Plant Molecular Biology **46:** 43–56, 2001.
© 2001 *Kluwer Academic Publishers. Printed in the Netherlands.*

A leucine-rich repeat region is conserved in pollen extensin-like (Pex) proteins in monocots and dicots

Suzanne Stratford[1,+], William Barnes[1,4,+], Diane L. Hohorst[1], Jason G. Sagert[1], Robyn Cotter[2], Alice Golubiewski[1], Allan M. Showalter[3], Sheila McCormick[2] and Patricia Bedinger[1,*]
[1]*Department of Biology and Program in Cell and Molecular Biology, Colorado State University, Fort Collins, CO 80523-1878, USA (*author for correspondence; e-mail: bedinger@lamar.colstgate.edu);* [2]*Plant Gene Expression Center, USDA/ARS-UC Berkeley, 800 Buchanan St., Albany, CA 94710, USA;* [3]*Department of Environmental and Plant Biology, Molecular and Cellular Biology Program, Ohio University, Athens, OH 45701-2979, USA;* [4]*Current address: Department of Biology, Clarion University of Pennsylvania, Clarion, PA 16214, USA;* [+]*these authors contributed equally to the work*

Received 19 June 2000; accepted in revised form 10 January 2001

Key words: cell wall protein, LRR, maize, pollen, tomato

Abstract

We previously isolated a pollen-specific gene encoding a pollen tube wall-associated glycoprotein with a globular domain and an extensin domain from maize (mPex1). To evaluate which protein domains might be important for function, we isolated a second monocot gene (mPex2) and a dicot gene (tPex). Each gene encodes a signal sequence, an N-terminal globular domain comprised of a variable region, a leucine-rich repeat (LRR) with an adjacent cysteine-rich region, a transition region and an extensin-like C-terminal domain. The LRRs of the maize and tomato Pex proteins are highly conserved. Although the extensin domains in the maize and tomato proteins vary in length and in amino acid sequence, they are likely to be structurally conserved. Additional putative Pex gene sequences were identified by either GenBank search (*Arabidopsis*) or PCR (sorghum and potato); all encode conserved LRRs. The presence of a conserved LRR in the known and potential Pex proteins strongly suggests that this motif is involved in the binding of a specific ligand during pollen tube growth. Gene expression studies using RNA and protein blotting as well as promoter-reporter gene fusions in transient and stable transformation indicate that the tomato Pex gene is pollen-specific.

Introduction

In flowering plants, pollen germinates a pollen tube on the receptive stigma that then grows within the specialized transmitting tissue of the pistil to reach the ovule (Steer and Steer, 1989; Mascarenhas, 1993). The pollen tube transports the pollen cytoplasm, tube nucleus and sperm cells, entering the ovule through the micropyle. Once the pollen tube reaches the embryo sac, the tube tip ruptures, releasing the sperm for double fertilization. The extension of the pollen tube can be quite rapid, close to 1 cm/h in maize (Barnabas and Fridvasky, 1984). As it grows, the pollen tube must generate significant amounts of cell wall material, which must be rapidly organized into a resilient structure at the pollen tube tip to allow the migrating tube to penetrate the extracellular matrix of the transmitting tissue.

The cell wall of pollen tubes is bilayered; the thick inner wall is composed of callose and sometimes small amounts of cellulose and the outer layer is composed of pectins (Heslop-Harrison, 1987; Ferguson *et al.*, 1998). In addition to polysaccharides, the pollen tube cell wall contains proteins, including arabinogalactan proteins (Li *et al.*, 1992; Jauh and Lord, 1996) and

The nucleotide sequence data reported is in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers AF159297 (mPex2) and AF159296 (tPex).

extensin-like proteins (Pex, for pollen extensin-like) (Rubinstein *et al.*, 1995a, b). The extensins are a major class of plant cell wall proteins that contain a repeated serine-proline$_{3-7}$ (SP_n) motif (Lamport, 1965, 1977; Kieliszewski and Lamport, 1994). The Pex proteins, first reported in maize, have an N-terminal globular domain and a C-terminal extensin-like domain (Rubinstein *et al.*, 1995a, b); these proteins have been localized to the callose portion of the pollen tube cell wall (Rubinstein *et al.*, 1995a, b).

The globular domain-extensin domain chimeric Pex proteins may represent an ancient paradigm in the recognition of reproductive structures conserved during the evolutionary history of plants since the sexual agglutinins of *Chlamydomonas* have been shown to have a globular-rod structure (Adair *et al.*, 1983; Cooper *et al.*, 1983). Recent studies have revealed chimeric proteins with extensin domains as well as other protein domains in the female tissues of higher plants (Baldwin *et al.*, 1992; Chen *et al.*, 1992; Goldman *et al.*, 1992; Wu *et al.*, 1993; Lind *et al.*, 1994). The role of this class of proteins in pollen-pistil interactions is clearly a focus for further investigation.

By identifying conserved protein motifs in diverse plant species, the possible role(s) of Pex proteins in plant reproduction may be suggested. We report here the analysis of a second maize Pex gene (mPex2) and a dicot (tomato) Pex gene (tPex), as well as potential Pex genes from *Arabidopsis*, and portions of potential Pex genes from potato and sorghum. All of the known and potential Pex proteins contain a conserved leucine-rich repeat (LRR), a protein motif that mediates specific protein-protein binding (Kobe and Deisenhofer, 1994, 1995). The conservation of LRRs supports the idea that the Pex proteins are involved in protein-ligand interactions during pollen tube growth.

Materials and methods

mPex2 genomic clones

mPex2 genomic sequences were identified in a maize genomic library (Clontech) by hybridizing plaque lifts with a digoxygenin-UTP-labeled (Boehringer Mannheim) partial cDNA specific to mPex2 (pSF21) that had been identified in a screen for pollen-specific genes. Inserts were isolated from the EMBL3 vector and cloned into BlueScript (Stratagene), to produce pAG3 and pAG4, which together contained the entire mPex2 coding sequence. Sequencing was performed at the MacroMolecular Resources facility at Colorado State University.

Tomato cDNA and genomic clones

A tomato anther cDNA library (McCormick *et al.*, 1987) was screened at low stringency with a maize cDNA probe (pSF66) that included the extensin domain. Out of 25 000 plaques screened, 40 recombinant phage were isolated and rescreened. Two positive clones, pWB1 and pWB2 (Figure 2B), were analyzed further. cDNA inserts were PCR-amplified with λgt11-specific primers by PCR, and cloned into BlueScript (Stratagene). DNA sequencing revealed a striking similarity to a previously published partial sequence of a tomato extensin gene, Tom 5 (Showalter *et al.*, 1985). The 10 kb genomic insert of Tom 5 was recloned into BlueScript (pSS100, Figure 2A), subcloned and sequenced, revealing the presence of the predicted Pex globular domain. Additional cDNAs were obtained and sequenced (pSS201 and pSS202, Figure 2B). A restriction enzyme map of genomic clone pSS100, relative locations of cDNA clones, probes and promoter-GUS/nos fusions are shown schematically in Figure 2.

PCR cloning of part of the N-terminal region of potato and sorghum Pex genes

The polymerase chain reaction contained 20 mM Tris pH 8.3, 50 mM KCl, 0.1% Tween 20, 5% DMSO, 0.2 mM of each dNTP, 0.5 μmol of the upstream and downstream primers, 1.5 mM $MgCl_2$, 2.5 U *Taq* DNA polymerase, and 200 ng of sorghum (cv. Shanqi Red) genomic DNA or 80 ng of potato (cv. Désíree) genomic DNA. The sequence of the upstream primer was 5′-AT(C/T)GCCC(A/G)(C/T)CGCCA(A/G)CTG-3′ for sorghum and 5′-AT(C/T)GCCC(A/G)(C/T)CGCCA-(A/G)CTT-3′ for potato. For both species, the sequence of the downstream primer was 5′-CC(C/T)TC-(A/G)AAGTCGTTGAACC-3′. For each reaction, the denaturing temperature was 94 °C, the annealing temperature was 55 °C, and the extension temperature was 72 °C. The PCR fragments between 300 and 600 bp were isolated from an agarose gel and ligated into the pGEM-T-easy vector (Promega) to create a plasmid library. Colonies were screened using a digoxigenin-11-dUTP-labeled probes according to the manufacturer's protocol (Boehringer Mannheim). Probes consisted of the amplified region of mPex1 for the sorghum library and the amplified region of tPex for the potato library.

```
 AAGCTTCCAGCAGCATGTATATTTTGGGGTTGCACGAGGTGGACCAAGCACGACCGCTCGCACGCAGCCATGCAGATAGGCCAGGCTCCAGGAACGCCAATGCCGCATGCATGCACCGACAATCCCTGAAACAAGGGCCCCATGCTGC   148
CAGAGTGCCAGCTCCACCCTCTTCTGTCGTCGTCCTCGTCCGTCCGTCCCTGGAGCCTGGACAGGCGCAGGCCGCCAATGCCGCGGAGCCAGCGAGCGGCGGCCACCGTCCACCACCACCACCACCGGCCTGACGTACGATCAGCGCGCG   298
CATGCAACCGCACGCTGTTACCGCCCCCCGCCCGGCGTCCATAAAGCACCAGGGCGACGGGGGGTGTCGACGTCAGTGTGCACGCCCACCGCCGCCCGCTCGCTCTCCCTCATGGACCCTCCTCTCCGGCTGCTGCCCGATGGGCGCCTC   448
                                                                                                                     M  D  P  P  L  R  L  L  P  D  G  R  L     13
ATGGCAGCGCTGTTGCTCCTGGCAGCCTGCATGTCCGCCTGCAGCGTGCAGGCGGTGACGAGCGCGGAGGCCTCCTACATCGCGCGCCGCCAGCTGCTGGCCATGAAGGAGGCCGGCGCCGGCGAGGGCGGCGACCTGCCGGCGGACTTC   598
 M  A  A  L  L  L  L  A  A  C  M  S  A  C  S  V▼Q  A  V  T  S  A  E  A  S  Y  I  A  R  R  Q  L  L  A  M  K  E  A  G  A  G  E  G  G  D  L  P  A  D  F     66
GAGTTCGACGACCGCGTCGGCGCCGCCAACTTCCCCAACCCGCGCCTCCGCCGCGCGTACATCGCGCTGCAGGCCTGGCGCCGCGCCTTCTACTCCGACCCCAAGGGCTACACCGCCAACTGGGTCGGCGAGGACGTGTGCAAGTACAAC   748
 E  F  D  D  R  V  G  A  A  N  F  P  N  P  R  L  R  R  A  Y  I  A  L  Q  A  W  R  R  A  F  Y  S  D  P  K  G  Y  T  A  N  W  V  G  E  D  V  C  K  Y  N    116
GGCGTCATCTGCACGGAGGCGCTGGACGACCCCAAGGTCACGGTCGTGGCCGGGATCGACCTCAACGGCGCCGACATCGCGGGGTACCTGCCGCCCGAGCTCGGCCTGCTCACCGACCTCGCCTTCTTCCACATCAACACCAACCGCTTC   898
 G  V  I  C  T  E  A  L  D  D  P  K  V  T  V  V  A  G  I  D  L  N  G  A  D  I  A  G  Y  L  P  P  E  L  G  L  L  T  D  L  A  F  F  H  I  N  T  N  R  F    166
TGCGGCATCATCCCCAAGAGCATGTCGCGGCTGTCGCTGCTGCACGAGTTCGACGTGAGCAACAACCGCTTCGTCGGCGTCTTCCCCTACGTCTGCCTCGAGATGGTGTCGCTCAAGTACCTCGACCTCCGCTTCAACGACTTTGAGGGC  1048
 C  G  I  I  P  K  S  M  S  R  L  S  L  L  H  E  F  D  V  S  N  N  R  F  V  G  V  F  P  Y  V  C  L  E  M  V  S  L  K  Y  L  D  L  R  F  N  D  F  E  G    216
GAGCTGCCGCCAGCGCTCTTCGACAAGGACCTGGACGCCATCTTCGTCAACACCAACCGATTCGTGGGCTATATCCCGGAGAACCTCGGCAACTCCACGGCGTCCGTCATCGTTTTCGCCAACAACGCCTTCGTCGGATGCATCCCCAAG  1198
 E  L  P  P  A  L  F  D  K  D  L  D  A  I  F  V  N  T  N  R  F  V  G  Y  I  P  E  N  L  G  N  S  T  A  S  V  I  V  F  A  N  N  A  F  V  G  C  I  P  K    266
AGCATCGGCCGCATGGTGAAGACGCTGGACGAAATCATCTTCCTCAACAACAAACTCGACGGGTGCCTGCCGCTGGAGATGGGCCTGCTGGTCAACACCACCGTCATCGACGTCAGCGGGAACATGCTCGTGGGCACCATCCCCGAGCAG  1348
 S  I  G  R  M  V  K  T  L  D  E  I  I  F  L  N  N  K  L  D  G  C  L  P  L  E  M  G  L  L  V  N  T  T  V  I  D  V  S  G  N  M  L  V  G  T  I  P  E  Q    316
CTCTCCAACATCGCCAAGCTGGAGCAGCTCGACGTGTCCCGGAACGTGTTCACCGGCATCGTGCAGGAGTCCATGTGCGAGCTCCGGGCGCTGGTCAACTTCAGGTTCGCCTTCAACTTGTTCAACTCGGAGGCCGTCGTGTGCATGCCG  1498
 L  S  N  I  A  K  L  E  Q  L  D  V  S  R  N  V  F  T  G  I  V  Q  E  S  M  C  E  L  R  A  L  V  N  F  R  F  A  F  N  L  F  N  S  E  A  V  V  C  M  P    366
TCCGACAACGCGCTCGTCAACTTCGACGACAGGGACAACTGCCTCGGCGCGCTGCGCCCGGCGCAGAAGACCGCGCTGCAGTGCGCTCCCGTGCTCGCGCGCCCCGTCGAGGTCGACTGCAGCAAGCACGTGTGCGCTGGCTACCCCACA  1648
 S  D  N  A  L  V  N  F  D  D  R  D  N  C  L  G  A  L  R  P  A  Q  K  T  A  L  Q  C  A  P  V  L  A  R  P  V  E  V  D  C  S  K  H  V  C  A  G  Y  P  T    416
CCCGGAGGGCCACCATCGTCGTCGGTGCCAGGGAAGCCGCCTTCGGTGCCAGGGAAGCCTGCTGCACCTGCGCCCATGCCAACCCCGCATACACCGCCGGATGTGTCGCCAGAACCGCTTCCAGAACCTTCGCCGGTGCCTGCACCTGCA  1798
 P  G  G  P  P  S  S  S  V  P  G  K  P  P  S  V  P  G  K  P  A  A  P  A  P  M  P  T  P  H  T  P  P  D  V  S  P  E  P  L  P  E  P  S  P  V  P  A  P  A    466
CCCATGCGGATGCCGACCCTTCGTTCACCGCCGGCGGACGAATACATTCCAACGCCACCTGTACCCGCGAAGTCACCGCCAGGGACTTCGCCACCGGCATCACGTGGTGCTCCACCGCTTCAGGCACAACCTCCAGCTGCCTCCTCCCCG  1948
 P  M  R  M  P  T  L  R  S  P  P  A  D  E  Y  I  P  T  P  P  V  P  A  K  S  P  P  G  T  S  P  P  A  S  R  G  A  P  P  L  Q  A  Q  P  P  A  A  S  S  P    516
CCTGCAACACCTGTGAAATCGTCTCCTCCGCCAGCAGCAGTTGTTTTACCACCACCTGCGAAGACTCCTTCTCCTCCGGCCCCTGTTGCTTCACCACCACCAGAAGCACCAGTGAGCTCTCCACAGCCTCAAGTAAAATCACCTCCTCCA  2098
 P  A  T  P  V  K  S  S  P  P  P  A  A  V  V  L  P  P  P  A  K  T  P  S  P  P  A  P  V  A  S  P  P  P  E  A  P  V  S  S  P  Q  P  Q  V  K  S  P  P  P    566
CCAGCACCGGTGGCCTCACCACCACCACCCATGAAGTCGCCTCCCCCACCAGCGCGAGTGGCATCACCTCCGCCGCTGATGAAATCACCTCCACCACCAGCACCAGTGGCCTCTCCACCCCAACCGTTGAAATCTCCACCACCGCCAGTA  2248
 P  A  P  V  A  S  P  P  P  P  M  K  S  P  P  P  P  A  R  V  A  S  P  P  P  L  M  K  S  P  P  P  P  A  P  V  A  S  P  P  Q  P  L  K  S  P  P  P  P  V    616
CTGTGGCTCTCCACTCCATCGGTGAAATCTCCACCACCGCCGGTACCAGTGGCCTCTCCACCTCCACCGGTGAAATCTCCACCACCGCTGGCTCCAGTGAGCTCACCTTCTCCTCCGGTGAAGTTGCCACCACTACCAGCTCCTGGGAAA  2398
 L  W  L  S  T  P  S  V  K  S  P  P  P  P  V  P  V  A  S  P  P  P  P  V  K  S  P  P  P  L  A  P  V  S  S  P  S  P  P  V  K  L  P  P  L  P  A  P  G  K    666
TCAACCCCACCTCCAGAAGAGGAAAAACCAACACCACCTACTCCAGTAAAGTCATCGCCGCCTCCAGAGAAGTCTCTACCTCCACCCACTCTGACTACCTCTCCACCACCACAAGAGAAGCCTACTCCTCCATCAACTCCATCCAAACCA  2548
 S  T  P  P  P  E  E  E  K  P  T  P  P  T  P  V  K  S  S  P  P  P  E  K  S  L  P  P  P  T  L  T  T  S  P  P  P  Q  E  K  P  T  P  P  S  T  P  S  K  P    716
CCACCACCATCTCCAGTTGAGACACTACCTCCTCCATCAAAGTCTTCTCCACCGGAAGAACCAGTTAGCTCACCACCTCAAGCACCAAAGTCGTCATCTCCACCAGCTCCTGTAAGCTCACCACCTCCACTGAAGTCTTCTCCGCCTCCA  2698
 P  P  P  S  P  V  E  T  L  P  P  P  S  K  S  S  P  P  E  E  P  V  S  S  P  P  Q  A  P  K  S  S  S  P  P  A  P  V  S  S  P  P  P  L  K  S  S  P  P  P    766
GTACCAGAGAGCTCACCACCACCGACACCTAAGTCATCACCTCCACTGGCTCCAGTGAGCTCACCACCTCAAGTGGAGAAGACTTCTCCACCACCAGCACCAGTCAGCTCACCACCACCGACACCTAAGTCATCACCTCCACTAGCTCCA  2848
 V  P  E  S  S  P  P  P  T  P  K  S  S  P  P  L  A  P  V  S  S  P  P  Q  V  E  K  T  S  P  P  P  A  P  V  S  S  P  P  P  T  P  K  S  S  P  P  L  A  P    816
GTGAGCTCACCACCTCAAGTAGAGAAGACTTCTCCACCACCAGCACCAGTCAGCTCACCACCACCAACACCTAAGTCATCACCTCCACTAGCTCCAGTGAGCTCACCACCTCAAGTAGAGAAGACTTCTCCACCACCAGCACCAGTCAGC  2998
 V  S  S  P  P  Q  V  E  K  T  S  P  P  P  A  P  V  S  S  P  P  P  T  P  K  S  S  P  P  L  A  P  V  S  S  P  P  Q  V  E  K  T  S  P  P  P  A  P  V  S    866
TCACCACCATTGGAGCCTAAGCCATCATCTCCACCTAGTTCTGTGAGCTCTCCACCTACGACGGTGAAGTCTTCTCCACCACCAGCACCACTTAGCTCGCCACCAATGACGCCTAAGTCATCAAGCCCACCTGCACACGTGAGCTCGCCT  3148
 S  P  P  L  E  P  K  P  S  S  P  P  S  S  V  S  S  P  P  T  T  V  K  S  S  P  P  P  A  P  L  S  S  P  P  M  T  P  K  S  S  S  P  P  A  H  V  S  S  P    916
CCTGAAGCAGAGAAGTCATCCCCGCCACTGGCACCGATCAGCTCACCGCCATCAGAGCCCAAGTCACCGTCGTCACCTCCAATGGTAGAGAAAACTTCTCCACCACCGGCAACGGTTAGCTCACCACCTACGCCTAAGTCGTCACCT  3298
 P  E  A  E  K  S  S  P  P  L  A  P  I  S  S  P  P  S  E  P  K  S  P  S  S  P  P  M  V  E  K  T  S  P  P  P  A  T  V  S  S  P  P  P  T  P  K  S  S  P    966
CCACCTGCTCCTGTGAGCTCACCACCTCCAGTGGTGAAGTCTTCTCCACCACCGGCACCAGTCAGCTCACCACCACCGACACCTAAGCCATTACCTCCACCTGCTCCTGTGAGCTCACCACCTCCGGTGGTGAAGTCTTCTCCACCGCCG  3448
 P  P  A  P  V  S  S  P  P  P  V  V  K  S  S  P  P  P  A  P  V  S  S  P  P  P  T  P  K  P  L  P  P  P  A  P  V  S  S  P  P  P  V  V  K  S  S  P  P  P   1016
ACACCAGTCAGCTCACCACCACCGACACCTAAGCCATTGCCTCCACCGCCGACACCAGTCAGCTCACCACCACCGACACCTAAGCCATTGCCTCCACCTGCTCCTGTGAGCTCACCACCTCCAGTGGTGAAATCTTCTCCACCACCGGCA  3598
 T  P  V  S  S  P  P  P  T  P  K  P  L  P  P  P  P  T  P  V  S  S  P  P  P  T  P  K  P  L  P  P  P  A  P  V  S  S  P  P  P  V  V  K  S  S  P  P  P  A   1066
CCAGTCAGCTTACCACCACCGACGCGTAAGCCATCTCCACCACGTACTCGTGTGAGCTCACCACGTCCAGTAGTAAAGTGTTGTCCACCACCAACACTAGTCAGCTCACCACCACCAGCACCTAAGTCATTACCTCCACCTACTCCCGTG  3748
 P  V  S  L  P  P  P  T  R  K  P  S  P  P  R  T  R  V  S  S  P  R  P  V  V  K  C  C  P  P  P  T  L  V  S  S  P  P  P  A  P  K  S  L  P  P  P  T  P  V   1116
AGCTCGCCGCCTCCAGAAGTGAAGTCCTCTCCTCCACCAACCCCAGTTAGCTCACCACCACCAGCGCCGAAGTCGTCACCTCCACCTACTCCAGTGAGCTCGCCACCTCCAGAACTGAAGTCCTCTCCACCACCAGCCCCAGTTAGCTCA  3898
 S  S  P  P  P  E  V  K  S  S  P  P  P  T  P  V  S  S  P  P  P  A  P  K  S  S  P  P  P  T  P  V  S  S  P  P  P  E  L  K  S  S  P  P  P  A  P  V  S  S   1166
CCACCATCAGCGCCAAAGTCGTCACCTCCACCTGCTCCTGTAAGCTTGCCACCTCCTGAAGTAAAGTCTTCCCCACCACCAGCACCGATTAGCTCTCCACCACCTCCGGCCAAGTCACCACCTCCACCGGCCCCTATGAGCTCACTTCCT  4048
 P  P  S  A  P  K  S  S  P  P  P  A  P  V  S  L  P  P  P  E  V  K  S  S  P  P  P  A  P  I  S  S  P  P  P  P  A  K  S  P  P  P  P  A  P  M  S  S  L  P   1216
CCCCCTGTGAAGTCTCCACCTCCACCAGCGCCGGTCAGTTCGCCACCACCTCCAATGAAGTCCCCACCGCCGCCAGCACCAATTAGCTCTCCGCCACCAGCACCAGTGAAGCCACCATCGCTACCACCACCAGCCCCGGTAAGCTCACCT  4198
 P  P  V  K  S  P  P  P  P  A  P  V  S  S  P  P  P  P  M  K  S  P  P  P  P  A  P  I  S  S  P  P  P  A  P  V  K  P  P  S  L  P  P  P  A  P  V  S  S  P   1266
CCTCCTGCGGTCACCTCTGCCCCACCAAAGAAAGAAGAAGATTCAACGGCACCACCGGCAGAAGCCCTACCTCCACCTTCATTCAACGACATCATCCTTCCACCTATCATGGCCAACAAATACGCATCTCCGCCTCCTCCTCTATTCCAA  4348
 P  P  A  V  T  S  A  P  P  K  K  E  E  D  S  T  A  P  P  A  E  A  L  P  P  P  S  F  N  D  I  I  L  P  P  I  M  A  N  K  Y  A  S  P  P  P  P  L  F  Q   1316
GGATATTAAGCTCTATGGAGAGATGGTCGACCNAGAATGAATTGAACAACAATATNTAAGGTCGCATGCGCAACCAAACAGCTACATGTACATTGCAGGCGTGTGTGAAATTGTTTATTATTATTTTTGCATGTGCACGCGTGACTATTTG  4498
 G  Y  *                                                                                                                                                   1318
```

Figure 1. Sequence of mPex2 gene and encoded protein. TATA box is in bold. Arrow shows site of putative signal peptide cleavage. Repeats of SP_{3-6} are underlined. Insert relative to mPex1 is highlighted.

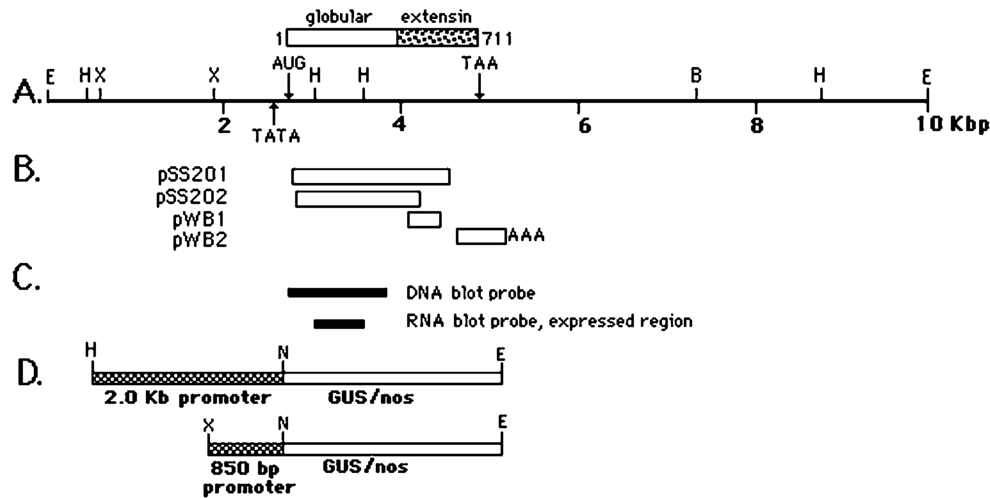

Figure 2. Schematic representation of genomic tPex clone and tPex protein, tPex cDNA clones, probes, and tPex promoter-reporter gene fusions. A. Restriction map of pSS100 genomic clone, with scale in kb. tPex protein is shown above aligned with reading frame, with numbers of amino acids. B. Relative position of cDNA clones isolated from tomato anther cDNA library. C. Probes used in DNA and RNA blot analysis. Both contain only globular domain-encoding sequences, and the fragment used as an RNA blot probe was also used in a bacterial expression vector to obtain antibodies D. Promoter-GUS/nos fusions used in transient and stable expression in pollen. E, *Eco*RI; H, *Hin*dIII: X, *Xba*I; B, *Bam*HI: N, *Nco*I.

Tomato genomic DNA blots

Southern analysis was performed with tomato nuclear DNA of cv. Cal-Ace. The probe consisted of a 824 bp fragment from 34 bp upstream of the AUG start codon to near the end of the LRR-encoding region; thus the probe did not contain any extensin domain-encoding sequences (Figure 2C). The probe was labeled with ^{32}P by random priming (Rediprime kit, Amersham). Hybridization was performed in 6× SSC, 5× Denhardt's reagent, 0.5% SDS and 100 μg/ml single-stranded salmon sperm DNA for 18 h at 61 °C. Blots were washed 3 times in 0.1% SDS, 0.1× SSC at 61 °C before exposing to film.

Tomato RNA isolation and RNA blots

Total RNA was isolated from immature flower, mature anther, pollen, pistil, root, leaf, immature fruit, sepal and petal by using a guanidium thiocyanate method (Logeman *et al.*, 1987). 15 μg of total RNA per lane was electrophoresed and blotted according to standard methods (Sambrook *et al.*, 1989). The blot was UV-crosslinked (Stratalink 2400, Stratagene), then probed with a ^{32}P random-prime-labeled 477 bp *Hin*dIII fragment from subclone pSS101, representing part of the globular domain-encoding part of the tPex gene (Figure 2C). Hybridization was carried out at 65 °C. The blot was washed three times with 0.3× SSPE and 0.1% SDS at room temperature for 15 min before exposing to film.

Fusion protein construction and expression

A 477 bp *Hin*dIII fragment of pSS101 (same as RNA blot probe) was cloned into the pET21a expression vector (Novagen). This fragment corresponds to the region of the tPex gene that encodes amino acids 45–204 of the tPex protein, highlighted in Figure 2. In order to generate antibodies that may recognize Pex proteins in a variety of plant species, the variable, putative species-specific region was not included in the expressed protein. Once the correct orientation and frame of the insert were confirmed, the construct was designated pET21a(SS101). pET21a(SS101) was transformed into *Escherichia coli* BL21 (DE3) (Novagen) in order to induce production of the tPex fragment (BEtPex). A protein of ca. 26 kDa, corresponding to the expected size for bacterially expressed tPex protein, was detected by SDS-PAGE.

Inclusion bodies were purified to recover the BEtPex protein. Cells were resuspended in 10 ml of inclusion body solution (50 mM Tris HCl pH 8, 10 mM EDTA, 1 mM PMSF, 0.1% BME). To this, 200 μl of 10 mg/ml lysozyme and 1 ml of 10% Nonidet P-40 were added. After 30 min of incubation at 30 °C, the cells were sonicated three times for 10 s each. The sonicate was centrifuged at 18 500 × *g* for 20 min. The

supernatant was discarded and the pellet was resuspended in inclusion body solution. This process was repeated 5 times. The final pellet containing inclusion bodies was resuspended in 1.4 ml solubilization buffer (8 M urea, 0.5 M NaCl, 20 mM Tris-HCl pH 7.9, 1 mM EDTA, 10 mM 2-mercaptoethanol).

Antibody production

0.8 mg of BEtPex protein was isolated from SDS-PAGE gels fixed in 50% methanol and 10% acetic acid. Another 0.2 mg of BEtPex protein was isolated from an SDS-PAGE gel stained in water in an attempt to preserve a variety of native tPex epitopes. Protein samples were pooled and sent to Macromolecular Resources (Colorado State University) for the generation of antibodies in rabbits. One rabbit produced serum that reacted at a 1:10 000 dilution with BEtPex and a pollen protein of ca. 200 kDa.

Protein blotting analysis

All protein samples were prepared by grinding tissues with a pestle in the presence of 37 °C 2× sample buffer (0.1 M Tris pH 6.8, 20% glycerol, 4% SDS, 2% 2-mercaptoethanol, 0.001% bromophenol blue) and proteinase inhibitors (leupeptin, pepstatin, and chymostatin, 5 μg/ml each, final concentration). Crude extracts were boiled three times for 5 min, centrifuged to remove insoluble material, and electrophoresed on a 12% polyacrylamide gel for 1.5 h at 100 V. Kaleidoscope prestained standards (BioRad) were used to assess the apparent molecular mass of electrophoresed proteins. Proteins were electroblotted onto a nitrocellulose membrane for 1 h at 100 V. The membrane was blocked in 5% dry milk overnight then washed and incubated in primary antibody (1:10 000 dilution in incubation buffer; 3% non-fat milk, 0.1% Tween-20, 13 mM Tris pH 7.5, 0.15 M NaCl) for 1 h. The membrane was washed then incubated for 1 h in the secondary antibody (goat anti-rabbit alkaline phosphatase conjugate) diluted 1:5000 in incubation buffer. The membrane was washed, then rinsed in 10 ml developing buffer (100 mM Tris pH 9.5, 100 mM NaCl, 5 mM $MgCl_2$). NBT (0.375 μg/ml) and BCIP (75 μg/ml) were added to the buffer for detection.

Transient transformation of tobacco pollen with promoter-GUS/nos fusions

Putative promoter regions of the tPex gene were fused to the β-glucuronidase (GUS) gene of *Escherichia coli* and the *Agrobacterium* nopaline synthase (nos) 3′ region from pLAT52-7 (Twell *et al.*, 1991). Short promoter constructs contained a *Xba*I-*Nco*I fragment consisting of 850 bp of the 5′ region of the tPex gene to 12 bp 3′ of the putative TATA box (Figure 2D). Long promoter constructs contained a *Hin*dIII-*Nco*I fragment ca. 2 kb of the 5′ region of the tPex gene to 12 bp 3′ of the putative TATA box.

In order to assay transient expression of the tPex promoter-GUS fusions in pollen, 10 μg of each tPex promoter-GUS plasmid construct was co-bombarded with 10 μg of a reference plasmid, LAT59-LUC into a monolayer of tobacco cv. Samsun pollen as previously described (Twell *et al.*, 1991). A promoterless GUS construct and a well-characterized tomato pollen promoter fused to GUS (pLAT52-7) (Twell *et al.*, 1991) were used as negative and positive controls, respectively. Three independent bombardments were performed for each construct. After 15–17 h incubation at room temperature, pollen was extracted and assayed for both luciferase (Amersham Luciferase Assay Kit) and GUS (Twell *et al.*, 1991). Units of GUS were normalized to units of luciferase to control for variation in transformation efficiency of each experiment.

Stable transformation and expression studies

Transgenic tomato plants (cv. VF36) containing either the long or short tPex promoters fused to the GUS gene were generated as described (McCormick, 1991) and grown under standard greenhouse conditions. GUS activity was measured in plant extracts essentially as described (Jefferson *et al.*, 1987) with some modifications. Pollen from a single flower of each of the thirteen transgenic plants containing the short tPex promoter-GUS transgene, as well as from two untransformed VF36 plants were collected and ground in extraction buffer with a small glass mortar and pestle on ice. Leaf, root and reproductive tissues (pistil, petal and sepal tissue from a single flower and pollen from a single flower) from five randomly chosen transgenic plants containing the short tPex promoter-GUS fusion construct and from two untransformed VF36 plants were processed in the same manner. Pollen from nine transgenic plants containing the long tPex promoter-GUS transgene and from one untransformed VF36 plant were collected and extracted in the same manner as in the previous experiments. All samples were assayed in duplicate for each experiment. The extracts were centrifuged at 14 000 rpm for 15 min at 4 °C in a microfuge

```
 TCTAGATGGTGTTAAACTGATGCCTTTTCCTTCAGTTTCCTTGAAGTACATACACAAAATACATCACAATAAAA      74
GAGCTTGGTGCACGAAGCATCTTACATATTCCGTANAGTTTGGTAAAGTAAAGTATCTCACTCGGAGGGATATGA     149
TGTGTATGCAATTATGTAAGTATCAGTGCTGACTGACTTCCTCCAAATCCCTATTTTTTCTGGTAAATCAGTTGC     224
AATTACGGCCATAATTTCTTTGAATAGCATAAGCCTTATCTTTATGTTTACTTTGATTACTGTCCAAACGTGAAT     299
ACATCCACTTTCCACTGTCACTGGCATCATGCCTAACTGTATCACAAAAGCCATTACAAATATTTCCCTTTCTCG     374
ATTCCTATTTCATAGGTACAATGTTTGAATCAATAGATTATCTTTTTTCTACCTGAACTGATATTATTCGATTAG     449
TAGTTTGTGTCATTTTAACTGTGAATTTAAGAATAAACAAGAGTAAGTCACTTGGAACCTTAATAATGGAAAAAT     524
CAACCAATTATTACTAGCTTAAAGAAAAACAAAGCATGTTTATGAAAATAATGTCCGTCTGTTTTAGACGGTAAA     599
TGAACTTGTGCAACTGGTGAATTTCCATTCTAATAGGTAAAAGANGGATCAAGAATTAACTCTGTTCACTCAAAT     674
TTTTAATTCTAAGAAACTCAACAAACCGCTAAAATTTGTTTCAAGTTTATTTTGACCCTCCAAATTCTATGGTTC     749
TCAAAATTCATGACAGCTTATCAACCATTTAATTTCCTTCCTTTGCCTACATATATAATGCCCCCCATTTTGCTT     824
TGAAAAATATATTGCAATAAGCAGCCTTGGGCCTTACCCACAGGTGGACTTACAAACATGGTGGCCTTGGGTTGC     899
                                                             M  V  A  L  G  C      6
TTTCTCTTTTTCCTTGTATCTTTCTGTTCTTTTTCTCCATCATATTTTGCATTATCTGATATCGAAGCTGCTTCT     974
 F  L  F  F  L  V  S  F  C▼ S  F  S  P  S  Y▼ F  A  L  S  D  I  E  A  A  S     31
ATTGCCCGTCGCCAACTTTTAAGTAACAATGGCCAGCTTTCAAATACATACGAGTCTGAAATGACCATTAACATG    1049
 I  A  R  R  Q  L  L  S  N  N  G  Q  L  S  N  T  Y  E  S  E  M  T  I  N  M     56
AAATTTGAAAATGCCAGGTTGAAGAAAGCTTATGTTGCTCTTCAGGCATGGAAGAAGTCTATTTACTCTGATCCA    1124
 K  F  E  N  A  R  L  K  K  A  Y  V  A  L  Q  A  W  K  K  S  I  Y  S  D  P     81
ACAAACTTCACAGCTAATTGGGAGGGTAGCAATGTTTGTGCCTATAACGGTGTCTTTTGTGACAATGCTCTTGAT    1199
 T  N  F  T  A  N  W  E  G  S  N  V  C  A  Y  N  G  V  F  C  D  N  A  L  D    106
GATCCCAACATCTCTGTTGTTGCTGGGATTGATCTTAACCATGCAGACATAGCTGGACATCTCCCTGTGGAGCTT    1274
 D  P  N  I  S  V  V  A  G  I  D  L  N  H  A  D  I  A  G  H  L  P  V  E  L    131
GGCCTCTTGGCTGATGTGAGTCTCATCCACATTAATTCTAACAGATTTTGTGGAATCATTCCCAAGAGCATTACA    1349
 G  L  L  A  D  V  S  L  I  H  I  N  S  N  R  F  C  G  I  I  P  K  S  I  T    156
AATCTTACCCTTTTGGATGAGATTGATTTCAGCAACAACCGTTTCGTGGGGCCATTCCCTGATGTTGTGCTTGAC    1424
 N  L  T  L  L  D  E  I  D  F  S  N  N  R  F  V  G  P  F  P  D  V  V  L  D    181
TTGCCTAAGCTCAACTATCTTGATCTTAGGTTCAACGACTTTGAAGGTCAAGTACCTTCCGCCCTATTTGAGAAG    1499
 L  P  K  L  N  Y  L  D  L  R  F  N  D  F  E  G  Q  V  P  S  A  L  F  E  K    206
AATCTTGATGCAATATTGATCAACAATAACCGATTCCATTCCACCATCCCTGAAAGCTTGGGCAATTCCAATGCT    1574
 N  L  D  A  I  L  I  N  N  N  R  F  H  S  T  I  P  E  S  L  G  N  S  N  A    231
TCCGTGGTGGTATTGGCCAATAACAAGTTCTATGGTTGCATCCCGAGCAGTATAGGTAAAATGGGCAACTCATTA    1649
 S  V  V  V  L  A  N  N  K  F  Y  G  C  I  P  S  S  I  G  K  M  G  N  S  L    256
GATGAGCTTGTATTCACCAACAACGAGCTCTCTGGTTGTTTGCCTGAAGAGATTACTAAGCTCACGAGCCTAACA    1724
 D  E  L  V  F  T  N  N  E  L  S  G  C  L  P  E  E  I  T  K  L  T  S  L  T    281
CTGCTTGATATTAGTGGAAATAAGTTTGTTGGATCGTTGCCTCAGGATTTGAAATCGATGCAGAAAGTTGAAATC    1799
 L  L  D  I  S  G  N  K  F  V  G  S  L  P  Q  D  L  K  S  M  Q  K  V  E  I    306
TTTGACATTGCGTCGAACAAGTTCATGGGAAATGTTCCAAAGAACCTTTGCACGTTGCCTAGTTTGAAGAATTTC    1874
 F  D  I  A  S  N  K  F  M  G  N  V  P  K  N  L  C  T  L  P  S  L  K  N  F    331
ACATTCTCAAAGAACTACTTTGAGAGCATGGATGAAACATGTAGGCCATCAGAGTCGAAGCAAGTTAAGATAGAT    1949
 T  F  S  K  N  Y  F  E  S  M  D  E  T  C  R  P  S  E  S  K  Q  V  K  I  D    356
GGTAATGAGAACTGTTTGGGAGGAAGATCAGAGCAGAGGACGGAGAAAGAGTGTTTCCCAGTAGTGAGCAAGCCT    2024
 G  N  E  N  C  L  G  G  R  S  E  Q  R  T  E  K  E  C  F  P  V  V  S  K  P    381
GTTGATTGCAGTAAAGGTCATTGCGGAGTTTCAAGGGAAGGGCAGAGTCCAAAAGATCCACCTAAAACCGTGACT    2099
 V  D  C  S  K  G  H  C  G  V  S  R  E  G  Q  S  P  K  D  P  P  K  T  V  T    406
CCTCCAAAACCTTCAACACCAACCACGCCAAAACCAAATCCATCTCCACCACCACCAAAAACCTTGCCTCCACCA    2174
 P  P  K  P  S  T  P  T  T  P  K  P  N  P  S  P  P  P  P  K  T  L  P  P  P    431
CCACCAAAAACCTCGCCTCCACCTCCTGTCCACTCACCACCACCACCACCGGTAGCATCACCTCCCCCCCCCGTG    2249
 P  P  K  T  S  P  P  P  P  V  H  S  P  P  P  P  P  V  A  S  P  P  P  P  V    456
CACTCACCACCACCACCAGTAGCATCACCTCCACCTCCCGTCCACTCACCACCACCACCACCAGTAGCATCACCT    2324
 H  S  P  P  P  P  V  A  S  P  P  P  P  V  H  S  P  P  P  P  P  V  A  S  P    481
CCACCTCCTGTCCACTCACCACCACCACCGGTAGCATCACCTCCCCCTCCCGTCCACTCACCACCACCTCCAGTT    2399
 P  P  P  V  H  S  P  P  P  P  V  A  S  P  P  P  P  V  H  S  P  P  P  P  V    506
CACTCACCACCACCACCAGTAGCATCACCTCCACCTCCAGTTCACTCACCACCACCTCCCGTTCACTCACCTCCA    2474
 H  S  P  P  P  P  V  A  S  P  P  P  P  V  H  S  P  P  P  P  V  H  S  P  P    531
CCTCCAGTCCACTCACCTCCACCTCCTGTTCACTCACCACCACCTCCCGTTCACTCACCACCACCACCGGTAGCA    2549
 P  P  V  H  S  P  P  P  P  V  H  S  P  P  P  P  V  H  S  P  P  P  P  V  A    556
TCACCTCCACCTCCAGTCCACTCACCTCCACCTCCTGTCCACTCACCACCACCTCCCGTCCACTCACCACCACCA    2624
 S  P  P  P  P  V  H  S  P  P  P  P  V  H  S  P  P  P  P  V  H  S  P  P  P    581
CCAGTGGCCTCACCTCCACCTCCCGTCCACTCACCACCACCACCACCACCAGTAGCATCACCTCCACCTCCTGTC    2699
 P  V  A  S  P  P  P  P  V  H  S  P  P  P  P  P  P  V  A  S  P  P  P  P  V    606
CACTCACCACCACCACCAGTGGCCTCACCTCCACCTCCCGTCCACTCACCACCACCACCGGTAGCATCACCTCCA    2774
 H  S  P  P  P  P  V  A  S  P  P  P  P  V  H  S  P  P  P  P  V  A  S  P  P    631
CCTCCAGTCCACTCACCACCACCACCAGTCCATTCACCACCACCACCAGTCCACTCACCTCCACCACCAGTCGCC    2849
 P  P  V  H  S  P  P  P  P  V  H  S  P  P  P  P  V  H  S  P  P  P  P  V  A    656
TCACCTCCACCTGCTCTGGTATTCTCTCCTCCACCGCCAGTCCACTCTCCACCACCACCAGCTCCCGTAATGTCT    2924
 S  P  P  P  A  L  V  F  S  P  P  P  P  V  H  S  P  P  P  P  A  P  V  M  S    681
CCACCTCCACCCACTTTTGAGGATGTCGCTCTTCCCCCAACATTGGGCTCGCTATACGCCTCACCACCACCACCA    2999
 P  P  P  P  T  F  E  D  V  A  L  P  P  T  L  G  S  L  Y  A  S  P  P  P  P    706
ATTTTCCAAGGTTATTAAGAGGATCGCCATTNTGATTTCAATGGCACAAAAACTACATAAAAAATGTGTACCAATC    3074
 I  F  Q  G  Y  *                                                                711
AATCTGTAAACACACATATCAACTGNAAAAACAAACATTTACATTATCAATTGTACACTGAAAAATTGGAAATTT    3149
GCTATTAGATGGTGCTGATCATTNTACAGTTTCAACATATTTCAGTATTA                             3199
```

Figure 3. Sequence of tPex gene and encoded protein. Promoter elements Lat56/59 (TGAANTTGTG) and TATA box are in bold. Arrows show two possible sites of signal peptide cleavage. The peptide expressed in bacteria for antibody production is highlighted. Repeats of SP_{3-6} are underlined.

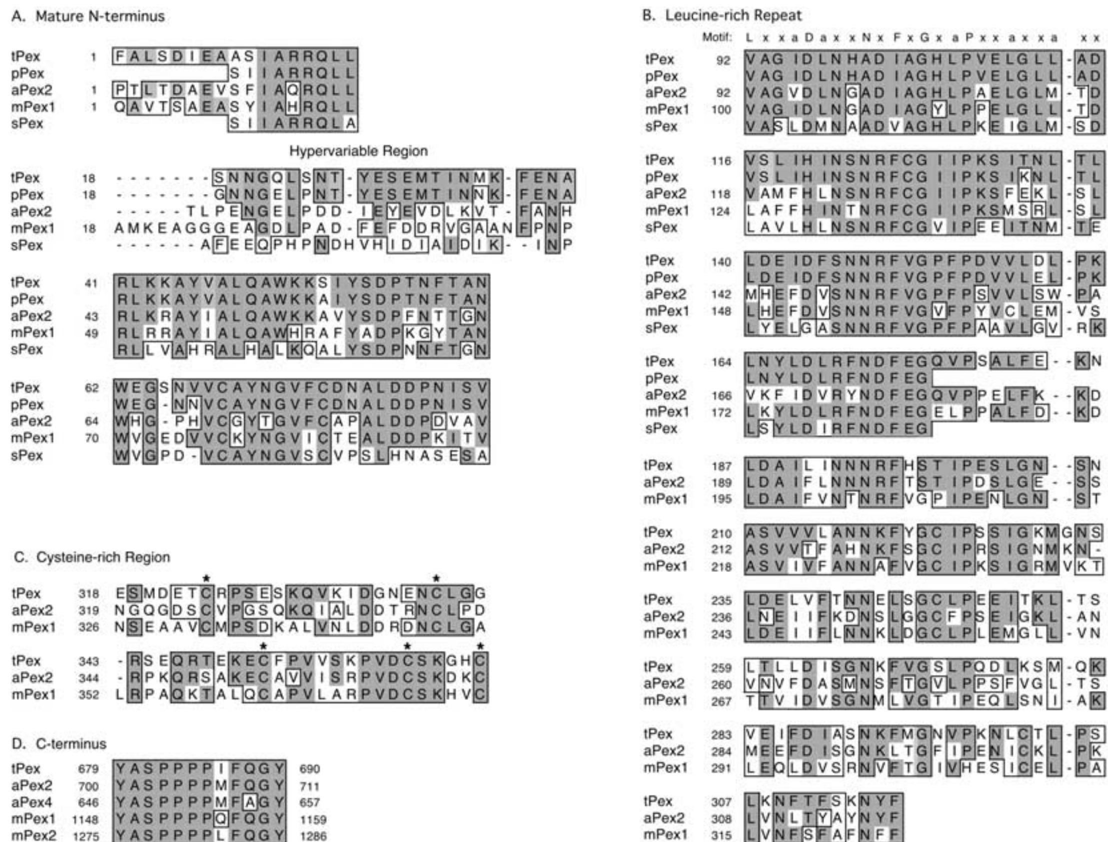

Figure 4. Clustal X alignment of amino acids in tPex protein, mPex1 protein, potential *Arabidopsis* Pex protein encoded by gene on chromosome 2 (aPex2), and regions of potential potato (pPex) and sorghum (sPex) Pex proteins as identified by genomic PCR. Note that numbering reflects amino acid residues in the mature proteins. Amino acids identical to tPex amino acids are shaded, similar amino acids are included in boxes. A. N-terminal region of mature proteins. B. LRRs, with the consensus for the Pex protein LRRs at the top C. Cysteine-rich regions, with conserved cysteine residues marked with * D. C-terminal amino acids of tPex, aPex2, aPex4, mPex1 and mPex2 proteins.

tube and a 100 μl aliquot of the supernatant was assayed with 1 mm 4-methylumbelliferyl glucuronide at 37 °C for 1 h. Fluorescence of the reaction product, 4-methylumbelliferone, was measured with a Hoefer TKO fluorometer (Hoefer Scientific Instruments, San Francisco, CA) using excitation and emission wavelengths of 365 nm and 460 nm respectively. The specific activity of GUS in each sample from the transgenic plants was calculated as nmol 4-MU per hour per mg protein. Protein concentration in plant extracts was determined with Bradford reagent (BioRad, Hercules, CA) according to the manufacturer's instructions.

Results

The two maize Pex genes

Previous genomic DNA blot studies and cDNA analysis indicated that there are two Pex genes in the maize genome (Rubinstein *et al.*, 1995a, b). RNA and protein blots indicated that both were expressed exclusively in pollen. The complete sequence of one gene (mPex1) has been published (Rubinstein *et al.*, 1995a, b). To compare the two maize genes, the genomic version of mPex2 was cloned and sequenced (Figure 1). The amino acid sequence of the proteins encoded by the two maize genes, mPex1 and mPex2, are highly conserved. Both contain a signal sequence, an N-terminal 'globular' domain including a leucine-rich repeat (LRR) with an adjacent cysteine-rich region, a transition zone, and an extensin-like domain. The entire N-terminal 'globular' domain is highly conserved,

being 96% identical and 97.4% identical or similar (Figure 4). In contrast, the transition and extensin-like domains are less conserved (71–75% identical and 89–90% similar) and the mPex2 protein has a 115 amino acid insert relative to mPex1 at the transition zone/extensin-like domain junction.

A dicot Pex gene

The Pex proteins of maize have been proposed to have a role in plant reproduction, either as a structural component of the pollen tube wall, or as a pollen-pistil recognition protein. Although monocots and dicots were separated evolutionarily 200 million years ago, pollen in both monocots and dicots must perform very similar functions to deliver the sperm to the egg (Bedinger, 1992). It was therefore of interest to determine whether a homologous gene exists in a dicot and, if so, to compare the gene sequences to ascertain which regions are highly conserved and thus may be functionally important. A partial cDNA encoding an extensin-like protein was identified in a tomato anther cDNA library using a maize Pex extensin domain gene region as a probe, and a genomic clone corresponding to the cDNA was identified (Figure 2A). Nucleotide sequencing revealed an open reading frame of 711 codons (Figure 3). Genomic sequences were used to obtain additional cDNAs encoding almost the entire tomato Pex (tPex) reading frame (Figure 2A, B). The tomato gene, like the maize genes, contains no introns. While the maize and tomato genes encode signal sequences, allowing for secretion of the proteins, there is no similarity between the signal sequences of the dicot and monocot proteins (as opposed to 97% identity between the two maize proteins). The predicted signal peptides of the maize proteins are 29 amino acids in length, while the tomato protein signal peptide is either 16 or 20 amino acids according to the von Heijne (1986) scoring system.

Other potential Pex genes

Two potential Pex *Arabidopsis* genomic sequences (one on chromosome 2, accession number AC006438; one on chromosome 4, accession number CAA19879) were identified by searching by tBLAST the available *Arabidopsis* genomic sequence database (AtDB). A third potential *Arabidopsis* Pex gene on chromosome 1 was also identified (accession number AC007504), but this gene lacked the conserved N-terminal sequences including (A/S)XIA(R/Q)RQLL found in other Pex genes. The entire N-terminal globular domains of the aPex2 and aPex4 proteins were highly conserved (84.6% identical, 91% identical or similar). In addition, conserved regions of mPex and tPex were used to design primers for PCR amplification and cloning of Pex-like sequences from potato and sorghum genomic DNA. The amino termini of tomato Pex protein (tPex), *Arabidopsis* chromosome 2 putative Pex protein (aPex2), maize protein (mPex1) and PCR-identified regions of the potato (pPex) and sorghum (sPex) putative Pex proteins were aligned and compared by Clustal X (Jeanmougin *et al.*, 1998) as shown in Figure 4.

Amino-terminal region

Adjacent to the putative site of signal peptide cleavage, similarity is seen among all known and putative Pex proteins; in particular the sequence IA(R/Q)RQLL is highly conserved (Figure 4A). After this sequence, there is a hypervariable region (23 amino acids in tomato and potato, 24 in sorghum, 25 in *Arabidopsis*, 31 in maize). The hypervariable region is conserved within a species (30/31 identical amino acids between the two maize genes, 22/25 between the two *Arabidopsis* genes, not shown) and between two members of the Solanaceae family (19/23 identical amino acids between tomato and potato). All other pairwise matchings between species are highly variable. It is notable that this hypervariable region has the highest ‘surface probability’ (Emini *et al.*, 1985) within the globular domain of the Pex proteins, and is sandwiched between two predicted regions of α-helix (Chou and Fasman, 1978; McLelland and Rumelhart, 1988; Kneller *et al.*, 1990).

Leucine-rich repeats

An LRR region of 225 amino acids is found in all known and putative Pex proteins (Figure 4). LRRs function as protein-ligand (often protein-protein) interaction domains (Kobe and Deisenhofer, 1994, 1995). The general repeat sequence of the LRR superfamily is LxxLxLxxNx-axxxaxxxxaxxxx. In the Pex proteins, each repeat is 23–25 amino acids long and has the consensus sequence L/VxxaDaxxNxFxGxaPxxaxxaxx, where a is an aliphatic amino acid and x is a variable amino acid. This repeat sequence is very similar to those of the tomato pollen receptor-like protein kinases and several other plant LRR-containing proteins, where aliphatic amino acids other than leucine can occupy the leucine positions in the repeats (Muschietti *et al.*, 1998). There

are 9.5 repeats in each Pex protein. The vast majority (90%) of non-conservative amino acid differences within the LRRs of the Pex proteins are in the variable 'x' positions of the consensus sequence. This is consistent with the notion that the variable amino acids in LRRs are responsible for specific protein-ligand (often protein-protein) contacts, while the conserved amino acids confer a short β-sheet/β-turn-α-helix structure on each repeat (Kobe and Deisenhofer, 1994, 1995). It appears that the conservation of the position of aliphatic amino acids in the repeat as well as the Asp at position 9, the Gly at position 13 and the Pro at position 16 are especially important in preserving the structure of the LRRs of the Pex proteins.

Cysteine-rich region

Many LRRs, in particular those of extracellular proteins, have cysteine-rich regions adjacent to the LRR. The mPex, aPex and tPex proteins have a 51–54 amino acid cysteine-rich region C-terminal to the LRR (Figure 4C). Although the similarity of the three proteins in this region is not remarkably high, the positions of the five cysteines within this region are conserved.

Transition region

To the C-terminal side of the cysteine-rich region, the known and potential Pex proteins diverge to such a great extent that any alignments are difficult. All of the proteins contain a variable 'transition' region (31 amino acids for the tomato protein, 106 and 107 for the aPex proteins, 108 for the mPex1 protein, 111 for the mPex2 protein), containing a stretch of alternating single or double prolines.

Extensin domain

Extensins are cell wall proteins containing a repeated Ser-Pro$_n$(SP$_n$) motif (Lamport, 1965, 1977; Showalter, 1993; Cassab, 1998). The beginning of the extensin domain of the Pex proteins is defined as the first continuous set of SP$_n$ repeats. The extensin domain of the mPex1 protein has 53 of these SP$_{3-6}$ repeats, the mPex2 protein has 51 repeats, and the tPex protein has 37 repeats. The extensin domains of the mPex and tPex proteins are actually composed of larger repeats that contain two SP$_{3-6}$ repeats within them. These repeats include a Lys in the mPex proteins (SSPPPPAPVSSPPxxK) and a His in the tomato protein (SPPPPVASPPPPVH). In the potential Pex proteins of *Arabidopsis*, there is a repeat unit of SP$_{3-7}$V(H/Y/F). The aPex1 and aPex4 proteins have 25 and 18 SP$_{3-7}$ repeats, respectively. According to the analysis of extensin proteins by Kieliszewski (Kieliszewski and Lamport, 1994; Kieliszewski *et al.*, 1995), many of the prolines in all of the Pex proteins are predicted to be both hydroxylated and glycosylated. The maize Pex proteins are known to be glycosylated during pollen maturation (Rubinstein *et al.*, 1995a, b).

Carboxy terminus

The 12 carboxy-terminal amino acids of the tomato, *Arabidopsis* and maize Pex proteins are very highly conserved (Figure 4D). It is of interest to note that this C-terminal region contains the only tyrosine residues in the extensin domains of the tomato and mPex2 proteins (the mPex1 protein has one other tyrosine in its extensin domain, and the *Arabidopsis* proteins do contain a number of tyrosines in the extensin domains). Tyrosines have been proposed to be involved in the crosslinking of extensins to the cell wall (Fry, 1986; Waffenschmidt *et al.*, 1993). Consistent with this idea, the maize Pex proteins are tightly, perhaps covalently, associated with the pollen tube cell wall (Rubinstein *et al.*, 1995a, b). In any case, the high degree of conservation of the C-terminal amino acids in all of the known and potential Pex proteins suggests an important function for this region.

The tPex gene is a single-copy gene

Protein biochemistry, molecular cloning and genomic Southern blotting with extensin DNA probes indicate that the tomato genome contains multiple extensin genes (Showalter, 1993). In order to determine the copy number of the Pex gene(s) in tomato, genomic hybridization was performed using a portion of the globular domain-encoding region of the tPex gene as a hybridization probe, as shown in Figure 5. In all DNA digests performed, a single band was observed, as expected for a single-copy gene. In the *Eco*RI digest, the probe hybridized to a single ca. 10 kb fragment, and to a ca. 6.8 kb *Bam*HI/*Eco*RI fragment, as predicted from the restriction map of pSS100. A single hybridizing *Bam*HI fragment of about 16 kb was also detected. Since there is only one internal *Bam*HI site in the sequenced tPex gene, these results are consistent with the tomato genome containing only one Pex gene. In contrast, maize and *Arabidopsis* each have two known or potential Pex genes, perhaps not surprising considering that maize is an autotetraploid and that many

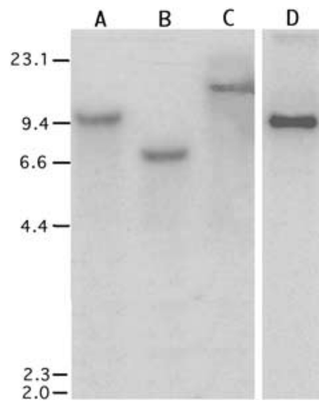

Figure 5. Tomato genomic DNA blots. 4 μg of genomic DNA from variety Cal-Ace was digested with: A, *Eco*RI; B, *Eco*RI/*Bam*HI; C, *Bam*HI; D, 25 pg of pSS100 digested with *Eco*RI. The hybridization probe was an 824 bp fragment representing globular domain-encoding sequences. Sizes of DNA markers in kb are shown to the left.

gene duplications in chromosomes 2 and 4 are being found in *Arabidopsis*.

Expression of the tPex gene

The maize Pex genes are expressed exclusively in pollen (Rubinstein *et al.*, 1995a, b). Although the tPex cDNAs were isolated from a tomato anther cDNA library, it was important to test the tissue specificity of the tPex gene, since various extensin genes are expressed in different organs of tomato (Showalter, 1993). As shown in Figure 6, RNA blotting and hybridization with the globular domain-encoding region of the tPex gene shows that the tPex gene is not expressed in any of the vegetative tissues tested, nor in immature flowers, sepals, petals, pistils or young fruits. tPex gene expression was detected in mature anthers containing pollen, and in pollen.

Part of the N-terminal and LRR regions of the tPex protein was expressed in bacteria and used to generate antibodies (see Figure 3). Protein extracts were made from the same tomato tissues as those for RNA blotting and electrophoresed on SDS-PAGE gels, blotted and probed with the tPex antibody. No specific signal was seen in any protein extracts except mature anther, pollen and pistil (Figure 7). A protein with an apparent molecular mass of about 200 kDa was detected in pollen, and likely represents the glycosylated form of the tPex protein. A large-molecular-mass protein was also detected in pistil extracts, which were made by using unpollinated pistils from flowers emasculated prior to anther dehiscence. No tPex expression in pistils

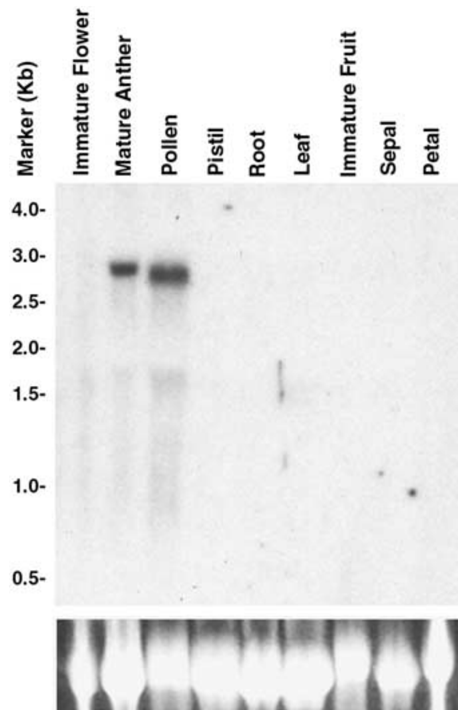

Figure 6. RNA blot analysis of tPex expression. 18 μg of total RNA from each indicated tissue was loaded per lane. Size markers in kb are to the left. Top panel: hybridization to tPex-specific probe representing globular domain sequences. Bottom panel: ethidium bromide stain to indicate loading of RNA in lanes.

was detected in RNA blots. We therefore propose that the pistil signal observed on protein blots represents recognition of a protein in pistil different from the Pex proteins at the gene level but with epitopes in common with the pollen Pex proteins at the protein level.

The tPex promoter: transient and stable expression of a reporter gene

To further confirm the expression of the tPex genes in pollen, 5′ genomic sequences were used to regulate transient and stable expression of a GUS reporter gene as shown schematically in Figure 2D. For transient expression, plasmids were co-bombarded into tobacco pollen with a well-characterized tomato pollen promoter fused to a luciferase reporter gene, which served as a transformation control (LAT59-LUC) (Twell *et al.*, 1991; Eyal *et al.*, 1995). Normalization of GUS to luciferase units yielded the expression data presented in Figure 8. While the short promoter of 850 bp produced a GUS signal 100-fold above background, the longer promoter of ca. 2 kb had an almost 3-fold higher level of expression, near

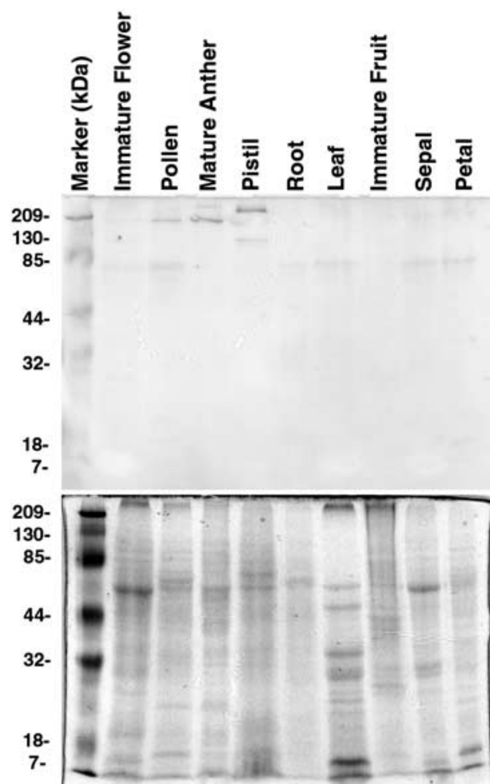

Figure 7. Protein blot analysis with antibodies to the tPex protein. Equal amounts of protein extracts from each tissue as indicated by Coomassie Blue staining (bottom panel) were electrophoresed, blotted and probed as described (top panel). Size markers in kDa are to the left. The faint signal at 85 kDa is non-specific, because it appears in secondary antibody-only control blots.

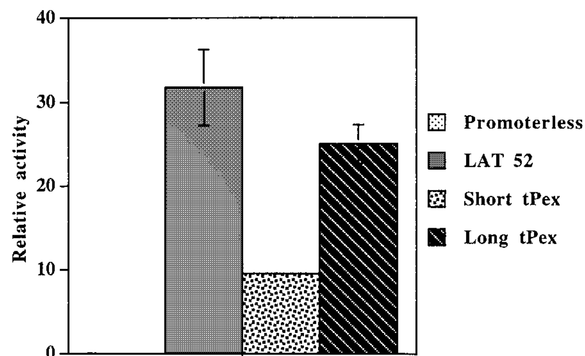

Figure 8. Transient expression in tobacco pollen. Plasmids with promoter sequences fused to the GUS gene were co-bombarded with plasmids containing the LAT59 promoter fused to the luciferase gene. Units expressed are GUS units relative to luciferase units. Constructs were pSS300 (promoterless GUS, negative control), pLAT52-7 (positive control), pSS301A (850 bp of tPex promoter region fused to GUS, short tPex), and pSS302 (2 kb of tPex promoter region fused to GUS, long tPex). Bars indicate one standard deviation.

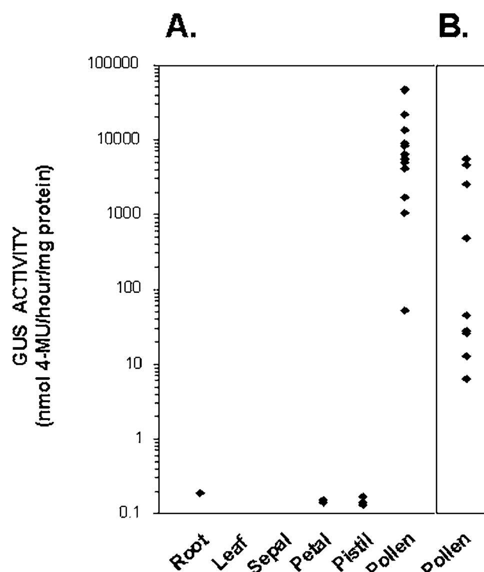

Figure 9. Expression of tPex promoter-GUS fusions in stable transformants. GUS activity in reproductive and vegetative tissues of 5 transgenic VF36 tomato plants harboring the short tPex-promoter fused to the GUS reporter gene (A) and in pollen from pollen of 13 transgenic plants and other transgenic VF36 plants harboring the long tPex promoter fused to a GUS reporter gene (B). Background GUS activity in extracts from untransformed plants was subtracted from GUS activity in transgenic plants. GUS activity in all tissue extracts was expressed as nmol of 4-methylumbelliferone (4-MU) generated per hour per milligram of protein. GUS activity below 0.1 nmol is not shown on the graph.

to that produced by a well-characterized strong tomato pollen promoter (positive control LAT52-GUS). These results demonstrate that the tPex gene is expressed in pollen. Comparison of the 5′ region of tPex genomic sequences with well-characterized tomato pollen promoters revealed the presence of a characteristic pollen promoter element, the LAT56/59 box (GAAT/A/CTTGTG) (Twell *et al.*, 1991; Eyal *et al.*, 1995) 193 bp 5′ to the TATA box.

In order to test the tissue specificity of the tPex promoter, tPex promoter-GUS fusions were stably transformed into tomato, as described in Materials and methods. Expression of GUS in various tissues of transformants is shown in Figure 9. Of the 21 transgenic plants assayed for GUS activity in pollen, those containing the short tPex promoter showed a considerably stronger expression of the GUS gene than did those from the long tPex promoter transgenic plants. This is in contrast to the transient expres-

sion results, where the long tPex promoter was more strongly expressed in pollen. Discrepancies between the expression of promoter-reporter gene fusions in transient as opposed to stable transformation have been documented (Hamilton *et al.*, 2000).

The tPex promoter is pollen-specific as shown by the high GUS activity in pollen extracts relative to other reproductive and vegetative tissues. It is conceivable that small but detectable GUS activity from petal and pistil extracts from some of the individual transgenic plants might be due to errant pollen contamination during the collection and dissection stages.

Discussion

In the study of genes with an unknown function, it is useful to compare homologous genes in similar and divergent species. Such comparisons can identify regions of genes that have been highly conserved, presumably as a result of selection for regions that are critical to preserve function. In this report, we describe the comparison of the mPex1 gene with a second maize gene and a dicot (tomato) gene. The conservation of an LRR, a cysteine-rich region, the extensin domain and the C-terminus, along with anther-specific gene expression, strongly suggests that we have isolated homologues, now called mPex1, mPex2 and tPex. In addition, we have identified potential Pex genes with highly conserved LRRs in *Arabidopsis*, potato and sorghum.

The LRRs in all known and potential Pex proteins are highly conserved, especially in the amino-terminal part of this motif. In the carboxy region, the amino acids thought to confer the α-helix/β-sheet repeat structure in LRRs are still highly conserved, but not the variable 'x' amino acids in the LRR motif. Specifically, 90% of the non-conservative amino acid differences between the mPex proteins and the tPex protein are in the positions of variable amino acids. The variable amino acids in LRRs are thought to confer specific protein-ligand contacts (Kobe and Deisenhofer, 1994, 1995). This model for LRR structure and function is supported by recent analysis of LRR proteins involved in pathogen resistance in plants (Meyer and Abel, 1975; Jones and Jones, 1996; Botella *et al.*, 1998; McDowell *et al.*, 1998). The high degree of divergence in the variable amino acids of the Pex LRRs implies that the Pex proteins in monocots and dicots recognize and bind to similar, but not identical, ligands. LRR-encoding regions have been identified in many plant genes (Muschietti *et al.*, 1998), and recent genetic and biochemical studies with the CLAVATA genes of *Arabidopsis* suggest that LRR-containing receptors (CLV1 and CLV2) interact with a small protein ligand (CLV3) (Fletcher *et al.*, 1999; Trotochaud *et al.*, 2000). The cysteine-rich region adjacent to the LRRs in monocot and dicot proteins is notable in the conservation of position of cysteines, particularly in the light of the importance of cysteines and disulfide bonds in the determination of protein structure.

The extensin domains in both the maize and tomato Pex proteins and potential *Arabidopsis* Pex proteins contain significant numbers of SP_{3-7} repeats. Studies of extensin and extensin-like proteins have shown that the prolines in these repeats are hydroxylated to hydroxyproline post-translationally (Lamport, 1963; Chrispeels, 1969), and are variably *O*-glyosylated (Kieliszewski and Lamport, 1994; Kieliszewski *et al.*, 1995). Runs of glycoslyated hydroxyprolines confer the extremely extended 'polyproline II' configuration on proteins, including extensins (van Holst and Varner, 1984). Thus, although alignment of amino acids in the extensin domains of the tomato, *Arabidopsis* and maize Pex proteins is more difficult than alignment of amino acids in the globular domains, the extensin domains in all of the Pex proteins are likely to be highly conserved at the structural level. It is notable that the carboxy-terminal 12 amino acids, which contain the only two tyrosines in the tPex and mPex2 extensin domains, are very highly conserved in mPex1, mPex2, tPex, aPex2 and aPex4. Tyrosines may be involved in the formation of chemical crosslinks of the proteins to the cell wall (Epstein and Lamport, 1984; Fry, 1986; Waffenschmidt *et al.*, 1993). It is possible that the tyrosines at the C-terminus of the Pex proteins play a role in the association of these proteins with the cell wall.

While the similarities between the monocot and dicot Pex proteins are instructive, the lack of similarity in a small region of the amino terminus is perhaps even more striking. This hypervariable region of 23–31 amino acids is sandwiched between two predicted regions of α-helix, and has a high predicted surface probability. Thus, it is possible that the Pex proteins have a variable segment protruding from the globular domain that may be involved in species-, family- or tissue-specific recognition.

While the exact function of the Pex proteins is still uncertain, future studies can now center on specific regions identified by the findings presented in

this report. Given the conserved motifs in the globular domain, the Pex proteins are likely to be interacting with ligands – possibly with molecules in the pistil, possibly to recognize the appropriate species (Heslop-Harrison *et al*., 1984) or tissues (Cheung *et al*., 1995). Another possibility is that the Pex proteins recognize and bind to each other in a self-assembly process during pollen tube growth (Hills *et al*., 1975). The hypervariable region and the LRRs will be the focus of searches for possible ligands for the Pex proteins.

Acknowledgments

The research in this report was supported by the National Science Foundation #IBN-9496338 to P.B., including REUs to support A.G. and J.G.S., and an ROA to support W.B., and a CAA NSF #IBN-9707384 to P.B. and USDA CRIS grant 5335-21000-011-00D to S.M. We thank Dr Hyun-Uk Kim for the RNA blot, Dr Amy Rubinstein for helpful discussions, and Amy Olson for help with figures.

References

Adair, W.S., Hwang, C. and Goodenough, U.W. 1983. Identification and visualization of the sexual agglutinin from the Mating-Type Plus flagellar membrane of *Chlamydomonas*. Cell 33: 183–193.

Baldwin, T.C., Coen, E.S. and Dickinson, H.G. 1992. The *ptl1* gene expressed in the transmitting tissue of *Antirrhinum* encodes an extensin-like protein. Plant J. 2: 733–739.

Barnabas, B. and Fridvasky, L. 1984. Adhesion and germination of differently treated maize pollen grains on the stigma. Acta Bot. Hung. 30: 329–332.

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

Botella, M., Parker, J., Frost, L., Bittner-Eddy, P., Beynon, J., Daniels, M., Holub, E. and Jones, J. 1998. Three genes of the *Arabidopsis RPP1* complex resistance locus recognize distinct *Peronospora parasitica* avirulence determinants. Plant Cell 10: 1847–1860.

Cassab, G.I. 1998. Plant cell wall proteins. Annu. Rev. Plant Physiol. Plant Mol. Biol. 49: 281–309.

Chen, C.-G., Cornish, E.D. and Clarke, A.E. 1992. Specific expression of an extensin-like gene in the style of *Nicotiana alata*. Plant Cell 4: 1053–1062.

Cheung, A.Y., Wang, H. and Wu, H.-M. 1995. A floral transmitting tissue-specific glycoprotein attracts pollen tubes and stimulates their growth. Cell 82: 383–393.

Chou, P. and Fasman, G. 1978. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. 47: 45–148.

Chrispeels, M. 1969. Synthesis and secretion of hydroxyproline containing macromolecules in carrots I. Kinetic analysis. Plant Physiol. 44: 1187–1193.

Cooper, J.B., Adair, W.S., Mecham, R.P., Heuser, J.E. and Goodenough, U.W. 1983. *Chlamydomonas* agglutinin is a hydroxyproline-rich glycoprotein. Proc. Natl. Acad. Sci. USA 80: 5898–5901.

Emini, E., Hughes, J., Perlow, D. and Boger, J. 1985. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. J. Virol. 55: 836–839.

Epstein, L. and Lamport, D. 1984. An intramolecular linkage involving isodityrosine in extensin. Phytochemistry 23: 1241–1246.

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

Ferguson, C., Teeri, T., Siika-aho, M., Read, S. and Bacic, A. 1998. Location of cellulose and callose in pollen tubes and grains of *Nicotiana tabacum*. Planta 206: 452–460.

Fletcher, J.C., Brand, U., Running, M.P., Simon, R. and Meyerowitz, E.M. 1999. Signaling of cell fate decisions by CLAVATA3 in *Arabidopsis* shoot meristems. Science 283: 1911–1194.

Fry, S.C. 1986. Cross-linking of matrix polymers in the growing cell walls of angiosperms. Annu. Rev. Plant Physiol. 37: 165–186.

Goldman, M.H., Pezzotti, M., Seurinck, J. and Mariani, C. 1992. Developmental expression of tobacco pistil-specific genes encoding novel extensin-like proteins. Plant Cell 4: 1041–1051.

Hamilton, D.A., Schwarz, Y.H., Rueda, J. and Mascarenhas, J.P. 2000. Comparison of transient and stable expression by a pollen-specific promoter: the transformation results do not always agree. Sex. Plant Reprod. 12: 292–295.

Heslop-Harrison, J. 1987. Pollen germination and pollen tube growth. Int. Rev. Cytol. 107: 1–78.

Heslop-Harrison, Y., Reger, B.J. and Heslop-Harrison, J. 1984. The pollen-stigma interaction in the grasses. 6. The stigma ('silk') of *Zea mays* L. as host to the pollens of *Sorghum bicolor* (L.) Moench and *Pennisetum americanum* (L.) Leeke. Acta Bot. Neerl. 33: 205–227.

Hills, G.J., Phillips, J.M., Gay, M.R. and Roberts, K. 1975. Self-assembly of a plant cell wall *in vitro*. J. Mol. Biol. 96: 431–441.

Jauh, G.Y. and Lord, E.M. 1996. Localization of pectins and AGPs in lily pollen tube and style and their possible roles in pollination. Planta 199: 251–261.

Jeanmougin, F., Thompson, J., Gouy, M., Higgins, D. and Gibson, T. 1998. Multiple sequence alignment with Clustal X. Trends Biochem. 23: 403–405.

Jefferson, R., Kavanagh, T. and Bevan, M. 1987. GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. 6: 3901–3907.

Jones, D. and Jones, J. 1996. The roles of leucine rich repeats in plant defences. Adv. Bot. Res. Plant Path. 24: 90–167.

Kieliszewski, M.J. and Lamport, D.T.A. 1994. Extensin: repetitive motifs, functional sites, post-translational codes, and phylogeny. Plant J. 5: 157–172.

Kieliszewski, M.J., O'Neill, M., Leykam, J. and Orlando, R. 1995. Tandem mass spectrometry and structural elucidation of glycopeptides from a hydroxyproline-rich plant cell wall glycoprotein indicate that contiguous hydroxyproline residues are the major sites of hydroxyproline *O*-arabinosylation. J. Biol. Chem. 270: 2541–2549.

Kneller, D., Cohen, F. and Langridge, R. 1990. Improvements in protein secondary structure prediction by an enhanced neural network. J. Mol. Biol. 214: 171–182.

Kobe, B. and Deisenhofer, J. 1994. The leucine-rich repeat: a versatile binding motif. Trends Biol. Sci. 19: 415–421.

Kobe, B. and Deisenhofer, J. 1995. Proteins with leucine-rich repeats. Curr. Opin. Struct. Biol. 5: 409–416.

Lamport, D.T.A. 1963. Oxygen fixation into hydroxyproline of plant cell wall protein. J. Biol. Chem. 238: 1438–1440.

Lamport, D.T.A. 1965. The protein component of primary cell walls. Adv. Bot. Res. 2: 151–218.

Lamport, D.T.A. 1977. Structure, biosynthesis, and significance of cell wall glycoproteins. In: F.A. Loewus and V.C. Runeckles (Eds.) Rec. Adv. Phytochem. 11: 79–115.

Li, Y.-Q., Bruun, L., Pierson, E.S. and Cresti, M. 1992. Periodic deposition of arabinogalactan epitopes in the cell wall of pollen tubes of *Nicotiana tabacum* L. Planta 188: 532–538.

Lind, J.L., Bacic, A., Clarke, A. and Anderson, M.A. 1994. A style-specific hydroxyproline-rich glycoprotein with properties of both extensins and arabinogalactan proteins. Plant J. 6: 491–502.

Logeman, J., Jeanmougin, F., Schell and Willmitzer, L. 1987. Improved method for the isolation of RNA from plant tissues. Anal. Biochem. 163: 16–20.

Mascarenhas, J.P. 1993. Molecular mechanisms of pollen tube growth and differentiation. Plant Cell 5: 1303–1314.

McCormick, S. 1991. Transformation of tomato with *Agrobacterium tumefaciens*. In: K. Lindsey (Ed.) Plant Tissue Culture Manual: Fundamentals and Applications, Kluwer Academic Publishers, Dordrecht, Netherlands.

McCormick, S., Smith, A., Gasser, C., Sachs, K., Hinchee, M., Horsch, R. and Fraley, R. 1987. Identification of genes specifically expressed in reproductive organs of tomato. In: D. Nevins and R. Jones (Eds.) Tomato Biotechnology, Alan R.Liss, New York, pp. 255–265.

McDowell, J., Dhandayham, M., long, T., Aarts, M., Goff, S., Holub, E. and Dangl, J. 1998. Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance in the RPP8 locus of *Arabidopsis*. Plant Cell 10: 1861–1874.

McLelland, J. and Rumelhart, D. 1988. Explorations in parallel distributed processing: a handbook of models, programs and exercises, MIT Press, Cambridge, MA.

Meyer, Y. and Abel, W.D. 1975. Importance of the cell wall for cell division and in the activity of the cytoplasm in cultured tobacco protoplasts. Planta 123: 33–40.

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

Rubinstein, A.L., Broadwater, A.H., Lowrey, K.B. and Bedinger, P.A. 1995a. Pex1, a pollen-specific gene with an extensin-like domain. Proc. Natl. Acad. Sci. USA 92: 3086–3090.

Rubinstein, A.L., Marquez, J., Suarez-Cervera, M. and Bedinger, P.A. 1995b. Extensin-like glycoproteins in the maize pollen tube wall. Plant Cell 7: 2211–2225.

Sambrook, J., Fritsch, E. and Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Plainview, NY.

Showalter, A., Bell, J., Cramer, C., Bailey, J., Varner, J. and Lamb, C. 1985. Accumulation of hydroxyproline-rich glycoprotein mRNAs in response to fungal elicitor and infection. Proc. Natl. Acad. Sci USA 82: 6551–6555.

Showalter, A.M. 1993. Structure and function of plant cell wall proteins. Plant Cell 5: 9–23.

Steer, M.W. and Steer, J.M. 1989. Tansley Review No. 16. Pollen tube tip growth. New Phytol. 111: 323–358.

Trotochaud, A.E., Jeong, S., and Clark, S.E. 2000. CLAVATA3, a multimeric ligand for the CLAVATA1 receptor-kinase. Science 289: 613–617.

Twell, D., Yamaguchi, J., Wing, R.A., Ushiba, J. and McCormick, S. 1991. Promoter analysis of genes that are coordinately expressed during pollen development reveals pollen-specific enhancer sequences and shared regulatory elements. Genes Dev. 5: 496–507.

van Holst, G. and Varner, J. 1984. Reinforced polyproline II conformation in a hydroxyproline-rich cell wall glycoprotein from carrot root. Plant Physiol. 74: 247–251.

von Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. Nucl. Acids Res. 14: 4683–3690.

Waffenschmidt, S., Woessner, J.P., Beer, K. and Goodenough, U.W. 1993. Isodityrosine cross-linking mediates insolubilization of cell walls in *Chlamydomonas*. Plant Cell 5: 809–820.

Wu, H.-M., Zou, J., May, B., Gu, Q. and Cheung, A.Y. 1993. A tobacco gene family of flower cell wall proteins with a proline-rich and a cysteine-rich domain. Proc. Natl. Acad. Sci. USA 90: 6829–6833.