For the long-type-specific sequence of *CjMPI*, we amplified the insert of the long type of *CjMPI* by PCR using the sense primer 5'-CACTTCTCACCAGATGAGAA-3' and the antisense primer 5'-GCTCCGTCGATAAGGAATG-3'. We labeled each product of PCR with the Megaprime DNA labeling system using a mixture of random nonameric primers and two sequence-specific nonameric primers (5'-CCGTCGATA-3' and 5'-CGATAAGGA-3').

Production of *CjMPI*-specific antiserum

The synthetic peptide, CEEVEHSTPQKSVEK, was conjugated to keyhole limpet hemocyanin as carrier protein. Two rabbits were immunized with the conjugated protein eight times at 2-week intervals. Whole blood was collected 18 weeks after the first injection. Serum was obtained by centrifugation of clotted blood.

Production of recombinant *CjMPI* fusion protein in *E. coli*

A cDNA fragment encoding *CjMPI* (the cDNA Clone CP10-460) was amplified by PCR with the primers 5'-CCGGAATTCATGGGGAAATCTATGGTCATCA-3' and 5'-CCGCTCG-

AGTTAAGCCGTTTTTTCATCTGCTGA-3'. After digestion with *EcoR* I and *Xho* I, the fragment was inserted into pGEX-4T-1 (Amersham Biosciences) in-frame with cDNA for glutathione-S-transferase (GST). The resulting expression plasmid was sequenced and then introduced into the *E. coli* Rosetta-gami B strain (Novagen, Madison, WI).

Cells harboring the expression plasmid were grown in 500 ml of 2 × YT medium (Sambrook et al. 1989) with ampicillin (50 µg ml⁻¹), chloramphenicol (34 µg ml⁻¹) and 2% glucose at 37 °C for 5 h. Isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.4 mM to induce expression of the fusion protein. After further incubation at 37 °C for 3 h, cells were harvested by centrifugation. Cell pellets were suspended in 12.5 ml of CellLytic BII buffer (Sigma-Aldrich, St. Louis, MO) that contained 1 mM EDTA and 1 × complete proteinase inhibitor cocktail (Roche Diagnostics, Mannheim, Germany) and disrupted by sonication on ice. The cell lysate was centrifuged at 25,000 g for 15 min at 4 °C. The clarified supernatant was loaded onto a GSTrap FF column (Amersham Biosciences), which had been equilibrated in buffer A (50 mM Tris-Cl (pH 7.5) and 0.3 M NaCl). The column was washed with 20 ml of buffer A. Bound proteins were eluted with a mixture (1:1, v/v) of buffer A and buffer B (50 mM Tris-HCl (pH 8.0), 0.3 M NaCl and 20 mM reduced glutathione). The pooled fractions of interest were dialyzed against PBS (10 mM Na₂HPO₄, 2 mM KH₂PO₄, 137 mM NaCl and 2.7 mM KCl).

Immunochemical procedures

A whole-cell lysate and the cytoplasmic fraction of *E. coli* and purified recombinant GST-*CjMPI* were subjected to sodium dodecyl sulphate (12%) polyacrylamide gel electrophoresis (SDS-PAGE). Proteins were transferred to a PVDF membrane

(Millipore) by electroblotting. Immunoblot analysis was performed with a 1:3750 dilution of the CjMP1-specific antiserum and ECL Plus Western Blotting Detection Reagents (Amersham Biosciences).

The CjMP1 protein in a crude extract of *C. japonica* pollen was detected by ELISA with the CjMP1-specific antiserum according to a previously described procedure (Takahashi-Omoe et al. 2004). The allergenicity of the CjMP1 protein was examined by fluorometric ELISA as described previously (Sakaguchi et al. 1990). IgE antibodies in the sera of 19 patients with *C. japonica* pollinosis were used to examine binding of human antibodies to GST-CjMP1 and to proteins in the crude extract of *C. japonica* pollen.

Results

Sequencing and assembly of ESTs

We picked clones at random from the cDNA library generated from mature pollen grains of *C. japonica* and sequenced both the 5'- and 3'-ends of the cDNA sequences using the upstream M13 reverse primer, the T7 primer and a mixture of dT anchor primers (see Materials and methods). After elimination of inappropriate EST sequences, we obtained both 5'- and 3'-end sequences from 1726 clones, and either the 5'- or the 3'-end sequence from 203 clones. The average read length of the ESTs was 527 bp. A total of 3655 ESTs were deposited in the DDBJ/EMBL/GenBank databases (Accession Nos. BJ9366-38–BJ940292).

Clustering analysis based on the CAP3 algorithm revealed that the 3655 ESTs could be assembled into 1838 transcripts that consisted of 696 contigs and 1142 singlets. Of the 1142 singlets, 946 were derived from both ends of individual clones. As a result, we obtained 1365 transcripts (696 contigs, 473 non-assembled clones in which both sides were singlets and 196 clones in which one side was a singlet).

Classification of ESTs by putative product functions

We classified all 1929 cDNA clones and 1365 transcripts functionally on the basis of their assignments in the COG database. The COG database contains proteins encoded by seven eukaryotic genomes: three from animals (*Caenorhabditis elegans*, *Drosophila melanogaster* and *Homo sapiens*); one from a plant (*Arabidopsis thaliana*); two from fungi (*Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*); and one from an intracellular microsporidian parasite (*Encephalitozoon cuniculi*) (Tatusov et al. 2003). Of the cDNA clones and transcripts, 53% (1024 clones) of the former and 50% (686 transcripts) of the latter were assigned to 23 putative functional categories by BLASTX (E -value $< 10^{-5}$; Figure 1). In both cases, genes for "translation, ribosomal structure and biogenesis" and "post-translational modification" were abundant in *C. japonica* pollen libraries. None of the ESTs were included in the category of "extracellular structures."

Designation of ESTs

We compared the sequences deduced from the 1365 transcripts with known amino acid sequences using BLASTX al-

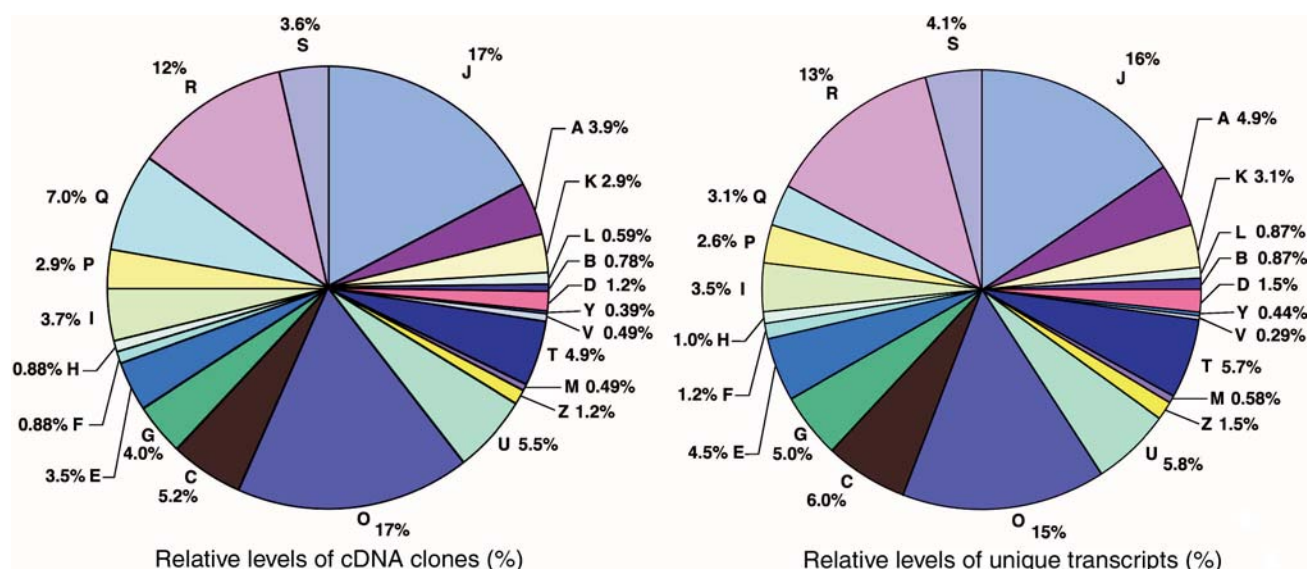


Figure 1. Functional classification and relative quantities (as percentages) of cDNA clones and unique transcripts in the pool of ESTs derived from *C. japonica* pollen. We assigned 1024 cDNA clones (left) and 686 transcripts (right) to COGs using a BLAST-based algorithm (E -value $< 10^{-5}$). Designations of functional categories and percentages: (A) RNA processing and modification; (B) chromatin structure and dynamics; (C) energy production and conversion; (D) cell cycle control and mitosis; (E) amino acid transport and metabolism; (F) nucleotide transport and metabolism; (G) carbohydrate transport and metabolism; (H) coenzyme transport and metabolism; (I) lipid transport and metabolism; (J) translation, ribosomal structure and biogenesis; (K) transcription; (L) replication and repair; (M) cell wall/membrane/envelope biogenesis; (O) post-translational modification, protein turnover and chaperone functions; (P) inorganic ion transport and metabolism; (Q) secondary metabolites biosynthesis, transport and catabolism; (R) prediction of general function only; (S) function unknown; (T) signal transduction; (U) intracellular trafficking, secretion and vesicular transport; (V) defense mechanisms; (Y) nuclear structure; and (Z) cytoskeleton.

gorithms. We found that 1090 (79.9%) of the 1365 transcripts encoded peptides that were significantly similar to known amino acid sequences, 984 (72.1%) were similar to *Arabidopsis* proteins and 949 (69.5%) were similar to proteins in japonica rice (*Oryza sativa* L.) at a BLASTX *E*-value of 10^{-5} . When compared with sequences of ESTs from *Pinus*, *Picea* and *Populus* by TBLASTX, 1101 (80.7%), 1008 (73.8%) and 1002 (73.4%) transcripts were similar, respectively, at *E*-values of 10^{-5} . A larger overlap was found in the case of coniferous species. When we compared the ESTs from pollen grains with sequences of transcripts from the inner bark of *C. japonica* (Ujino-Ihara et al. 2000), we found that 690 (50.5%) transcripts were similar at a TBLASTX *E*-value of 10^{-10} . These results suggested that the ESTs from pollen grains contained information about pollen-specific transcripts.

We grouped the assembled sequences that represented similar genes on the basis of BLASTN scores > 250, and then compared them to known amino acid sequences. The abundant transcripts represented by the pollen cDNA library are listed in Table 2. We excluded ESTs that were similar to sequences of mitochondrial and chloroplastic DNA and rDNA sequences. The sequence encoded by the most redundant EST was different from all known amino acid sequences. Peptides encoded

by four clusters of the 20 most abundant transcripts exhibited strong similarity to several ribosomal proteins. Three of them are classified as “translation, ribosomal structure and biogenesis proteins” (Figure 1). Two clusters of the 20 most abundant transcripts (encoding peptidyl-prolyl *cis-trans* isomerase and polyubiquitin) are classified as “post-translational modification proteins.” These observations may explain why the categories of “translational, ribosomal structure and biogenesis proteins” and “post-translational modification proteins” account for large fractions of the ESTs generated from *C. japonica* pollen (Figure 1).

To determine whether expression of the enriched transcripts was pollen-specific, we examined the patterns of gene expression associated with the 19 most redundant ESTs in each organ of *C. japonica* by semi-quantitative reverse transcriptase polymerase chain reaction (RT-PCR, Figure 2). We detected the expression of all the genes in pollen and in developed male flowers that included mature pollen grains. These genes could be divided into two groups according to their patterns of expression. The first group contained four genes, which encoded an oxidoreductase-like protein, SF18 protein, a desiccation-related protein and oleosin, and these genes were expressed only in pollen and developed male flowers. The second group con-

Table 2. Putative functions of the products of the 20 most abundant transcripts in pollen.¹

No. of clones (redundancy)	Accession No. (top BLASTX hit)	Tentative designation (top BLASTX hit)	AGI code (top hit in <i>Arabidopsis</i> genome)	<i>E</i> -value ²	Classification by COG
91		No hit ³			No hit
51	AAG52269	Oxidoreductase-like protein (<i>Arabidopsis thaliana</i>)	At1g80320 ⁴	5E-30	Q
46	Q39613	Peptidyl-prolyl <i>cis-trans</i> isomerase (<i>Catharanthus roseus</i> (L.) G. Don)	At2g21130	1E-75	O
28	P22357	Anther-specific protein SF18 precursor (<i>Helianthus annuus</i> L.)	At2g26010	2E-06	No hit
18	AY091116	Photoassimilate-responsive protein (putative) (<i>A. thaliana</i>)	At5g52390 ⁴	7E-38	No hit
15	AAP21434	40S Ribosomal protein S2 (putative) (<i>Oryza sativa</i>)	At3g57490 ⁴	3E-83	J
13	AY081851	60S Acidic ribosomal protein P2 (<i>Prunus dulcis</i> (Mill.) D.A. Webb)	At2g27720 ⁴	1E-10	J
10	AAT38003	Nuclear RNA binding protein A (putative) (<i>O. sativa</i>)	At4g17520 ⁴	2E-16	R
10	AJ575747	Arginine decarboxylase (<i>Lotus corniculatus</i> var. <i>japonicus</i> L.)	At2g16500	4E-83	No hit
10	O65351	Cucumisin-like serine protease (<i>A. thaliana</i>)	At5g67360 ⁴	2E-73	No hit
10	X98063	Polyubiquitin (<i>Pinus sylvestris</i> L.)	At5g20620	1E-113	O
9	AY087598	Similar to desiccation-related protein (<i>A. thaliana</i>)	At1g47980 ⁵	2E-55	No hit
8	T07139	Cysteine proteinase inhibitor (<i>Glycine max</i> L.)	At3g12490 ⁴	5E-50	No hit
8	AY644645	Nucleolin (putative) (<i>O. sativa</i>)	At1g48920 ⁴	1E-23	K
7	T02037	60S Acidic ribosomal protein P3 (<i>Zea mays</i>)	At4g25890 ⁴	5E-21	No hit
7	AF466103	Oleosin (<i>Theobroma cacao</i> L.)	At2g25890 ⁵	1E-12	No hit
6	AY164838	RTN/Nogo gene (<i>Cryptomeria japonica</i>)	At4g23630	9E-41	U
6	AJ224160	Delta-8 sphingolipid desaturase (<i>Brassica napus</i> L.)	At2g46210	2E-65	I
6	AY052692	Hexose transporter-like protein (<i>A. thaliana</i>)	At5g26340 ⁴	8E-89	R
6	BAD33092	60S Acidic ribosomal protein P1a (<i>O. sativa</i>)	At5g24510 ⁴	8E-23	J

¹ ESTs with similarity to mitochondrial and chloroplastic DNA sequences and rRNA sequences (BLASTN score > 80) are not included.

² The *E*-value of an EST that exhibited strongest similarity to the top hit *Arabidopsis* protein.

³ There were no reported amino acid sequences with BLASTX scores above 50. Proteins had BLASTX scores > 280 when compared with each other, and they were considered to be the products of transcripts derived from *CjMPI*-related genes.

⁴ Genes that were not expressed or whose expression was suppressed in pollen according to two studies with ATH1 Genome Arrays (Honys and Twell 2004, Pina et al. 2005).

⁵ Genes that were selectively expressed in pollen and were not expressed in vegetative tissues (cotyledon, leaves, petiole, stems and roots) according to results in *A. thaliana* (Honys and Twell 2004).

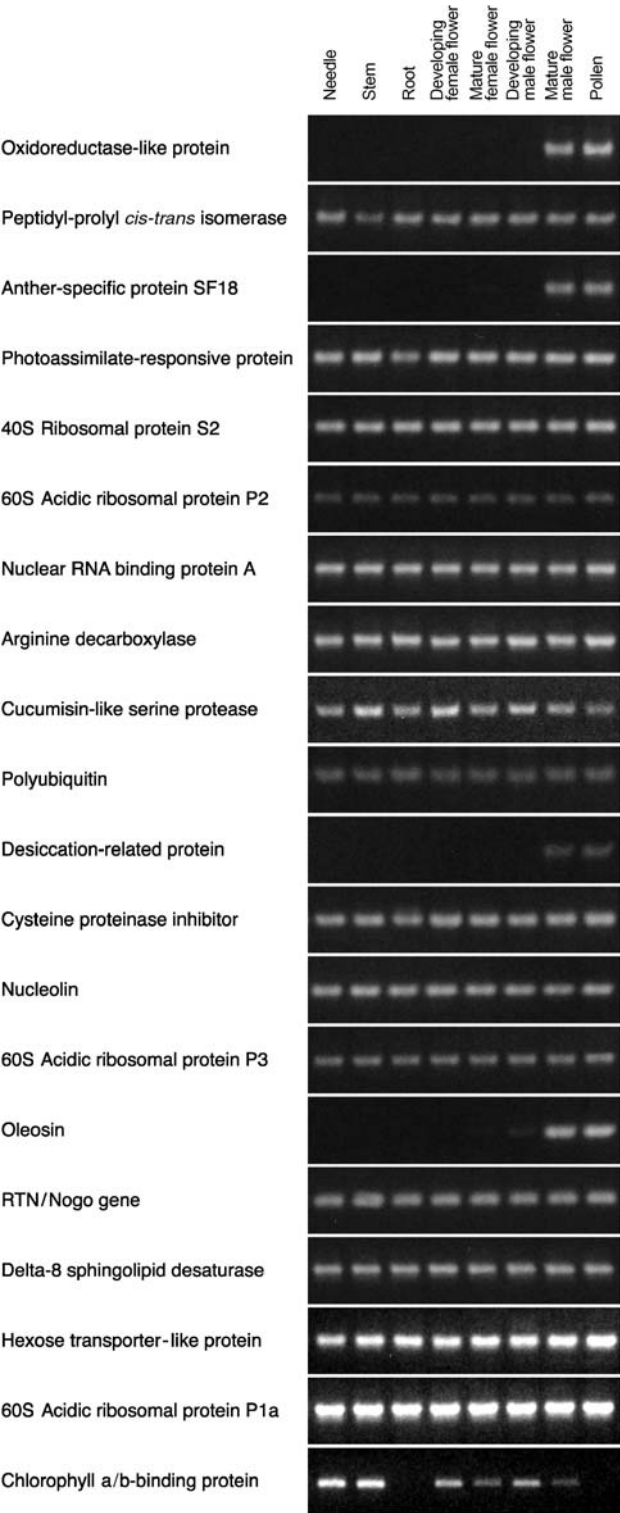


Figure 2. Expression of abundant transcripts in various organs of *C. japonica*. The expression of the 20 most abundant transcripts, with the exception of the most abundant transcript, was analyzed by RT-PCR. Chlorophyll a/b-binding protein was a negative control.

tained all the remaining genes, which were expressed in all organs examined. The purity of the pollen population of transcripts was confirmed by the absence of gene expression for

the light-harvesting chlorophyll a/b-binding protein in pollen grains (Mukai et al. 1998).

Similarity of EST products to allergenic proteins from plants

We compared the peptides encoded by all the ESTs with known plant allergens in SDAP (Ivanciuc et al. 2003) in an attempt to identify novel candidate allergens in *C. japonica* pollen. Proteins encoded by 48 cDNAs (2.5% of total cDNAs) exhibited partial sequence homology to plant allergens by BLASTX analysis. Table 3 shows a list of plant allergens that are similar to proteins encoded by ESTs derived from *C. japonica* pollen. More than a few deduced proteins were similar to known allergens of plant origin, which included the two major allergens of *C. japonica* pollen, Cry j 1 and Cry j 2 (Namba et al. 1994, Sone et al. 1994, Kawamoto et al. 2002). These newly deduced proteins in *C. japonica* were potential candidates for novel plant allergens.

Characterization of the most abundant transcripts in *C. japonica* pollen

The deduced products of the most abundant transcripts in *C. japonica* pollen exhibited no homology to any known amino acid sequences (Table 2). The corresponding genes were designated *CjMP1*-related (*Cryptomeria japonica* transcript induced in mature pollen-related) genes. Figure 3 shows representative amino acid sequences deduced from sequences of *CjMP1*-related genes that included entire open reading frames. The *CjMP1*-related genes encoded a variety of proteins with strong homology to one another. The predicted amino acid sequences could be roughly divided into short types, consisting of 91, 92 or 99 amino acids, and long types, consisting of 146 or 147 amino acids, including 49 or 50 additional amino acids near the carboxyl terminus.

The predicted proteins of the short type have a calculated mass of approximately 11 kDa, whereas those of the long type have a predicted mass of approximately 17 kDa. The SignalP algorithm indicated that the *CjMP1*-related proteins had a putative amino-terminal signal peptide. The proposed site of cleavage in *CjMP1*-related proteins is located between Ala and Lys residues (Figure 3). A computational domain search using InterPro revealed that the region between the amino terminus and the cleavage site corresponds to a transmembrane region. No other functional domains were found in the *CjMP1*-related proteins.

We performed gel blot analysis with genomic DNA to estimate the number of *CjMP1*-related genes in the genome of *C. japonica* (Figure 4). When we used the entire cDNA of a short type of *CjMP1*-related gene, we obtained more than 15 signals (Figure 4A). We assumed that this probe would hybridize to both the short and the long type of *CjMP1*-related genes. When we used a probe specific to the long type of *CjMP1*-related genes, we detected 5–8 signals (Figure 4B), indicating that the *C. japonica* genome contains multiple genes that correspond to the short type and long type of *CjMP1*-related genes and form a multiple-gene family.

We examined the organ-specific expression of the *CjMP1*-

Table 3. Products of ESTs that resemble plant allergens.

Accession No. ¹	Allergen	Species	Type of allergen	Putative product	E-value ²	No. ³
BJ939937	Cry j 1	<i>Cryptomeria japonica</i> (sugi)	Tree pollen	Pectate lyase	1E-119	4
BJ937674	Cry j 2	<i>C. japonica</i>	Tree pollen	Polymethylgalacturonase	2E-115	5
BJ940078	Jun o 4	<i>Juniperus oxycedrus</i> L. (prickly juniper)	Tree pollen	Calmodulin	1E-17	2
BJ937871	Bet v 6.0102	<i>Betula verrucosa</i> Ehrh. (birch)	Tree pollen	Isoflavone reductase-like protein	9E-64	1
BJ937965	Aln g 1	<i>Alnus glutinosa</i> (L.) Gaertn. (alder)	Tree pollen	Pathogenesis related protein PR10	2E-08	1
BJ937056	Che a 1	<i>Chenopodium album</i> L. (lamb's-quarters)	Grass pollen	LAT52 protein	3E-28	2
BJ940275	Act c 1	<i>Actinidia chinensis</i> Planch. (kiwi)	Food	Cysteine protease	1E-21	2
BJ938033	Cap a 1w	<i>Capsicum annuum</i> L. (bell pepper)	Food	Thaumatococin-like protein (PR5)	5E-46	1
BJ937370	Cor a 10	<i>Corylus avellana</i> L. (hazel)	Food	Luminal binding protein	3E-102	2
BJ939463	Cuc m 1	<i>Cucumis melo</i> L. (muskmelon)	Food	Serine protease	4E-48	15
BJ937533	Fag e 1	<i>Fagopyrum esculentum</i> Moench (buckwheat)	Food	13S globulin seed storage protein 3	10E-15	2
BJ938907	Mal d 2	<i>Malus pumila</i> Mill. (apple)	Food	Thaumatococin-like protein (PR5)	6E-63	1
BJ937760	Ory s 1	<i>Oryza sativa</i> (rice)	Food	Expansin	1E-33	3
BJ936819	Pru p 4.0201	<i>Prunus persica</i> (L.) Batsch (peach)	Food	Profilin	1E-62	1
BJ940097	Hev b 2	<i>Hevea brasiliensis</i> Müll. Arg. (rubber tree)	Rubber (latex)	β -1,3-glucanase	1E-20	1
BJ939513	Hev b 3	<i>H. brasiliensis</i>	Rubber (latex)	Small rubber particle protein	3E-20	2
BJ939948	Hev b 9	<i>H. brasiliensis</i>	Rubber (latex)	Enolase	4E-95	3

¹ The accession number of an EST that exhibited strongest similarity.

² The E-value of an EST that exhibited strongest similarity.

³ Total number of cDNA clones whose products were similar to the respective allergens (BLASTX score > 50 and E-value < 10⁻⁷).

related genes by gel blot analysis of total RNA extracted from various organs of *C. japonica* (Figure 5). The *CjMP1* transcripts were especially abundant in pollen grains, and were also detected in developed male flowers that included mature pollen; no transcripts were detected in the other organs examined. The patterns of signals in RNA gel blot analysis were very similar when we used the entire cDNA probe that corresponded to the short type of *CjMP1*-related genes and the probe specific to the long-type sequences. This result indicated that both the short type and the long type of *CjMP1*-related genes were expressed specifically in pollen grains.

To confirm that the abundant *CjMP1* transcripts could be translated into polypeptides in *C. japonica*, we prepared polyclonal antiserum against *CjMP1* (Figure 6). The *CjMP1*-specific antiserum was produced in rabbits immunized against a synthetic peptide of 16 amino acids (CEEVEHSTPQQKSV-EK), selected on the basis of the predicted primary structures of gene products (Figure 3). The GST-*CjMP1* fusion protein that had been expressed in *E. coli* was specifically recognized by the *CjMP1*-specific antiserum (Figure 6A). The fusion protein was found in the cytoplasmic fraction of *E. coli* cells and its molecular mass was almost equal to the sum of molecular masses of GST (26.7 kDa) and *CjMP1* proteins (11.2 kDa). After purification of the fusion protein by glutathione-affinity column chromatography, we detected three signals with the *CjMP1*-specific antiserum. However, binding of the antibodies to a protein of approximately 38 kDa gave the strongest signal. The results of ELISAs showed that the crude extract of *C. japonica* pollen reacted with the *CjMP1*-specific antiserum, as did the GST-*CjMP1* fusion protein (Figure 6B), suggesting that *CjMP1*-related proteins are present in pollen.

We also examined the allergenicity of *CjMP1* in patients with *C. japonica* pollinosis. None of the sera from 19 patients

reacted with the GST-*CjMP1* fusion protein, whereas all sera reacted with the crude extract of *C. japonica* pollen (Figure 7). Thus, the allergenicity of *CjMP1* was not demonstrated.

Discussion

In attempts to examine gene function at the genome level in conifers, EST sequencing projects have been undertaken for several gymnosperms (Allona et al. 1998, Ujino-Ihara et al. 2000, 2003, Brenner et al. 2003, Kirst et al. 2003, Pavy et al. 2005). In *C. japonica*, ESTs derived from the inner bark and developing male and female flowers were analyzed by Ujino-Ihara et al. (2000, 2003). The present report is the first, to our knowledge, of an analysis of ESTs derived from pollen of *C. japonica*.

Kirst et al. (2003) analysed 59,797 ESTs from a loblolly pine (*Pinus taeda* L.) and found that 50% of them exhibited no homology to angiosperms in public databases. An EST analysis in *Cycas* revealed that 694 (28%) of 2458 contigs exhibited no similarity to other plant genes (Brenner et al. 2003). These no-hit sequences might include genes unique to the respective species. Recent exhaustive analysis of datasets derived from pine cDNAs showed that among 7732 contigs, 61.5%, 59.4% and 55.0% matched sequences from *Arabidopsis*, rice and poplar, respectively (Pavy et al. 2005). Moreover, comparative analysis of ESTs generated from gymnosperms and angiosperms revealed that approximately 6% of tentatively unique genes of species in Cupressaceae had homologs in other conifers, but not in angiosperms, and about 70% had apparent homologs in angiosperms (Ujino-Ihara et al. 2005). In the present study, approximately 20% of transcripts had no apparent homologs in public databases and about 30% exhibited no similarity to sequences from *Arabidopsis* or rice. Comparison

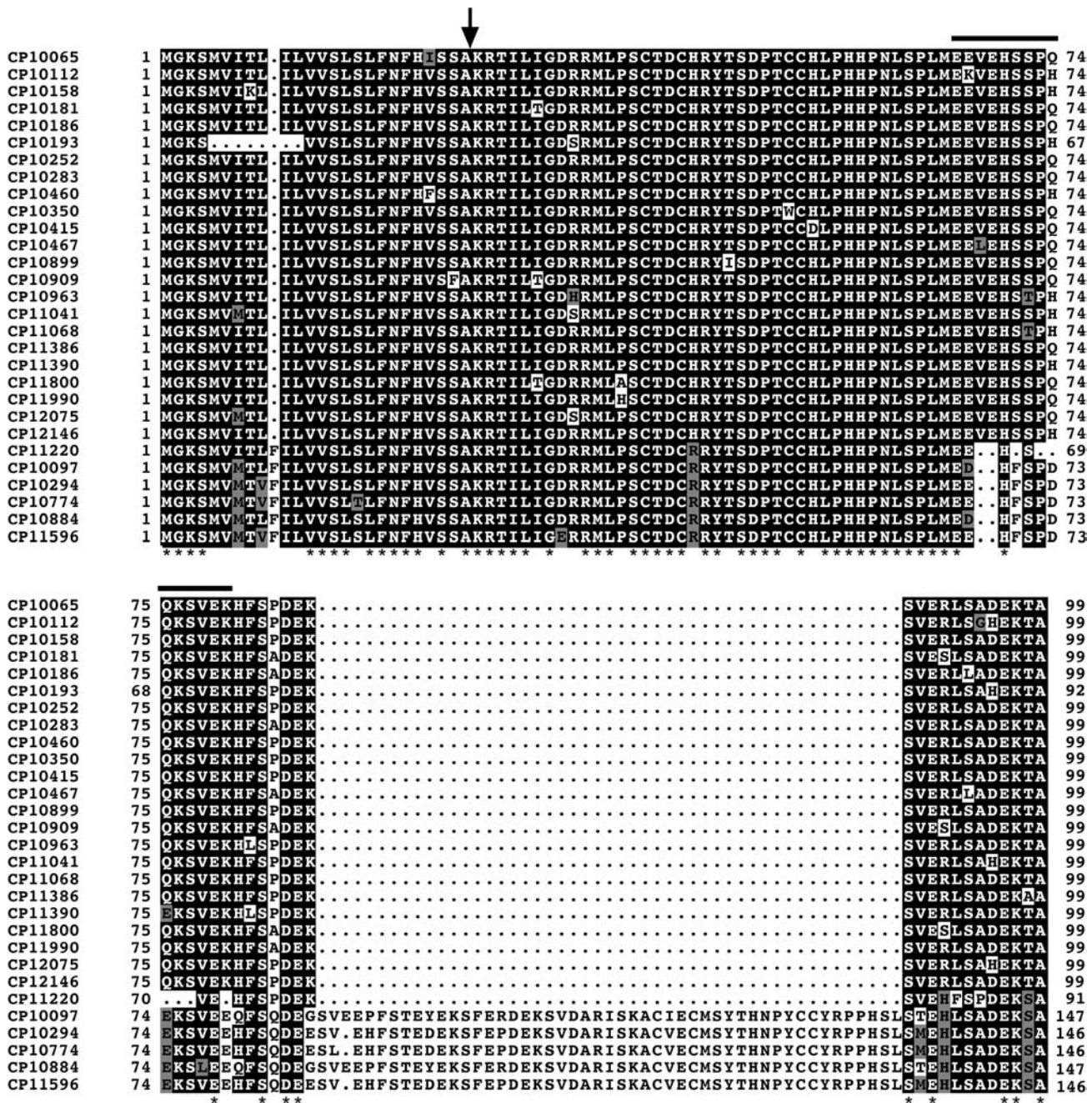


Figure 3. Alignment of the deduced amino acid sequences of CjMP1-related proteins. The diagram includes 29 predicted amino acid sequences of typical *CjMP1*-related cDNA clones. The name of each cDNA clone is indicated on the left. Amino acids in black and gray boxes are identical and similar, respectively, in at least 15 of the sequences. Conserved amino acid residues in all the sequences are indicated by asterisks. Dots indicate gaps introduced to maximize the extent of homology among sequences. A putative processing site is indicated by the vertical arrow. A line above the sequences indicates the peptide used for the production of CjMP1-specific antiserum.

of EST datasets revealed a larger overlap between coniferous species, a result that is consistent with the results of a comparative analysis of conifers and angiosperms (Ujino-Ihara et al. 2005).

To determine whether specific transcripts that are abundant in pollen are common to *Arabidopsis* and *C. japonica*, we searched for *Arabidopsis* genes that are homologous to the 20 most abundant transcripts in *C. japonica* pollen (Table 2).

Deduced products of 19 transcripts exhibited significant homology to *Arabidopsis* proteins, and all of the latter proteins, with the exception of At2g26010, had been identified with ATH1 Genome Arrays (Honys and Twell 2004, Pina et al. 2005). Two of them (At1g47980 and At2g25890) were selectively expressed in pollen of *Arabidopsis*, according to the data reported by Honys and Twell (2004). The genes encoding proteins of *C. japonica* that were homologous to At1g47980 and

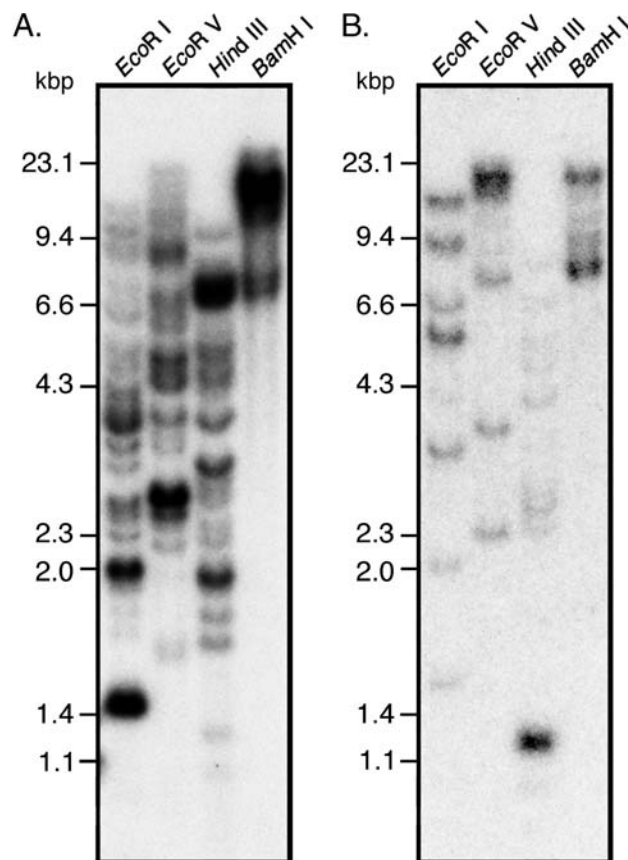


Figure 4. Genomic DNA gel blotting of *CjMPI*-related genes in the genome of *C. japonica*. Aliquots of 13 μ g of genomic DNA were digested with the indicated restriction enzymes. The positions of size markers are shown on the left, with sizes indicated in kilobase pairs (kbp). (A) The entire cDNA of a short *CjMPI*-related gene was used as the probe, which was assumed to hybridize to both the short type and the long type of *CjMPI*-related genes. (B) The long-type-specific insert of *CjMPI*-related genes was used as the probe.

At2g25890 were predominantly expressed in pollen (Figure 2). Desiccation-related protein in *Craterostigma plantagineum* Hochst. is homologous to At1g47980. This protein may be involved in the desiccation that occurs during pollen maturation. In floral microspores, At2g25890 is one of three oleosins expressed (Kim et al. 2002). This type of oleosin might be involved in preservation of the integrity of oil bodies in pollen. Our results indicate that a number of genes are expressed in common in the pollen of *Arabidopsis* and *C. japonica*.

In *Arabidopsis* pollen, mRNAs that encoded proteins involved in cell wall metabolism, such as polygalacturonases and pectinesterases, were among the 10 most abundant transcripts (Honys and Twell 2003). In *C. japonica*, we found cDNAs in our pollen library for proteins involved in cell wall metabolism, such as Cry j 1 and Cry j 2 (Table 3); however, the mRNAs were not among the 20 most abundant transcripts. Earlier studies on functional genes in the transcriptome of *A. thaliana* pollen revealed the overrepresentation of mRNAs for proteins involved in cell wall metabolism, the cytoskeleton

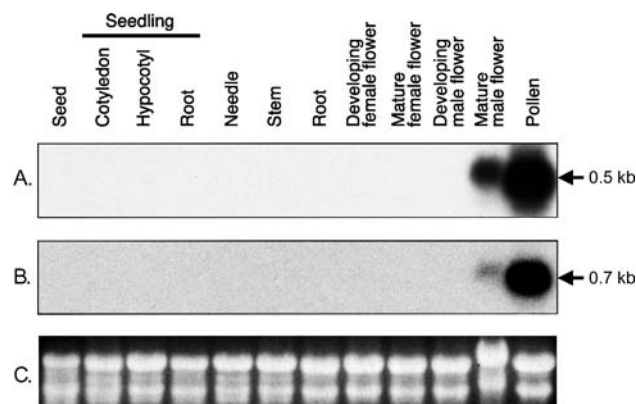


Figure 5. RNA gel blotting of *CjMPI*-related transcripts in various organs of *C. japonica*. Ten micrograms of total RNA were loaded in each lane of a 1% agarose gel that contained formaldehyde and subjected to electrophoresis. The lengths of transcripts are shown on the right. (A) The entire cDNA of a short *CjMPI*-related gene was used as a probe. (B) The long-type-specific insert of a *CjMPI*-related gene was used as the probe. (C) The results of staining the gel with ethidium bromide before blotting confirmed that the same amount of RNA had been loaded in each lane.

and signaling, and under-representation of mRNAs involved in translation and energy pathways (Honys and Twell 2003, Becker et al. 2003). Becker et al. (2003) indicated that 104 (37%) of 283 genes that were constitutively expressed in vegetative tissues of *Arabidopsis* were not expressed in *Arabidopsis* pollen grains, and 29 (27%) of 104 transcripts encoded putative or known ribosomal proteins. In many angiosperm species, rRNA gene transcription is inactivated at the late stage of pollen development and during the growth of pollen tubes, and there are large stores of ribosomal proteins in the pollen grains themselves (Mascarenhas 1975, 1990). The present study indicates that a large fraction of pollen transcripts encoded proteins that are involved in the synthesis and modification of proteins, and transcripts encoding ribosomal proteins were relatively abundant in *C. japonica* pollen (Figure 1; Table 2). Analysis by RT-PCR suggested that transcripts encoding ribosomal proteins were expressed in pollen grains at levels similar to those in vegetative tissues (Figure 2). The mRNAs for some ribosomal proteins in pollen seem to be as abundant as they were in sporophytic tissues. The reasons for the difference between reported results for angiosperms and our results in *C. japonica* are unclear. Further investigations are required to confirm the transcriptional activity that is related to ribosomal proteins in pollen grains of gymnosperms.

The 20 most abundant transcripts in *C. japonica* pollen and the expression of homologous genes in *Arabidopsis* presents several inconsistencies. The gene expression for seven proteins (At1g80320, At5g52390, At3g57490, At4g17520, At5g67360, At5g26340 and At5g24510) was not detected in *Arabidopsis* pollen by two groups of researchers (Honys and Twell 2004, Pina et al. 2005). In addition, gene expression for four proteins (At2g27720, At3g12490, At1g48920 and At4g25890) was suppressed in pollen grains (Pina et al. 2005). Several transcripts whose expression is suppressed in pollen

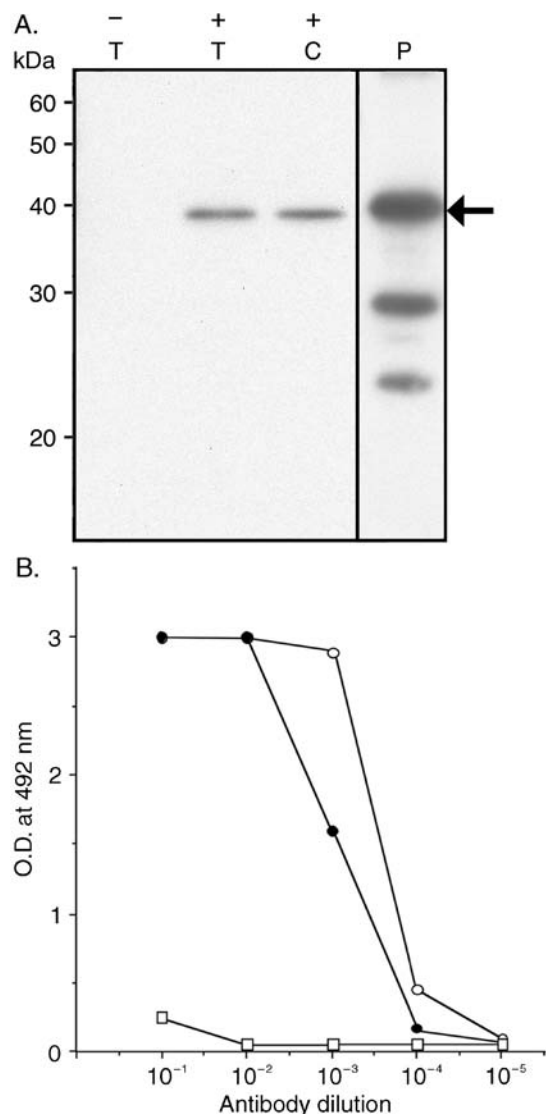


Figure 6. Reaction of CjMP1-specific antiserum with GST-CjMP1 and a crude extract of *C. japonica* pollen. (A) Induction of the synthesis of recombinant GST-CjMP1 in *E. coli* and purification of the protein. Total protein (T) and a cytoplasmic fraction (C) of *E. coli* before (–) and after (+) induction with IPTG, and purified GST-CjMP1 (P) were subjected to SDS-PAGE (12% polyacrylamide). Immunoblotting analysis was performed with the CjMP1-specific antiserum. The molecular masses of standard proteins are indicated on the left. The predicted mobility of the recombinant GST-CjMP1 is indicated by the arrow. (B) ELISA titration curves for reaction of the CjMP1-specific antiserum with purified GST-CjMP1 and a crude extract of *C. japonica* pollen. The crude extract of *C. japonica* pollen (●) and purified GST-CjMP1 (○) were immobilized separately at a concentration of $2 \mu\text{g ml}^{-1}$ on plastic microtiter plates. PBS (□) was used as a control in place of antigens. The CjMP1-specific antiserum was diluted exponentially from 10^{-1} to 10^{-5} , as indicated.

grains of *Arabidopsis* were abundant in pollen grains of *C. japonica* (Table 2). These differences can be assumed to be due mainly to the difference in species. Differences related to pollen collection may also have influenced the results. The pollen grains of *Arabidopsis* were collected from inflorescences and

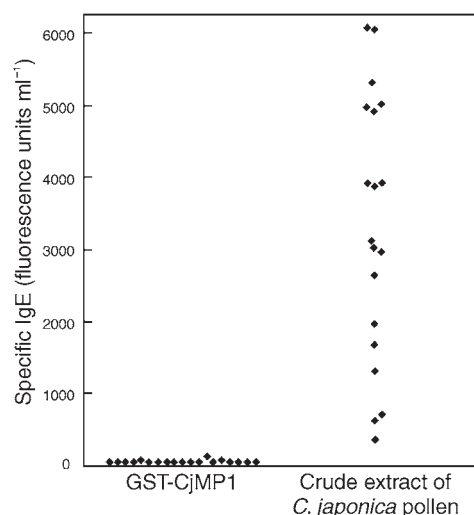


Figure 7. Binding frequencies of GST-CjMP1 and a crude extract of *C. japonica* pollen to IgE in serum from patients with pollinosis. Sera from 19 patients with pollinosis were tested in a fluorometric ELISA. The diamonds show the binding of IgE in the serum of each patient to plates coated with the purified GST-CjMP1 fusion protein (left) and with a crude extract of *C. japonica* pollen grains (right).

sorted in an appropriate buffer. We collected pollen grains directly from male flowers of *C. japonica* by gently tapping them and stored the pollen at -80°C . The hydration of *Arabidopsis* pollen might affect the pattern of gene expression. The number of cDNA clones (1929) in the present study might not be sufficient to provide a complete overview of gene expression in pollen grains of *C. japonica*. Larger-scale analysis of the transcriptome in pollen grains will clarify differences in patterns of gene expression among species.

The most abundant type of EST in *C. japonica* pollen encoded novel small proteins with strong mutual similarity (Table 2). Each predicted amino acid sequence contained a signal peptide for secretion. The corresponding genes, termed CjMP1-related genes, formed a multi-gene family, and their transcripts were abundant in pollen grains specifically (Figures 4 and 5). ELISAs using polyclonal antibodies against a sequence of 16 amino acid residues from CjMP1-related proteins indicated that related proteins were present in a crude extract of *C. japonica* pollen (Figure 6). This result suggests CjMP1-related genes were not only transcribed, but also translated in pollen. When we subjected the recombinant GST-CjMP1 fusion protein to ELISA with sera from 19 patients with pollinosis, we found no evidence of allergenicity (Figure 7). This result indicated that CjMP1 is not a major allergen, or that recombinant GST-CjMP1 had lost its intrinsic allergenicity. The functions of CjMP1-related proteins are unknown to date, and further study is needed to clarify their role in *C. japonica*.

We analyzed possible homologies between proteins encoded by ESTs derived from *C. japonica* pollen and known plant allergens in an attempt to identify the novel candidate genes for pollen allergens. Four known allergens in *C. japonica* pollen resemble pollen or food allergens from angio-

sperms. *Cry j 1* is homologous to the major short ragweed allergen Amb a 1 (Sone et al. 1994). *Cry j 2* is similar to pollen allergen phl p 13 of timothy grass (Suck et al. 2000). The third allergen, CJP-6, is homologous to members of the isoflavone reductase family, which includes the birch pollen allergen Bet v 6 and the pear allergen Pyr c 5 (Kawamoto et al. 2002). The fourth allergen from *C. japonica*, CJP-4, resembles class I chitinase allergens, such as Hev b 11, Pers a 1 and Mus a 1 (Fujimura et al. 2005). In addition, CJP-4 cross-reacts with latex C-serum (Fujimura et al. 2005). The cross-reactivity between *C. japonica* pollen and tomato (*Lycopersicon esculentum* Mill.) has been demonstrated in humans and dogs (Fujimura et al. 2002, Kondo et al. 2002). These findings indicate that unidentified allergens in *C. japonica* pollen might resemble known plant allergens.

In our study of the pollen of *C. japonica*, we found cDNA clones that encoded proteins similar to four types of pollen allergens, eight types of food allergens and three types of latex allergens, in addition to *Cry j 1* and *Cry j 2* (Table 3). The product of a cDNA clone (Accession Nos. BJ937870 and BJ937871) that encoded a Bet v 6-like protein was 85.6% identical to CJP-6, indicating this protein may be allergenic. It is likely that other proteins encoded by genes for allergen-like proteins are also allergenic. Further detailed information about ESTs derived from *C. japonica* pollen should contribute to studies of pollen development in gymnosperms and the attempts to isolate novel allergens.

Acknowledgments

The authors are grateful to Dr. Ovidiu Ivanciuc and Dr. Werner Braun of the University of Texas Medical Branch for providing access to the SDAP database. This work was supported in part by a Grant-in-Aid "Research Project for Utilizing Advanced Technologies in Agriculture, Forestry and Fisheries," by the Rice Genome Project (Grant No. CG1107) from the Research Council, Ministry of Agriculture, Forestry and Fisheries of Japan, in part by a Grant-in-Aid for Scientific Research (No. 18780123) from the Ministry of Education, Culture, Sports, Science and Technology of Japan and in part by a research grant (No. 200607) from the Forestry and Forest Products Research Institute.

References

- Allona, I., M. Quinn, E. Shoop et al. 1998. Analysis of xylem formation in pine by cDNA sequencing. *Proc. Natl. Acad. Sci. USA* 95:9693–9698.
- Becker, J.D., L.C. Boavida, J. Carneiro, M. Hauray and J.A. Feijó. 2003. Transcriptional profiling of *Arabidopsis* tissues reveals the unique characteristics of the pollen transcriptome. *Plant Physiol.* 133:713–725.
- Brenner, E.D., D.W. Stevenson, R.W. McCombie et al. 2003. Expressed sequence tag analysis in *Cycas*, the most primitive living seed plant. *Genome Biol.* 4:R78.
- Ewing, B., L. Hillier, M.C. Wendl and P. Green. 1998. Base-calling of automated sequencer traces using *phred*. I. Accuracy assessment. *Genome Res.* 8:175–185.
- Fujimura, M., K. Ohmori, K. Masuda, H. Tsujimoto and M. Sakaguchi. 2002. Oral allergy syndrome induced by tomato in a dog with Japanese cedar (*Cryptomeria japonica*) pollinosis. *J. Vet. Med. Sci.* 64:1069–1070.
- Fujimura, T., S. Shigeta, S. Kawamoto, T. Aki, M. Masubuchi, T. Hayashi, K. Yoshizato and K. Ono. 2004. Two-dimensional IgE-binding spectrum of Japanese cedar (*Cryptomeria japonica*) pollen allergens. *Int. Arch. Allergy Immunol.* 133:125–135.
- Fujimura, T., S. Shigeta, T. Suwa, S. Kawamoto, T. Aki, M. Masubuchi, T. Hayashi, M. Hide and K. Ono. 2005. Molecular cloning of a class IV chitinase allergen from Japanese cedar (*Cryptomeria japonica*) pollen and competitive inhibition of its immunoglobulin E-binding capacity by latex C-serum. *Clin. Exp. Allergy* 35:234–243.
- Futamura, N., H. Kouchi and K. Shinohara. 2002a. A gene for pectate lyase expressed in elongating and differentiating tissues of a Japanese willow (*Salix gilgiana*). *J. Plant Physiol.* 159:1123–1130.
- Futamura, N., Y. Mukai, M. Sakaguchi, H. Yasueda, S. Inouye, T. Midoro-Horiuti, R.M. Goldblum and K. Shinohara. 2002b. Isolation and characterization of cDNAs that encode homologs of a pathogenesis-related protein allergen from *Cryptomeria japonica*. *Biosci. Biotechnol. Biochem.* 66:2495–2500.
- Honys, D. and D. Twell. 2003. Comparative analysis of the *Arabidopsis* pollen transcriptome. *Plant Physiol.* 132:640–652.
- Honys, D. and D. Twell. 2004. Transcriptome analysis of haploid male gametophyte development in *Arabidopsis*. *Genome Biol.* 5:R85.
- Huang, X. and A. Madan. 1999. CAP3: A DNA sequence assembly program. *Genome Res.* 9:868–877.
- Ivanciuc, O., C.H. Schein and W. Braun. 2003. SDAP: database and computational tools for allergenic proteins. *Nucleic Acids Res.* 31:359–362.
- Kawamoto, S., T. Fujimura, M. Nishida et al. 2002. Molecular cloning and characterization of a new Japanese cedar pollen allergen homologous to plant isoflavone reductase family. *Clin. Exp. Allergy* 32:1064–1070.
- Kikuchi, S., K. Satoh, T. Nagata et al. 2003. Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. *Science* 301:376–379.
- Kim, H.U., K. Hsieh, C. Ratnayake and A.H.C. Huang. 2002. A novel group of oleosins is present inside the pollen of *Arabidopsis*. *J. Biol. Chem.* 277:22,677–22,684.
- Kirst, M., A.F. Johnson, C. Baucom et al. 2003. Apparent homology of expressed genes from wood-forming tissues of loblolly pine (*Pinus taeda* L.) with *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 100:7383–7388.
- Knox, R.B. and C. Suphioglu. 1996. Pollen allergens: development and function. *Sex. Plant Reprod.* 9:318–323.
- Komiyama, N., T. Sone, K. Shimizu, K. Morikubo and K. Kino. 1994. cDNA cloning and expression of *Cry j II*, the second major allergen of Japanese cedar pollen. *Biochem. Biophys. Res. Commun.* 201:1021–1028.
- Kondo, Y., R. Tokuda, A. Urisu and T. Matsuda. 2002. Assessment of cross-reactivity between Japanese cedar (*Cryptomeria japonica*) pollen and tomato fruit extracts by RAST inhibition and immunoblot inhibition. *Clin. Exp. Allergy* 32:590–594.
- Lee, J.-Y. and D.-H. Lee. 2003. Use of serial analysis of gene expression technology to reveal changes in gene expression in *Arabidopsis* pollen undergoing cold stress. *Plant Physiol.* 132:517–529.
- Lee, Y., J. Tsai, S. Sunkara et al. 2005. The TIGR Gene Indices: clustering and assembling EST and known genes and integration with eukaryotic genomes. *Nucleic Acids Res.* 33:D71–74.
- Mascarenhas, J.P. 1975. Biochemistry of angiosperm pollen development. *Bot. Rev.* 41:259–314.
- Mascarenhas, J.P. 1990. Gene activity during pollen development. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 41:317–338.

- Mukai, Y., M. Fukui, N. Futamura and K. Shinohara. 1998. Molecular cloning of a cDNA (Accession No. AB013728) encoding the light-harvesting chlorophyll a/b-binding protein of photosystem II from the sugi tree (*Cryptomeria japonica* D. Don). *Plant Physiol.* 117: 1126.
- Namba, M., M. Kurose, K. Torigoe, K. Hino, Y. Taniguchi, S. Fukuda, M. Usui and M. Kurimoto. 1994. Molecular cloning of the second major allergen, *Cry j* II, from Japanese cedar pollen. *FEBS Lett.* 353:124–128.
- Pavy, N., J. Laroche, J. Bousquet and J. Mackay. 2005. Large-scale statistical analysis of secondary xylem ESTs in pine. *Plant Mol. Biol.* 57:203–224.
- Pina, C., F. Pinto, J.A. Feijó and J.D. Becker. 2005. Gene family analysis of the *Arabidopsis* pollen transcriptome reveals biological implications for cell growth, division control, and gene expression regulation. *Plant Physiol.* 138:744–756.
- Pruitt, K.D., T. Tatusova and D.R. Maglott. 2005. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* 33: D501–504.
- Sakaguchi, M., S. Inouye, M. Taniai, S. Ando, M. Usui and T. Matuhasi. 1990. Identification of the second major allergen of Japanese cedar pollen. *Allergy* 45:309–312.
- Sambrook, J., E.F. Fritsch and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*. 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Shinohara, K. and A. Murakami. 1996. Changes in leaves of thylakoid components in chloroplasts of pine needles of different ages. *Plant Cell Physiol.* 37:1102–1107.
- Sone, T., N. Komiyama, K. Shimizu, T. Kusakabe, K. Morikubo and K. Kino. 1994. Cloning and sequencing of cDNA coding for *Cry j* I, a major allergen of Japanese cedar pollen. *Biochem. Biophys. Res. Commun.* 199:619–625.
- Suck, R., A. Petersen, S. Hagen, O. Cromwell, W.M. Becker and H. Fiebig. 2000. Complementary DNA cloning and expression of a newly recognized high molecular mass allergen Phl p 13 from timothy grass pollen (*Phleum pratense*). *Clin. Exp. Allergy* 30: 324–332.
- Takahashi-Omoe, H., K. Omoe, M. Sakaguchi, Y. Kameoka, S. Matsushita and T. Inada. 2004. Production of virus-specific antiserum corresponding to sequences in the lactate dehydrogenase-elevating virus (LDV) ORF6 protein. *Comp. Immunol. Microbiol. Infect. Dis.* 27:47–55.
- Taniguchi, Y., A. Ono, M. Sawatani, M. Nanba, K. Kohno, H. Usui, M. Kurimoto and T. Matuhasi. 1995. *Cry j* I, a major allergen of Japanese cedar pollen, has pectate lyase enzyme activity. *Allergy* 50:90–93.
- Tatusov, R.L., N.D. Fedorova, J.D. Jackson et al. 2003. The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* 4:41.
- Twell, D. 2002. The developmental biology of pollen. *In* *Plant Reproduction*. Eds. S.D. O'Neil and J.A. Roberts. Sheffield Academic Press, Sheffield, U.K., pp 86–153.
- Ujino-Ihara, T., K. Yoshimura, Y. Ugawa, H. Yoshimaru, K. Nagasaka and Y. Tsumura. 2000. Expression analysis of ESTs derived from the inner bark of *Cryptomeria japonica*. *Plant Mol. Biol.* 43: 451–457.
- Ujino-Ihara, T., Y. Taguchi, K. Yoshimura and Y. Tsumura. 2003. Analysis of expressed sequence tags derived from developing seed and pollen cones of *Cryptomeria japonica*. *Plant Biol.* 5:600–607.
- Ujino-Ihara, T., H. Kanamori, H. Yamane, Y. Taguchi, N. Namiki, Y. Mukai, K. Yoshimura and Y. Tsumura. 2005. Comparative analysis of expressed sequence tags of conifers and angiosperms reveals sequences specifically conserved in conifers. *Plant Mol. Biol.* 59:895–907.
- Wagner, D.B., G.R. Furnier, M.A. Saghai-Marooof, S.M. Williams, B.P. Dancik and R.W. Allard. 1987. Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids. *Proc. Natl. Acad. Sci. USA* 84:2097–2100.
- Willing, R.P. and J.P. Mascarenhas. 1984. Analysis of the complexity and diversity of messenger RNAs from pollen and shoots of *Tradescantia*. *Plant Physiol.* 75:865–868.
- Willing, R.P., D. Bashe and J.P. Mascarenhas. 1988. An analysis of the quantity and diversity of messenger RNAs from pollen and shoots of *Zea mays*. *Theor. Appl. Genet.* 75:751–753.