

PO149, a new member of pollen pectate lyase-like gene family from alfalfa

Yongzhong Wu, Xiao Qiu¹, Sarah Du and Larry Erickson*

Department of Crop Science, University of Guelph, Guelph, Ont., Canada, N1G 2W1 (*author for correspondence); ¹Present address: Plant Biotechnology Institute, National Research Council of Canada, 110 Gymnasium Place, Saskatoon S7N 0W9, Canada

Received 30 January 1996; accepted in revised form 12 June 1996

Key words: *Medicago*, pollen-specific, promoter, microgametogenesis

Abstract

PO149 is a low-copy-number gene expressed in the late stages of pollen development. The promoter region contains no similarities in DNA sequence to those of other pollen-specific genes, except for a tobacco sequence (AAATGA), which occurs four times in this alfalfa gene and much further upstream than in tobacco. Four distinct TATA boxes were detected in the promoter with the distal and proximal TATA boxes being separated by a spacer of 269 nucleotides. Hairpin loop structures were found in the 5'- and 3'-untranslated regions of PO149 mRNA. The coding region of PO149 is interrupted by two introns and encodes a putative prepeptide of 450 amino acids with homology to pollen pectate lyase-like proteins and pollen allergens. The coding region also contains sequences characteristic of both a signal peptide and a nuclear localization signal.

Introduction

The plant cell wall is both an inert mechanical support and barrier, and a dynamic, metabolically active organelle, which plays an important role in cell-cell communication, cell growth, and differentiation. The primary cell wall is composed of a network of cellulose microfibrils embedded in a matrix of pectic polysaccharides. The orientation of microfibril synthesis defines the direction of cell growth and thus establishes cell shape, whereas the dynamic interaction of the cellulosic structures and the jelly-like pectic polysaccharides determines the rate of cell growth [32].

There are two types of growth in plants: cell expansion, in which the growth takes place over the entire cell surface, and a more localized tip growth. Most plant cells undergo cell expansion, whereas root hairs and pollen tubes provide examples of tip growth. The growth of the pollen tube is extremely rapid and is associated with the presence of numerous vesicles in the apical tip [34]. The principal component in the

vesicles is pectic in nature [8, 31] and during pollen tube growth vesicles fuse with the cell membrane in the growing tip to contribute cell wall precursor materials [13]. Since pectin is a primary component of the cell wall, rapid growth of the pollen tube conceivably involves rapid mobilization of pectin molecules from these vesicles to form the new wall. It is not surprising, therefore, that a number of pollen-expressed genes with similarities to pectinases have been isolated, including polygalacturonase [2, 5, 23, 26, 29], pectin esterase [1, 22, 25] and pectate lyase [30, 36, 40]. We report here the cloning and characterization of PO149, an alfalfa pollen-specific pectate lyase-like gene which has several unique features.

Materials and methods

Plant material

Plants of alfalfa (*Medicago sativa* ssp. *sativa*) were grown under greenhouse conditions as described [9]. Plant tissues were collected and frozen in liquid nitrogen or stored at -80 °C prior to RNA and DNA isolation.

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number U41472.

tion. Microspores were isolated from flower buds classified into 3 developmental stages according to size: buds less than 1.5 mm contain primarily microspores at meiosis; buds between 1.5 mm and 5.0 mm contain early pollen, and those longer than 5.0 mm contain mature pollen.

Nucleic acid isolation and analysis

RNA from microspore/pollen extracts was isolated as previously reported [24] or with a Trizol RNA isolation kit (Gibco-BRL). RNA from other tissues was isolated according to Ausubel *et al.* [3] and the purification of poly(A)⁺ RNA from total RNA followed Maniatis *et al.* [17]. For northern analysis 10 µg total RNA per sample was fractionated on 1% agarose gels containing 0.66 M formaldehyde and transferred to GeneScreen Plus filters. Filters were hybridized with labelled PO149 cDNA insert in 50% formamide, 5× SSPE, 5× Denhardt's solution, 1% SDS and 100 µg/ml salmon sperm DNA at 42 °C for 18 h, and washed with 2× SSC and 1% SDS at 65 °C for 30 minutes, followed by 15 minutes in 0.1× SSC and 0.1% SDS.

Plant genomic DNA was isolated according to Davies *et al.* [7] and plasmid and phage DNA, according to Ausubel *et al.* [3]. For Southern blot analysis, 15 µg DNA (per sample), digested by *EcoRI* or *HindIII*, was electrophoresed in 0.7% agarose gels and transferred to GeneScreen Plus filters. Filters were hybridized with labelled PO149 cDNA insert in 1.45 M NaCl, 10% dextran sulphate, 50% formamide and 1% SDS overnight. The washing conditions were as described above for the northern blot.

Library construction, screening and subcloning

cDNA was synthesized from 5 µg poly(A)⁺ RNA from a microspore/pollen extract using a kit from Stratagene; fragments longer than 500 bp were collected and cloned into λZap, and the recombinant phage DNA was packaged with Gigapack Gold extracts (Stratagene). The resulting cDNA library was differentially screened using cDNA probes made from seedling and microspore/pollen poly(A)⁺ RNA. An alfalfa genomic library was prepared from DNA partially digested with *MboI* and fractionated by sucrose gradient centrifugation. Fragments (16–20 kb) were inserted into the *BamHI* site of λDASH II arms, and packaged in vitro with Gigapack Gold extracts. The resulting library was screened with labelled cDNA inserts according to the manufacturer's instructions. Genomic sub-clones (in

pBluescript II) were sequenced by the dideoxy chain termination method [33] using a T7 Sequencing Kit (Pharmacia). Computer analysis of the nucleotide and deduced amino acid sequences was performed using PCGene (IntelliGenetics).

Primer extension analysis

A 20 bp synthetic oligonucleotide (GGTTTGTAAGAC AAGGATGG) complementary to nucleotides between 110 and 129 in the first exon was end-labelled and annealed to 20 µg total RNA. After extension of the primer by AMV reverse transcriptase, the products were analyzed on a sequencing gel using the sequencing reaction of the PO149 genomic clone as a reference.

Results

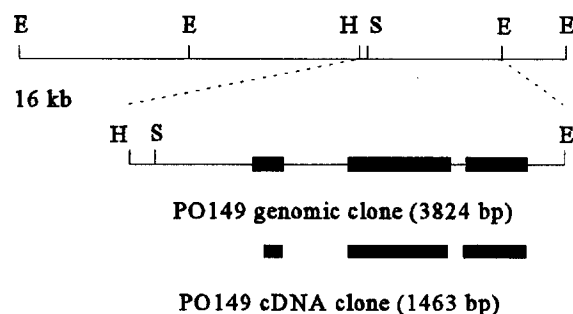
Isolation of PO149 cDNA clone and corresponding genomic clones

PO149 was one of several clones which hybridized only to the microspore/pollen cDNA following the differential screening of 5×10^5 independent clones. The size of this cDNA clone was 1463 bp with an 18 bp poly(A) tail. Sequence analysis of the reading frame, together with Northern blot data, suggested that PO149 was a partial cDNA clone.

In order to obtain the full coding region and the 5'-flanking sequence of PO149, ca. 1.5×10^5 original recombinant clones from a genomic library were screened with the PO149 cDNA insert, yielding two positive clones. Restriction mapping, Southern blot and PCR analysis of one clone indicated that a 3.8 kb *HindIII/EcoRI* fragment contained the entire coding region and ca. 1 kb of the 5'-flanking region (Fig. 1 A). This 3.8 kb fragment was subcloned into pBluescript II for sequencing.

Expression and copy number of PO149

The spatial and temporal expression patterns of PO149 were examined by gel blots of total RNA (Fig. 2). PO149 transcripts were detected only in mature pollen. A Southern blot of alfalfa genomic DNA, digested with *EcoRI* or *HindIII*, and probed with the cDNA insert, hybridized to two DNA fragments in each lane, suggesting that PO149 corresponds to a low-copy-number gene (Fig. 3).



PO149	TRRELSQKSRGGKGRR
Lat59: tomato [40]	TRRGLGTKKYT
G10: tobacco [30]	PRRQLSNKKYK
Lat56: tomato [40]	PRRKLTKKYR
Zm58.2 maize [36]	TRRALARY

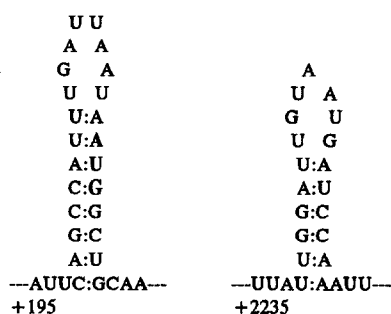


Figure 1. A (top). Physical map of genomic and cDNA clones of PO149. The exons are shown as black boxes. Restriction enzymes: *EcoRI* (E), *HindIII* (H), *SalI* (S). B(middle). Possible nuclear localization signals in some pollen pectate lyase-like proteins; conserved basic amino acid domains are in bold. C(bottom). Predicted hairpin loop structures in PO149 mRNA 5' region (left) and 3' region (right). The bold AUG is the translational start codon.

Structure of PO149

Both the genomic and cDNA clones of PO149 were completely sequenced. One nucleotide difference was detected between the genomic and cDNA clones: the T-to-C transition at position 400 is silent and probably results from a cloning artifact. A comparison of the genomic sequence with the cDNA sequence revealed the presence of two introns and the exon-intron junctions obey the GT-AG rule. The introns are AT-rich (72.1%) relative to the coding region (58.4%).

5'-end mapping of PO149 mRNA using primer extension produced 6 major and several minor exten-

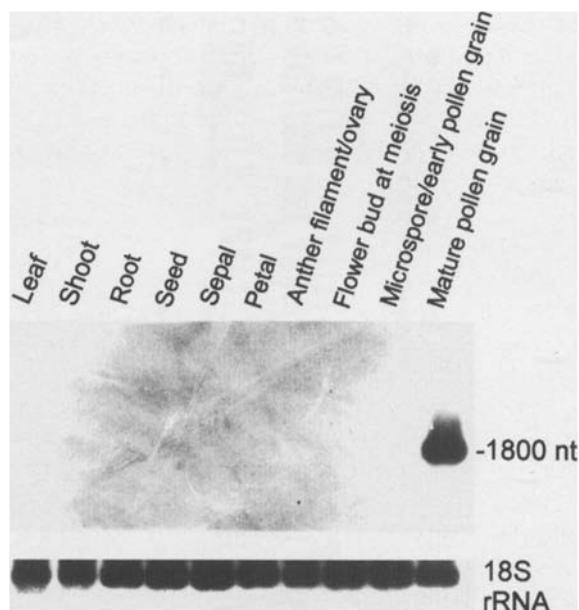


Figure 2. Northern blot analysis of PO149 gene expression. Blot was first probed with PO149 cDNA insert (upper portion of autoradiogram). After autoradiography, the blot was stripped and re-probed with an 18S rDNA clone from wheat (lower portion of autoradiogram).

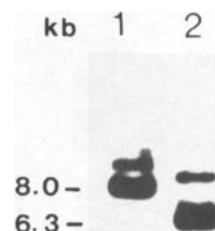


Figure 3. Southern blot of genomic DNA digested by *EcoRI* (1) and *HindIII* (2), probed with PO149 cDNA insert.

sion products (Fig. 4). The promoter region contains several repeating sequences: two 13 bp direct repeats at positions -453 and -403, and two 10 bp inverted repeats at positions -565 and -125. The polyadenylation site of PO149, deduced from the cDNA sequence, is located 2265 bp downstream of the proximal transcriptional start site. Two overlapping AATAAA-like sequences were found just after the stop codon, and another canonical AATAAA polyadenylation signal is located at position 2142, 118 bp upstream of the polyadenylation site.

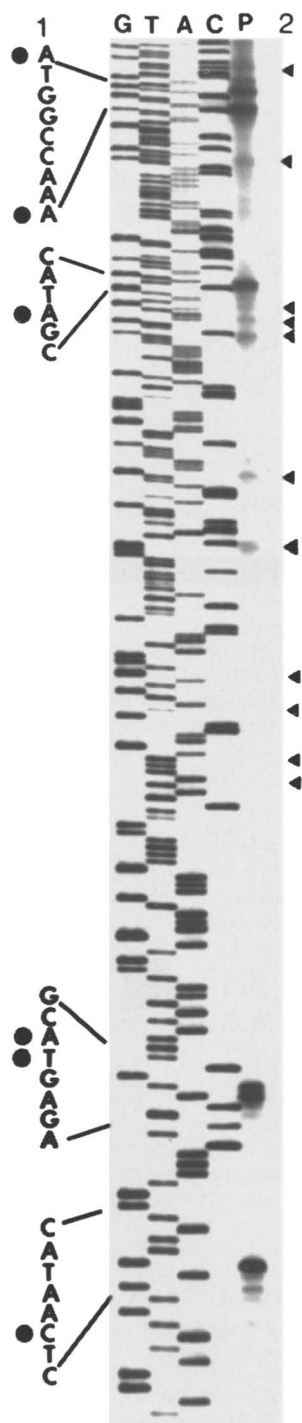


Figure 4. Mapping transcriptional start sites by primer extension. A 20 bp primer complementary to a sequence in the first exon was hybridized to 20 μ g total RNA from a microspore/pollen extract. The extension products (P) were separated on a sequencing gel along with the products of the sequencing reactions of the genomic clone (G, T, A, C) using the same primer. The DNA sequences presented vertically to the left of the autoradiograph (column 1) are of the sense strand with the black dots marking the positions of the major extension products. Other minor extension products (some not visible on the photograph) are marked by '▲' (column 2).

Coding region of PO149

Computer analysis of the reading frame [4] detected a signal peptide of 24 amino acids with the cleavage site between amino acids 24 (glycine) and 25 (arginine), thereby conforming to the $(-3, -1)$ rule [11]. A bipartite nuclear localization signal was also detected between amino acids 79 and 95, in which two pairs of arginines are separated by 12 amino acids. The signal peptide and the nuclear localization signal are separated by a spacer of 55 amino acids, and cleavage of the signal peptide would result in a protein composed of 426 amino acids with a molecular mass of 47 641 Da and a pI of 10.05. The resulting protein is relatively rich in glycine (9.11%) and arginine (8.2%), and has two potential N-linked glycosylation sites at amino acids 54 and 76 and one protein kinase phosphorylation site at amino acid 132.

A computer search of GenBank showed that PO149 has homology to several pollen proteins, a style protein and the pectate lyases from *Erwinia* species. The PO149 protein is 51–54% identical to reported pollen pectate lyase-like proteins [30, 36, 40]. PO149 might thus be included as a new member of the pectate lyase gene family. PO149 was also 50% identical to a pollen allergen [27] and 47% identical to a style protein [6].

Discussion

Pectate lyases belong to a large gene family which is divided into four subfamilies, one of which includes plant pectate lyase-like proteins [12]. Pectin is a primary component of the matrix of the plant cell wall and is also present in the wall surrounding pollen cells [1, 19]. All pectate lyase-like proteins expressed in pollen are considered to be secreted for extracellular function. However, in addition to a signal peptide sequence in PO149, there is possibly also a bipartite nuclear localization sequence consisting of 16 amino acids. Sequence alignment revealed conserved basic domains which are similar to the NLS for MAT α 2 (KIPK), and for SV40 (PKKKRKV) [28] in four members of eight previously-reported pollen pectate lyase-like proteins (Fig. 1B). Bipartite NLSs combine two regions of basic amino acids separated by a spacer of more than four residues. The occurrence of both a signal peptide sequence and a nuclear localization sequence in the same protein has been reported in other systems such as the mammalian gene for the FGF3 protein, which, as a result of competition between signals

for secretion and nuclear localization, is differentially localized within the cell depending on the tissue [14]. It is interesting to note that a recent analysis of the pectate lyase gene family has revealed sequence patterns with two possible enzymatic functions [12], implicating other functions for these proteins in addition to that of pectate lyase. In plant-microbe interactions, cell wall degrading enzymes are considered to be the basic compatibility factors which are required by pathogens for penetration of plant tissue [38]. Given the similarities between host-pathogen interactions and pollen-pistil interactions, pectate lyase-like proteins in pollen might also function as basic compatibility factors required by pollen for penetration into the style. It is conceivable that pectate lyase-like proteins in pollen may have as many as three functions: initial loosening of the pollen cell wall to enable pollen tube emergence and growth; breakdown of the cell wall of transmitting tissue in the style to facilitate penetration of pollen tube; interaction as ligands with receptor proteins in the stigma.

Although the numerous major and minor products observed following primer extension may be indicative of multiple transcriptional start sites in PO149, they may also have resulted from artifacts such as premature termination or post-transcriptional processing. However, four distinct TATA boxes, extending over 269 bp, correspond with the major extension products. It is thus possible that the PO149 promoter consists of multiple, overlapping and independently functional promoters. No common sequence was found upstream of the different TATA boxes and the promoter region does not contain the pollen boxes defined in tomato [10, 37]. A subsequence (AAATGA) identified as a *cis*-element required for pollen specificity in tobacco [39] occurs 4 times in this promoter; however, in contrast to tobacco, where this sequence is very close to the TATA box, all 4 copies in PO149 are 292 bp or more upstream of the distal TATA box and 468 bp or more upstream of the proximal TATA box. The significance of AAATGA in the promoter region of PO149 may therefore be questionable.

PO149 is a 'late' pollen-expressed gene. The transcripts of some of the 'late' genes are not translated, but are accumulated and stored for pollen germination and pollen tube growth [18,19]. The mechanism for repressing translation and maintaining mRNA stability of such pollen genes is unknown, but evidence from other studies implicates secondary structures in the 5' and 3' untranslated regions [15, 16, 21, 35]. Two hairpin structures which have a repressive effect on gene

expression [20] have been found in the 5' untranslated region of LAT59, pollen pectate lyase-like gene from tomato. A perfect hairpin loop structure (free energy -10.8 kcal) (Fig. 1C), detected in PO149 mRNA in which the translational start codon is sequestered as part of the stem structure, may thus play a similar role in repressing translation. Another perfect hairpin loop structure was detected just before the 3' end of PO149 mRNA (Fig. 1C). We have begun studies to determine the effect of such untranslated regions from pollen-specific genes from alfalfa on the expression of marker genes and genes of agronomic interest in this crop plant.

Acknowledgements

The authors are grateful to B. Livingstone and J. King-swell for assistance with the manuscript. Funds for this research were provided by the Natural Sciences and Engineering Research Council and the Ontario Ministry of Agriculture, Food and Rural Affairs.

References

1. Albani D, Altosaar I, Arnison PG, Fabijanski SF: A gene showing sequence similarity to pectin esterase is specifically expressed in developing pollen of *Brassica napus*. Sequences in its 5' flanking region are conserved in other pollen-specific promoters. *Plant Mol Biol* 16: 501-513 (1991).
2. Allen RJ, Lonsdale DM: Molecular characterization of one of the maize polygalacturonase gene family members which are expressed during late pollen development. *Plant J* 3: 261-271 (1993).
3. Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K: Current protocols in molecular biology. Wiley, NY (1987).
4. Bairoch A: The PROSITE dictionary of sites and patterns in proteins - its current status. *Nucl Acids Res* 21(13): 3097-3103 (1993).
5. Brown SM, Crouch ML: Characterization of a gene family abundantly expressed in *Oenothera organensis* pollen that shows sequence similarity to polygalacturonase. *Plant Cell* 2: 263-274 (1990).
6. Budelier KA, Smith AG, Gasser CS: Regulation of a stylar transmitting tissue-specific gene in wild-type and transgenic tomato and tobacco. *Mol Gen Genet* 224: 183-192 (1990).
7. Davies LG, Dibner MD, Battey JF: Basic methods in molecular biology. Elsevier, NY (1986).
8. Drashek WV, Rosen WG: Electron-microscopical localization of chemical components in the growth zone of *Lilium* pollen tubes. *Protoplasma* 61: 192-204 (1966).
9. Du S, Erickson L, Bowley S: Effect of plant genotype on the transformation of cultivated alfalfa (*Medicago sativa*) by *Agrobacterium tumefaciens*. *Plant Cell Rep* 13: 330-334 (1994).

10. Eyal Y, Curie C, McCormick S: Pollen specificity elements reside in 30 bp of the proximal promoters of two pollen-expressed genes. *Plant Cell* 7: 373–384 (1995).
11. Heijne G von: A new method for predicting signal sequence cleavage sites. *Nucl Acids Res* 14: 4683–4690 (1986).
12. Henrissat B, Heffron SE, Yoder MD, Lietzke SE, Jumak F: Functional implications of structure-based sequence alignment of proteins in the extracellular pectate lyase superfamily. *Plant Physiol* 107: 963–976 (1995).
13. Heslop-Harrison J, Heslop-Harrison Y: The growth of the grass pollen tube: 1. Characteristics of the polysaccharide particles ('P-particles') associated with apical growth. *Protoplasma* 112: 71–80 (1982).
14. Kiefer P, Acland P, Pappin D, Peters G, Dickson C: Competition between nuclear localization and secretory signals determines the subcellular fate of a single CUG-initiated form of FGF3. *EMBO J* 13: 4126–4136 (1994).
15. Kozak M: Regulation of translation in eukaryotic systems. *Ann Rev Cell Biol* 8: 179–225 (1992).
16. Manen JF, Simon P: A possible explanation for the multiple polyadenylation sites in transcripts coding for a winged-bean leghemoglobin. *Planta* 191: 289–292 (1993).
17. Maniatis T, Fritsch E, Sambrook J: *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982).
18. Mascarenhas JP: The male gametophyte of flowering plants. *Plant Cell* 1: 657–664 (1989).
19. Mascarenhas JP: Molecular mechanisms of pollen tube growth and differentiation. *Plant Cell* 5: 1303–1314 (1993).
20. McCormick S, Catherine C, Eyal Y, Muschietti J, Dircks L, Kulikaukas R: Molecular biology of male gametogenesis. *Euphytica* 79: 245–250 (1994).
21. Merrick WC: Mechanisms and regulation of eukaryotic protein synthesis. *Microbiol Rev* 56: 291–315 (1992).
22. Mu J, Stains JP, Kao T: Characterization of a pollen-expressed gene encoding a putative pectin esterase of *Petunia inflata*. *Plant Mol Biol* 25: 539–544 (1994).
23. Niogret MF, Dubald M, Mandaron P, Mache R: Characterization of pollen polygalacturonase encoded by several cDNA clones in maize. *Plant Mol Biol* 17: 1155–1164 (1991).
24. Qiu X, Erickson L: A simple and effective method for isolating RNA from alfalfa pollen. *Plant Mol Biol Rep* 12: 209–214 (1994).
25. Qiu X, Erickson L: A pollen-specific cDNA (P65, GenBank accession number U28148) encoding a putative pectin esterase in alfalfa. *Plant Gene Register, Plant Physiol* 109(3): 1127 (1995).
26. Qiu X, Erickson L: A pollen-specific polygalacturonase-like cDNA from alfalfa. *Sex Plant Reprod* 9: 123–124 (1996).
27. Rafnar T, Griffith U, Kuo M, Bond JF, Rogers BL, Klapper D: Cloning of *Amb al* (Antigen E), the major allergen family of short ragweed pollen. *J Biol Chem* 15: 1229–1236 (1991).
28. Raikhel N: Nuclear targeting in plants. *Plant Physiol* 100: 1627–1632 (1992).
29. Robert LS, Allard S, Gerster JL, Cass L, Simmonds J: Isolation and characterization of a polygalacturonase gene highly expressed in *Brassica napus* pollen. *Plant Mol Biol* 23: 1273–1278 (1993).
30. Rogers HJ, Harvey A, Lonsdale DM: Isolation and characterization of a tobacco gene with homology to pectate lyase which is specifically expressed during microsporogenesis. *Plant Mol Biol* 20: 493–502 (1992).
31. Rosen WG, Gawlick SR: Fine structure of lily pollen tubes following various fixation and staining procedures. *Protoplasma* 61: 181–191 (1966).
32. Salisbury FB, Ross CW: *Plant physiology*, pp. 334–336. Wadsworth I Press, Belmont, CA (1992).
33. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74: 5463–5467 (1977).
34. Sassen MA: Fine structure of *Petunia* pollen grain and pollen tube. *Acta Bot Neerl* 13: 175–181 (1964).
35. Stem DB, Radwansld ER, Kindle KL: A 3' stem/loop structure of the *Chlamydomonas* chloroplast *atpB* gene regulates mRNA accumulation in vivo. *Plant Cell* 3: 285–297 (1991).
36. Turcich MP, Hamilton DA, Mascarenhas JP: Isolation and characterization of pollen-specific maize genes with sequence homology to ragweed allergens and pectate lyases. *Plant Mol Biol* 23: 1061–1065 (1993).
37. Twell D, Yamaguchi J, Wing RA, Ushiba J, McCormick S: Promoter analysis of genes that are coordinately expressed during pollen development reveals pollen-specific enhancer sequences and shared regulatory elements. *Genes Devel* 5: 496–507 (1991).
38. Walton JD: Deconstructing the cell wall. *Plant Physiol* 104: 1113–1118 (1994).
39. Weterings K, Schrauwen J, Wullems G, Twell D: Functional dissection of the promoter of the pollen-specific gene NTP303 reveals a novel pollen-specific, and conserved *cis*-regulatory element. *Plant J* 8: 5–63 (1995).
40. Wing RA, Yamaguchi J, Larabell SK, Ursin VM, McCormick S: Molecular and genetic characterization of two pollen-expressed genes that have sequence similarity to pectate lyases of the plant pathogen *Erwinia*. *Plant Mol Biol* 14: 17–28 (1989).