

Short communication

Isolation and characterization of pollen-specific maize genes with sequence homology to ragweed allergens and pectate lyases

Michael P. Turcich, Douglas A. Hamilton and Joseph P. Mascarenhas*

Department of Biological Sciences and Center for Molecular Genetics, University at Albany, State University of New York, Albany, NY 12222, USA (*author for correspondence)

Received 28 June 1993; accepted in revised form 16 August 1993

Key words: allergens, gene expression, microsporogenesis, pectate lyase, pollen, *Zea mays*

Abstract

A cDNA clone (Zm58.1) was isolated by differential screening from a cDNA library made to mature *Zea mays* pollen, and shown to be pollen-specific by RNA blot analysis. When this partial-length clone was used to probe a genomic library, a similar but distinct pollen-specific genomic clone (68% sequence identity) was isolated (Zm58.2). The putative proteins coded for by these two clones show sequence homology to several flower-expressed gene products from various plant species, including known pollen allergens from short ragweed (*Ambrosia artemisiifolia*), and to pectate lyases from the plant pathogenic bacteria *Erwinia* spp. The two genes map to different chromosomes.

The function of a mature pollen grain is to effect the delivery of its two sperm cells to the embryo sac. To do this, the pollen grain must germinate on the surface of a receptive stigma, quickly sending a pollen tube out through the pore in the pollen wall and down through the tissues of the style. The pollen tube extends rapidly at its tip as it grows through or between the walls of the stigma cells, and then between the cells of the transmitting tissue in the style [17]. This germination and rapid growth requires a considerable amount of cell wall synthesis and possibly breakdown of stylar components [12]. One might expect these

processes to be performed by the enzymes present in the mature pollen grain and developing tube [9]. It is not surprising that many of the pollen-specific genes reported to date appear to have homology to cell wall degradation enzymes such as polygalacturonase [2, 3, 11], pectin esterase [1], and pectate lyase [15, 21]. Based on the intrusive nature of the growth exhibited during germination and pollen tube elongation, one might expect that genes of this type would be highly expressed in pollen.

We report here the characterization of a pair of pollen-specific genes of this type isolated from the

The nucleotide sequence data reported here will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers L20139 (Zm58.1) and L20140 (Zm58.2).

mature pollen of *Zea mays*. These genes, designated Zm58.1 and Zm58.2, show significant homology to pectate lyases [EC 4.2.2.2] known from the plant pathogenic bacteria *Erwinia* spp., but are more homologous yet to several other genes isolated from the pollen of various plant species, including the major pollen allergen of short ragweed (*Ambrosia artemisiifolia*).

Differential screening of a maize cDNA library made to mature pollen resulted in the isolation of numerous pollen-expressed clones [18]. Several of these were shown to be pollen-specific based on their hybridization only to cDNA probe made to poly(A)⁺ RNA isolated from pollen. RNA blot analysis showed one clone (designated Zm58.1) which hybridized specifically to RNA from mature maize pollen, but not to the RNAs from the vegetative tissues assayed (data not shown). Subsequent sequencing and characterization showed that this was a partial cDNA clone.

To obtain the full coding sequence and additional flanking sequences of Zm58.1, a maize genomic library made in λ EMBL3 [7] was screened to recover the corresponding genomic clone. A plaque was isolated yielding a genomic fragment which was designated Zm58.2. RNA blot analysis using a gene-specific part of this fragment as a probe indicated that Zm58.2 was also pollen-specific (data not shown).

The entire coding region of the gene was present on the genomic clone. While the start of the Zm58.2 mRNA has not been unequivocally determined, the sequence context of the region at the putative 5' end of the message is typical for known plant genes. Downstream from a TATATATA motif there are at least three places which could be putative transcription start sites [8]. The first ATG after this region starts an open reading frame of 438 amino acids, which would

code for the putative Zm58.2 protein. Modified polyadenylation signal motifs (e.g. ATTAAA and AATATA) can be found in the 3' untranslated region (portions of both the 5' and 3' untranslated regions are included in sequence L20140). Taken together these structures predict a mRNA very close to the 1.8 kb size which was observed by RNA blot analysis (data not shown). The N-terminal region derived from the Zm58.2 sequence is indicative of a signal peptide. All of the homologous genes shown in Fig. 1 are reported to have putative signal peptide sequences and are presumably secreted for extracellular function.

Complete cloning and sequencing of hybridizing subfragments of the Zm58.2 genomic clone indicated that, while there was a region of considerable sequence homology to much of the putative Zm58.1-coding region, it represented a different gene. At the nucleotide sequence level the homology in the coding region was high (68% identity), but much lower in the 3'-untranslated region (41% identity). The amino acid identity of the shared coding region was 53% (69% similarity). The cDNA and genomic clones thus represent different members of a maize multi-gene family of a size that is not yet known. The recombinant inbred maize lines developed by Burr *et al.* [5] were used to map the chromosomal locations of both of the Zm58 genes. Zm58.1 maps to a position on the long arm of chromosome 4 between *umc52* and *npi203*, and Zm58.2 maps to a locus one map unit from *wxl* near *bnl3.06* on the short arm of chromosome 9. The loci of Zm58.1 and Zm58.2 have been designated *mgs2A* and *mgs2B* (male gametophyte-specific), respectively.

Subsequent comparison of the putative Zm58.2 protein to protein sequences residing in the SwissProt database revealed striking homology

Fig. 1. Amino acid sequence comparisons among Zm58.1, Zm58.2, tomato anther-expressed genes *lat56* and *lat59* [21], tobacco pollen-specific gene *G10* [15], style-expressed gene *9612* from tomato and tobacco [4], short ragweed allergens *Amb a* I.1 and *Amb a* II [13, 14], and pectate lyase F from *Erwinia chrysanthemi* (Ep1F) [20]. Closed circles represent absolute amino acid conservation among all of the sequences at the given position; open circles represent conservation among all of the plant genes; pluses represent conservation of a residue in all but one or two of the sequences. Horizontal lines indicate regions of conservation within known *Erwinia* spp. pectate lyase genes [21, 22]. Computer analyses were performed using the Genetics Computer Group (GCG) Sequence Analysis Software, Version 7.1 for UNIX [6].

to several published florally expressed genes: *lat56* and *lat59* are anther-expressed genes from tomato [10, 21]; *G10* is a pollen-expressed gene from tobacco [15], and *9612* is a gene expressed in the styles of tobacco and tomato [4] (Fig. 1). These genes have previously been reported to be homologous to genes encoding pectate lyases from the plant pathogenic bacteria *Erwinia carotovora* and *E. chrysanthemi*. Several of the other homologous genes revealed by the sequence search are members of the major family of allergenic proteins known from short ragweed (*Ambrosia artemisiifolia*). The homologies between Zm58.2 and the plant genes ranged from 75% similarity (55% identity) for *lat59* to 63% similarity (46% identity) for *Amb a* 1.1. The *E. chrysanthemi* pectate lyase F gene [20] was identified by a FASTA comparison as having the closest homology to Zm58.2 of several *Erwinia* spp. pectate lyases (47% similarity, 26% identity). Pectate lyase F is the designation for the unprocessed pectate lyase E protein [20].

The multiple amino acid sequence alignment of these proteins shown in Fig. 1 is an updated version of a similar comparison originally made by McCormick [10]. In general, the regions of greatest conservation were found to be similar to regions of conservation shown for several pectate lyase genes known from *Erwinia* (Fig. 1, horizontal lines [21]). While no definitive biochemical functions have been shown for any of the homologous plant genes, their similarity to pectate lyases indicates that they may also have pectate lyase activity. The homology among the proteins in Fig. 1 is very limited at their N-terminal ends, but is very pronounced in the central portion, especially in the pectate lyase-conserved regions previously discussed. As can be seen in Fig. 1, the Lat56 protein does not terminate with the highly conserved region common to the other plant genes. It is interesting to note that the introduction of a frameshift in the region of the putative Lat56 protein terminus restores the amino acid sequence homology and terminates the protein just following the last conserved cysteine; the same point as all of the other plant proteins. All of the plant proteins compared in Fig. 1 also have

homology to *Amb a* 1 in the order of 45–48% amino acid identity, and may share similar epitopes. An additional homology has also been noted between these genes and the partial protein sequence of the major allergen of Japanese cedar pollen (*Cry j*1; [10, 19]).

The three-dimensional structure of pectate lyase C from *E. chrysanthemi* as determined by X-ray crystallography has recently been published indicating that the protein folds with a newly identified configuration designated a parallel β -helix [22]. The protein sequence similarity between the *Erwinia* pectate lyases and the previously published pollen and style genes prompted Yoder *et al.* [22] to suggest that these plant proteins also exhibit this novel protein folding. By extension, this is also likely true for the Zm58 protein(s). While the plant genes do not maintain the same conserved cysteines that are apparently responsible for some of the final structure of the pectate lyases [22], there are at least five cysteines which are conserved among the plant genes and which suggest a consistent structure. The areas of high amino acid conservation indicated by the lines in Fig. 1 correlate with the reported substrate-binding region, including specific residues speculatively associated with the binding of a Ca^{2+} ion. The other motifs mentioned in regards to the pectate lyase C (i.e. putative nucleotide-binding site, homology to tubulins [22]) are not conserved in the plant genes.

Taken together, the high level of homology among all of the proteins in Fig. 1 leads to the suggestion that these proteins may all have pectate lyase activity, as well as representing an allergenic class of proteins. Enzymes such as β -1,4-glucanase and pectinase have been reported to stimulate pollen tube growth rate when added to pectin-free growth media in low concentrations [16]. Thus, the pectate lyase might have functions in pollen tube extension. It is interesting to note that the 9612 protein (which is not expressed in pollen) has been localized in the stylar tissue of tomato and tobacco pistils [4]. It would be interesting to transform plants with a pollen-expressed anti-sense construct of one of these genes to see if antigen levels could be reduced in

pollen, and to note the effect on pollen tube germination and growth.

Acknowledgements

We acknowledge the participation of Thomas Teichmann in isolating the genomic Zm58 clone as part of a project while in the lab. We would like to thank Drs Benjamin and Frances Burr for mapping Zm58.1 and Zm58.2, and Dr Donna Crone for a critical review of this manuscript.

References

- 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).
- 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).
- 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).
- 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).
- Burr B, Burr FA, Thompson KH, Albertson MC, Stuber CW: Gene mapping with recombinant inbreds in maize. *Genetics* 118: 519–526 (1988).
- Devereux J, Haeberli P, Smithies O: A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* 12: 387–395 (1984).
- Hamilton DA, Bashe DM, Stinson JR, Mascarenhas JP: Characterization of a pollen-specific genomic clone from maize. *Sex Plant Reprod* 2: 208–212 (1989).
- Joshi CP: An inspection of the domain between putative TATA box and translation start site in 79 plant genes. *Nucl Acids Res* 15: 6643–6653 (1987).
- Mascarenhas JP: Gene activity during pollen development. *Annu Rev Plant Physiol Plant Mol Biol* 41: 317–338 (1990).
- McCormick S: Molecular analysis of male gametogenesis in plants. *Trends Genet* 7: 298–303 (1991).
- Niogret M-F, Dubald M, Mandaron P, Mache R: Characterization of pollen polygalacturonase encoded by several cDNA clones in maize. *Plant Mol Biol* 17: 1155–1164 (1991).
- Pressey R, Reger BJ: Polygalacturonase in pollen from corn and other grasses. *Plant Sci* 59: 57–62 (1989).
- Rafnar T, Griffith IJ, Kuo M, Bond JF, Rogers BL, Klapper DG: Cloning of *Amb a I* (antigen E), the major allergen family of short ragweed pollen. *J Biol Chem* 266: 1229–1236 (1991).
- Rogers BL, Morgenstern JP, Griffith IJ, Yu X-B, Counsell CM, Brauer AW, King TP, Garman RD, Kuo M-c: Complete sequence of the allergen *Amb a II*. *J Immunol* 147: 2547–2552 (1991).
- 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).
- Roggen HPJR, Stanley RG: Cell wall hydrolysing enzymes in wall formation as measured by pollen tube extension. *Planta* 84: 295–303 (1969).
- Sanders LC, Lord EM: A dynamic role for the stylar matrix in pollen tube extension. In: Russel SD and Dumas C (eds) *Sexual Reproduction in Flowering Plants*. Academic Press, San Diego (1992).
- Stinson JR, Eisenberg AJ, Willing RP, Pe ME, Hanson DD, Mascarenhas J: Genes expressed in the male gametophyte of flowering plants and their isolation. *Plant Physiol* 83: 442–447 (1987).
- Taniai M, Ando S, Usui M, Kurimoto M, Sakaguchi M, Inouye S, Matuhasi T: N-terminal amino acid sequence of a major pollen allergen of Japanese cedar pollen (*Cry jI*). *FEBS Lett* 239: 329–332 (1988).
- Van Gijsegem F: Relationship between the pel genes of the pelADE cluster in *Erwinia chrysanthemi* strain B374. *Mol Microbiol* 3: 1415–1424 (1989).
- 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).
- Yoder MD, Keen NT, Jurnak F: New domain motif: The structure of pectate lyase C, a secreted plant virulence factor. *Science* 260: 1503–1507 (1993).