213

Isolation and characterization of a microspore-specific gene from tobacco

M.T. Oldenhof, P.F.M. de Groot, J.H. Visser, J.A.M. Schrauwen* and G.J. Wullems
Molecular Plant Physiology Research Group, Department of Experimental Botany, University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, Netherlands (author for correspondence)*

Abstract

The characterization of a gene with a unique microspore-specific expression pattern is reported. Isolated microspores from tobacco were used to synthesize a cDNA library. Clones that did not hybridize to leaf cDNA were further characterized by northern analysis. One clone proved to be a microspore-specific cDNA, representing a transcript of 650 nt. The corresponding gene, NTM19 (*Nicotiana tabacum* microspore-specific), was isolated and its sequence analysed. The gene encodes a protein of 10.8 kDa with a pI of 6.92 and a putative signal sequence at the N-terminus. A localization study revealed a unique spatial and temporal distribution. The transcript was only detected in the unicellular microspore. No hybridization signals were observed in other pollen developmental stages, nor in the surrounding anther tissues or other vegetative tissues of the plant. Therefore it can be concluded that NTM19 is a gene with a highly microspore-specific character according to both localization and stage of expression. Southern blot analysis demonstrated the presence of a small gene family. The occurrence of TNM19 was investigated in a range of closely and distantly related species and was found to be present in other solanaceous species, including the ancestors of tobacco and in a monocot species.

Introduction

In angiosperm plants, sexual reproduction requires the production of viable male and female gametophytes. Pollen, as the male gametophyte, is formed within the anther and is initiated from sporogenous cells, which develop into meiocytes. The meiocyte undergoes meiosis to form a tetrad of haploid microspores which are subsequently released into the anther locule. Upon expansion and vacuolation an asymmetrical mitosis of the microspore results in bicellular pollen, containing a vegetative and generative cell. In the majority of plant species, including tobacco, pollen is shed in a bicellular condition. In other plant species (cereals and *Brassica* sp.), pollen maturation includes the mitotic division of the generative cell such that pollen is shed in a tricellular condition.

The morphological aspects of pollen development have been studied in great detail but knowledge about the underlying molecular processes is relatively limited. Pollen formation requires coordinated gene expression in the gametophytic cells and the sporophytic tissues surrounding it. Two periods of temporal gene expression are defined in pollen development by Mascarenhas [20]. The early developmental stage starts after meiosis and ends with the microspore mitosis, the late stage covers the period after mitosis up to mature pollen [14]. During pollen development the mRNA content present in different stages changes qualitatively [27]. Genes which are transiently and specifically transcribed during pollen development are assumed to play a crucial role in the developmental processes. Summarizing, genes involved in pollen development can be classified, according to expression pattern as early [10, 28, 35], and late genes [4, 7, 21, 32, 39, 42]. Additionally, genes transcribed in both stages have been isolated [1, 24, 33].

This classification does not use the place of expression in consideration. If this localization is taken as an extra criterion, the previous classification can be refined, which leads to four categories of pollen developmental genes: first, a category of genes with exclusive expression in the unicellular microspore; second, a group with localized expression in the diploid anther tissues during development; third, a group with localized expression in pollen; and fourth, a group with expression in unicellar microspores, pollen and diploid anther tissues. Genes of the first category have not been identified yet.

The aim of our study was to obtain more information about the involvement of the haploid genome in early pollen development. From microspore culture experiments [36] it is known that microspores can develop autonomously into mature pollen *in vitro* and it can be concluded that pollen development can occur under the control of the haploid genome. We, therefore, were interested in the identification of early genes with localized expression exclusively in the unicellular microspore.

Previous studies of early genes restricted to mi-

crosporogenesis resulted in a group of genes of which expression [10, 28, 35] is localized in the tapetum or other anther tissues. A few genes have been characterized which are expressed in both developing pollen and in tapetum [24, 29, 33]. This is a logical consequence of an approach in which cDNA libraries were synthesized from mRNA isolated from extracts of total anther tissues. Due to the high contribution of tapetal cells to the total anther tissues, differential screening from such libraries results mainly in the isolation of genes with expression in the tapetum.

In our research we focused on genes characterized by transient expression during microsporogenesis with a microspore-specific localization. To minimise contamination of the mRNA with sporophytic material, a laborious isolation of microspores was undertaken and microspore RNA was used for cDNA library construction. This new approach with homogenous material according to cell type and developmental stage has proved to be successful. A microspore-specific gene with unique expression characteristics was isolated. The DNA sequence of the gene and its temporal and spatial expression pattern during pollen development have been determined, which is described and discussed here.

Materials and methods

Plant material

Plants of *Nicotiana tabacum* L. cv. Petit Havana SR1 [19], were grown in soil under greenhouse conditions. Flower buds of selected length were collected to obtain pollen of different developmental stages. The determination of developmental stage and isolation of the anther content was carried out as previously described [27].

Plants of *Nicotiana tomentosiformis*, *Nicotiana sylvestris*, *Nicotiana otophora*, *Petunia hybrida*, *Brassica oleracea* and *Tradescantia virginiana* were grown in the greenhouse, while *Verbascum thapsus* was grown in the field.

Nucleic acid isolation

RNA was isolated with guanidine thiocyanate as RNase inhibitor by the procedure of Cathala [5] with minor modifications. All plant tissues including the microspores were frozen in liquid nitrogen and ground in a prechilled mortar. The ground material was placed in 2 ml lysis buffer and, after centrifugation at $3000 \times g$ for 10 min at 15 °C, the supernatant was collected. The duration of all centrifugation steps in the procedure was modified from the original procedure to 20 min.

The mRNA to be used for the synthesis of the cDNA library was bound on the surface of oligo(dT) Dynabeads (Dynal) as described by the manufacturer [11]. The mRNA used to synthesize a single-stranded cDNA probe was isolated as reported [25]. Plasmid DNA was isolated according to standard procedures [25], genomic DNA as reported [6] and phage DNA was obtained with LambdaSorb Phage Adsorbent according to the manufacturer's protocol (Promega).

cDNA library construction and screening

A cDNA library in λZAP II [30] was made from mRNA isolated from microspores. The cloning kit and the Gold Packaging Extract of Stratagene were used according to the manufacturer's protocols. The cDNA library was amplified once prior to screening. Differential screening was carried out with ^{32}P-labelled single-stranded cDNA probes derived from either mRNA from microspores or young leaves. The screening procedure of Stratagene was followed at a hybridization and washing temperature of 55 °C. The filters were washed in steps of 30 min in SSC with 0.1% (w/v) SDS (1 × SSC is 150 mM NaCl, 15 mM sodium citrate pH 7). The SSC concentration was lowered in steps: 6 × SSC, twice in 2 × SSC followed by 0.5 × SSC and 0.2 × SSC. A primary library screening of 2×10^4 cDNA clones was done.

The single-stranded cDNA probes were made

as described [25] with the following changes: $2 \mu g$ mRNA, oligo(dT), 75 μCi [α-^{32}P]dATP (3000 Ci/mmol) and 400 units Superscript reverse transcriptase from BRL were used to prime and radiolabel the probe. Selected plaques were isolated and excised *in vivo* from the λZAP II phagemid to form a pBluescript SK– plasmid according to the procedure of Stratagene.

Genomic library screening

The genomic library of *Nicotiana tabacum* cv. Samsun in bacteriophage λ Charon 32 [12] was a generous gift of Dr R.B. Goldberg [15], *Escherichia coli* K802 cells were used as hosts. A cDNA of 240 bp comprising part of the 3'-end region was used to synthesize a random-primed probe labelled with [α-^{32}P]dATP as described previously [9]. The screening and hybridization were carried out using standard techniques [25]. Hybridization and washings were done at 65 °C, with stringent washing up to 0.1 × SSC, 0.1% SDS. A positive clone of 11 kb was digested with *Kpn*I and *Xho*I. A digested fragment of 3 kb containing the sequence of interest was subcloned into the plasmid pGEM 7Zf+ (Promega) using standard methods [25].

Northern and Southern analysis

Samples of 25 μg of total RNA were denatured, electrophoresed on a 1% (w/v) agarose formaldehyde gel, and capillary-blotted onto Hybond-N membrane (Amersham) as reported [25]. The rRNA bands were used as size markers. The RNA was bound to the membrane by baking for 2 h at 80 °C. The same probe was used as for the screening of the genomic library. Hybridization was carried out overnight at 65 °C in 6 × SSC, 5 × Denhardt's reagent, 0.5% SDS and 100 μg/ml denatured herring sperm DNA. Washing was done at the same temperature during 30 min in 6 × SSC, 0.1% SDS followed by 2 × SSC, 0.1% SDS. An actin probe was hybridized to the stripped northern blot as a control.

Southern analysis of 10 μg genomic DNA was done after restriction digestion and electrophoresis on a 0.8% agarose gel using standard methods [25] except the depurination, which was carried out by UV irradiation. Southern blots were hybridized overnight at 60 °C, and washed to a stringency of 0.4 × SSC at 60 °C. The probe used in Southern analysis was made of a fragment of 250 bp including partial 5′ non-coding and 5′ coding regions. Filters were exposed to Kodak X-Omat AR film with intensifying screens at −80 °C.

Sequence analysis

DNA sequencing was performed by the dideoxy chain termination method [26] with the use of T7 polymerase (Pharmacia). The genomic clone was subcloned in pGEM 7Zf+ (Promega). Sequencing was performed with part single-stranded and double-stranded template DNA [43]. Sequence data were analysed with the PC/Gene program from Intelligenetics, Geneva (Switzerland).

In situ *hybridization*

Anthers were collected for *in situ* hybridization from tobacco flower buds of varying lengths (6, 8, 12, 20, 30 and 48 mm) covering a broad range of pollen developmental stages, from pollen mother cells to mature pollen. By use of a light microscope, the selected pollen developmental stages are distinguished clearly by their morphological characteristics visible after a DAPI staining [27].

The procedure as described by Reijnen *et al.* [23] was used with the following modifications. Prior to fixation the top of the anther was cut to optimise penetration of the fixative. Fixation was done for 12 h and initiad with degassing for 15 min. Washing after hybridization was started with 2 × SSC for 15 min at room temperature and coverslips removed. Subsequently, an RNase A treatment (20 μg/ml in 0.5 M NaCl, 10 mM Tris-HCl pH 7.5, 1 mM Na_2EDTA) was carried out for 45 min at 37 °C. The slides were washed twice in 2 × SSC, 1 mM DTT (dithiotreitol) for 30 min at 37 °C, followed by 0.2 × SSC, 1 mM DTT for 1 h at 45 °C.

The NTM19 RNA probes were synthesized by *in vitro* transcription in sense and antisense orientation from a cDNA subclone in pBluescript, that contained a 200 nt part of the coding region. The sense probe was used as a negative control. The protocol of Promega was followed with the use of $[5,6^{-3}H]UTP$ and for each slide 2×10^5 c.p.m. were used in hybridization. After processing of the slides, anthers were stained in 1 μg/ml ethidium bromide for 5 min and washed in distilled water for 20 s. A confocal laser scanning microscope (BioRad MRC-500) was used to analyse the sections. Pictures were made combining those made from a fluorescent image resulting from ethidium bromide staining, with a reflection image that visualised silver grains which are formed if the probe used hybridizes to the section.

Determination of the transcription initiation point

To determine the transcription start site, the primer extension technique was used as previously described [34]. A 30-mer primer was chosen 30 nt downstream of the translation start point, with the help of a primer program (Biosciences, Plymouth, USA) and the PC Gene program (Intelligenetics, Geneva, Switzerland). The latter programme was used to localize secondary structures in the RNA. The longest possible secondary structure was prevented from forming by blocking this region using hybridization of the 30-mer primer.

Results

Isolation of cDNA and genomic clone; characteristics of the transcript

Microspores of tobacco at an early stage were isolated as described in Materials and method. Poly(A)$^+$ RNA was prepared from these cells

and used to synthesize a cDNA library. The primary library contained 1.6×10^6 recombinants and was amplified once. Differential screening was done with cDNA from leaves as a negative probe. From the selected microspore-positive and leaf-negative clones, 10 clones were tested for tissue and pollen developmental stage-specific expression by northern analysis. This analysis resulted in 2 microspore-specific cDNA clones; a more detailed characterization of the NTM19 clone (*Nicotiana tabacum* microspore-specific) is reported here.

Figure 1 shows the results of a northern blot containing total RNA prepared from isolated microspores and pollen at different stages of development hybridized with NTM19. Other plant tissues, namely pistil, flower bud of which anthers were removed, seedling, leaf, root tips and stem, were also included in this analysis. No hybridization signal was observed in the lanes with RNA from the bicellular pollen stage or any other plant tissue. The position of the signal corresponded to a transcript length of about 650 nt. The NTM19 cDNA clone has a length of about 500 bp, and is therefore not a full-length clone.
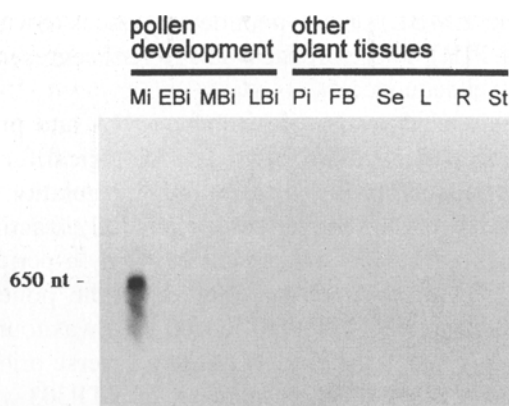

Fig. 1. Northern blot analysis of the expression of gene NTM19. A25 μg portion of total RNA from pollen at different developmental stages and different plant tissues was hybridized with a ^{32}P-labelled NTM19 cDNA probe. Abbreviations: Mic, microspores; EBi, early binucleate pollen; MBi, mid binucleate pollen; LBi, late binucleate pollen; Pi, pistil; FB, flower buds without anthers; Se, seedling; L, leaf; R, root tips; St, stem. The size of the mRNA was determined with the ribosomal bands as markers.

To complete the sequencing of the transcript a corresponding genomic clone was isolated. The available and presumed 3' non-coding region of cDNA clone NTM19 was used as a probe to screen a genomic tobacco library in λ Charon 32. In the non-coding parts of the transcript the restraint on evolutionary divergence is low and therefore this sequence accurately identifies the corresponding gene. The use of a non-coding region as a probe diminishes the chance of isolating a related gene from a gene family. The genomic clone corresponding to the NTM19 cDNA clone was successfully isolated using this procedure.

After the isolation of hybridizing genomic clones, the entire cDNA was used to probe restriction fragments of the selected clones. A subclone was made that covered the region of a genomic clone hybridizing to the cDNA clone. Another subclone overlapping the former one and protruding in the 5' direction was also constructed.

Sequence characterization and start of transcription of the NTM gene

To identify the NTM19 gene, a sequence analysis of the two genomic subclones and the cDNA clone was performed (Fig. 2). The longest open reading frame covering the cDNA region was selected. An alternative open reading frame would have resulted in a 65 amino acid polypeptide, but this did not fit with other criteria such as the position of the transcription start. From the cDNA region and from the corresponding genomic clone an open reading frame was deduced ranging from position + 40 to + 316 bp, relative to the transcription start, encoding a protein of 92 amino acids. The translation initiation region, 6 nucleotides immediately upstream and 3 nucleotides downstream of the ATG start codon, conforms well to the consensus in other plant genes [16] and thus supported the choice of reading frame. The putative polyadenylation signal [12] is underlined in Fig. 2 and it is at a distance of 20 nucleotides followed by a poly(A)tail of 18 nucleotides in the cDNA clone.

```
GATCCAGATTTATAGGGTCCTAATGCGGGTACTGAACACCAGGTGGGAAACAAAAAATAT  -893
ACAGAACAACTCCTTTAGAATTTACAATTTTTGAGCGTGTTGGCTTGGTACGATTCTACT  -833
TTTCATATCTCTCGTCATCTCCTAACTCCTATGGTTCACCAGCCACCGATTAATTATGAC  -773
ACCGCTAACAAAAATCTTGCGACGACATTGAGAGAAATTTCTTTTCATAAATTGGTAATT  -713
CGTACATCATTTATAGGCGTTAGCTATAACCTTTTAGTTAGTGAATACAATACTTTTTGC  -653
TATTATTATGTAACTTTTAGATATGAATTTACTTTCAAAAAAAAAAAAAGGATCGATGTT  -593
GGTTATCAACTAAGGACCAACCACTTTGGACGTCTCACCACTAAGTTAAATAAATCACTT  -533
TGTTCTCGAAAAAAACCCCAAAAGTGTTAAAATGCTTTTCATATCATAATCAAACAACGT  -473
GATTAATAAAATCTATTAAGTTAATAGAAGTAGGGAATAAATCGGGCAAAAGAATTTGAT  -413
ACAAACCAAACCGGTCAAAAAAGCTAGTATTCATATAAATGGACTATACAAGTTAATACC  -353
AGCTAGCAGAAATTAAATAGTTTATTAAGTTGATTACAAAACAATTCCTCATTTAAAAAA  -293
AGTTAATGTAATCAAGAGATCTTTTGCTTCTAATTGATCAGACGAGGACCCCTCTTATTT  -233
ATTTTCTTTTTCATATAAGATTTTGAATAGATATAGGGAAATCTTGTTCACTCTTTATCT  -173
ACTTCAAATTGCATGCATTTTAAGAATTCTCTTTGTATGCAAACTTCAGTATTTATGATT  -113
GACATAAATCAATATTCATATCTTCGATAAAGTTAATAACTCTCCTAATACTTATGAATA  -  53
TCTCTTCCTTTACAACCCTATAAAACCCCCCCACTATAGCTACCTTCATAATTCATCTTAG  +  7
                        ▲

AGTACCAACCCTAAATTTCTTAGTGATTAACCATGGCTAAGAAAAGTCTCACTTTTCTCA  + 67
                            M  A  K  K  S  L  T  F  L      aa9

TTTGCATTTTCCTACTTCTCAATTTATGTTTTGCAATTGAGAACGTAGAAACTATGCAAA  +127
I  C  I  F  L  L  L  N  L  C  F  A ↔I  E  N  V  E  T  M  Q     aa29

AATCGGATTCATCGTCACAAGATAAAGAATTAGATTGGTTTCATCCGTGGTTCCATCCAC  +187
K  S  D  S  S  S  Q  D  K  E  L  D  W̲ F̲ H̲ P̲ W̲ F̲ H̲ P̲        aa49
                                        *  *

ATCCATGGTGGCTACATCCACATCCATGGCCATTCGTTCATCCGCCAATGCCAGCTGGCG  +247
H  P  W  W  L  H  P  H  P  W  P  F  V  H  P  P  M  P  A  G     aa69
*  *  *              *  *  *  *  *  ------------------------

GTTTTCATCATGCATGGCCATTCCCCCATCCACCGATGCCTGCTGGTGGTTTTAAGTTTC  +307
G  F  H  H  A  W  P  F  P  H  P  P  M  P  A  G  G  F  K  F     aa89
-------  (V)              ----------------------------

CTCATCAATAATTTCATCGTCATCCATGGCCATTCATGCATCCACCAGTTCCATCTCCAC  +367
P  H  Q                                                        aa92
CTAAAGGAGACAAGAATTAATTGAAAATATGAAGAGAAGTGTTGGATCAACATCTTATTG  +427
ATCACATATTTTTCTTTAGGTTAATATCTTTAGGATTTATGTCTTAGGTTATTTTTGATA  +487
AAAATTAAAATAAATGATCGTTCTAGGGTAGTTATTATAATTTCTTAGATTTTTCCAAGT  +567
AGCTTTCGATGGTAGAAATGTTATTAATTTGATTCGGCTTATCATGAAATAAAATCCGTA  +627
GTATTATTGCTTTAGCTTTTATGATTTGTAGTTATTTTATGTTGATTGTTCTCCATTT    +687
```

Fig. 2. Nucleotide and deduced amino acid sequence of the NTM19 gene. The TATA box and polyadenylation site are underlined, the transcription start is marked with an arrowhead. The change in amino acid in the cDNA clone from cv. Petit havana relative to the genomic clone from cv. Samsun is bracketed. A possible cleavage site for the signal sequence is indicated by a double arrow. Three repeats in the amino acids with a length of 9, 5 and 4 residues are shown by resp. dashes lines, asterisks and continuous lines. A possible phosphorylated serine is double-underlined.

The transcription start was determined by primer extension analysis (Fig. 3). The first transcribed nucleotide is an adenine (A) at a distance of 39 nt from the start codon, marked in Fig. 2 by an arrowhead and denoted as position +1. This is in accordance with the fact that in plant genes, adenine is the most common nucleotide at the transcription start site [13]. The putative TATA box is 35 nt upstream from the transcription start and underlined in Fig. 2. The distances between TATA box, transcription start and first AUG codon fit well within the range determined for a large number of plant genes [13].

The coding region of the genomic clone is almost identical to that present in the cDNA. Two nucleotide differences are noted at nucleotide positions +74 and +132, and no intervening sequences were seen. The transcribed region of the genomic clone has a length of 658 bp, including
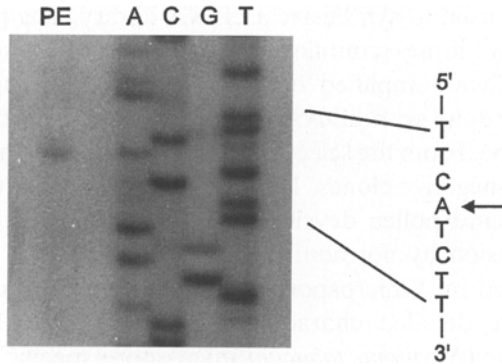

Fig. 3. Identification of the transcription start of the NTM19 gene. Primer extension analysis (PE) and plasmid sequencing (ACGT) were performed using the same NTM19-specific synthetic oligo-nucleotide with sequence 5'-TGCAAAACAT-AAATTGAGAAGTAGGAAAAT-3' (complementary to the nucleotides between positions (+)66 and (+)103 and ^{32}P-labelled). The nucleotide sequence at the 5' end is reported. The arrow indicates the position of the primer extension product.

the 18 nt poly(A) tail that starts at sequence position +641. This is in good agreement to the estimated transcript length of 650 nt determined from the northern blot. Therefore it is inferred that no introns are present in the NTM19 gene.

Database searches did not reveal any significant homology to nucleotide sequences (GenBank, EMBL) or polypeptide sequences (Swiss-Prot, PIR). Therefore the NTM19 gene represents a novel sequence.

A detailed search of available DNA and protein sequence motifs, related to biological functions such as cell organelle targeting, regulation of the cell cycle, and pollen developmental cis-acting elements [8, 40] was carried out. Only a perfect match with a cis-acting element of the pollen-specific gene NTP303 in tobacco [40] was found in the NTM19 gene. In NTM19 a reverse orientation of the AAATGA sequence of NTP303 was found twice. The TCATTT sequence of NTM19 began at positions -706 and -305 from the transcription start. In the case of NTP303 the distance to the transcription start was 100 nt and between the two boxes 29 nt. This cis-acting element sequence was also detected in Bp10 [3], a highly homologous gene to NTP303 in Brassica napus, in a similarly organised manner to the

boxes of NTP303. The sequence box is in fact present twice in all three mentioned genes but differs in its orientation. In the case of NTM19 both boxes are orientated in the reverse direction compared to NTP303, in Bp10 the two AAATGA boxes are directed forward, whereas NTP303 has two boxes, one direct and one oriented oppositely.

Characteristics of the NTM19 protein

The properties of the NTM19 protein were deduced by computer analysis. In Fig. 2 the deduced amino acid sequence of the NTM19 gene is given. The NTM19 protein has a molecular mass of 10.8 kDa and a pI of 6.92. A putative signal transit peptide sequence conforming to the rules formulated by von Heijne [38] is present, this signal is commonly associated with the targeting of membrane proteins or transport of secreted proteins. The cleavage point is between the amino acids alanine and isoleucine at positions 21/22. A presumed phosphorylation site at a serine residue is double-underlined [22]. The hydropathy plot, according to Kyte and Doolittle [17] (Fig. 4), shows a rather smooth hydrophylic and hydrophobic character when the hydrophobic signal sequence portion is excluded.
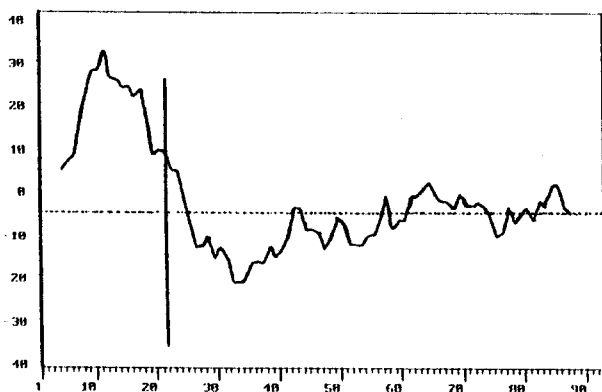

Fig. 4. Hydropathy plot of the NTM19 protein averaged over a window of 9 residues. The putative cleavage site of the signal sequence part is indicated with a vertical line.

The two differences in nucleotide sequence between the genomic and cDNA clone result in a change of one amino acid. One C in the genomic clone is replaced by a T in the cDNA clone, resulting in a change from alanine in the genomic sequence (bracketed at amino acid 74) to valine in the cDNA sequence. Because both amino acids have a similar chemical character, this substitution would not greatly change the properties of the protein. The second difference is at nucleotide position 132, a G in the genomic clone, is replaced by a T in the cDNA clone without affecting the encoded serine residue (amino acid position 31). The sequence differences between genomic and cDNA clone were repeatedly found in independent sequencing experiments. The cDNA library was made from cv. Petit Havanna and the genomic clone from cv. Samsun. This may be the explanation for the sequence differences.

The NTM19 protein contains a perfect direct repeat of 9 amino acids at positions 63–71 and 79–87, and a second one of 5 amino acids at the positions 42–45 and 46–49. Finally, there is a third repeat of 4 amino acids that partly overlaps with the second.

Spatial and temporal expression patterns of NTM19 in anthers

The precise temporal and spatial expression patterns of NTM19 were revealed by in situ hybridization. This technique allows the accurate discrimination between sporophytic and gametophytic localization of expression. Figure 5 shows the localization of NTM19 mRNA in longitudinal sections of tobacco anthers at different developmental stages. The antisense NTM19 RNA probe hybridized specifically with the microspore RNA and not with tapetum RNA. A hybridization signal was not found in anthers with pollen mother cells undergoing meiosis, nor in tetrads, nor in bicellular pollen. The NTM19 sense RNA probe used as a negative control in all experiments gave no hybridization signal.
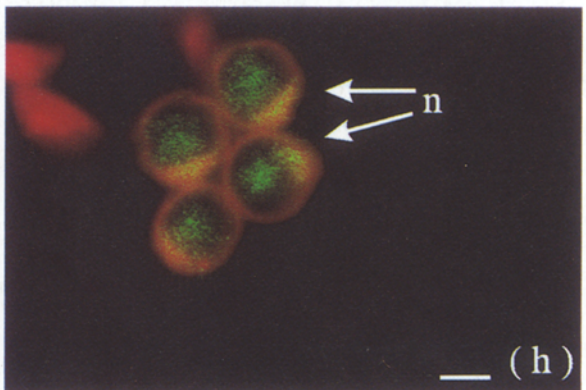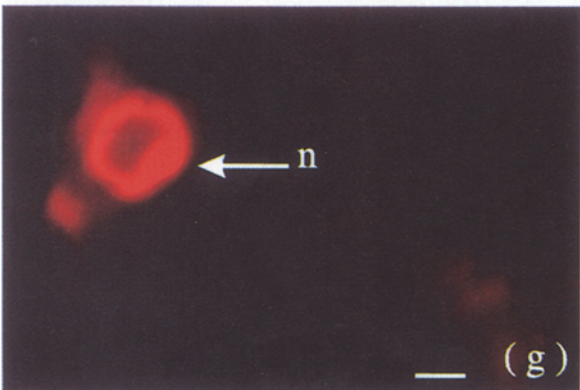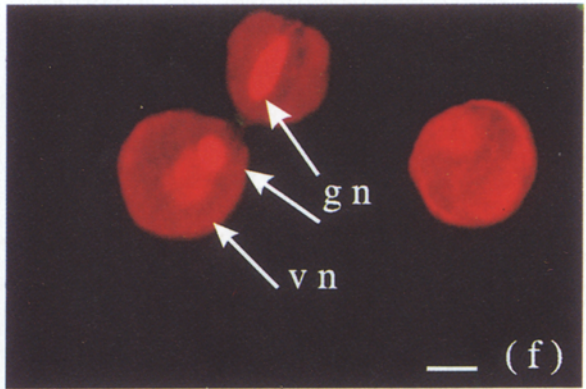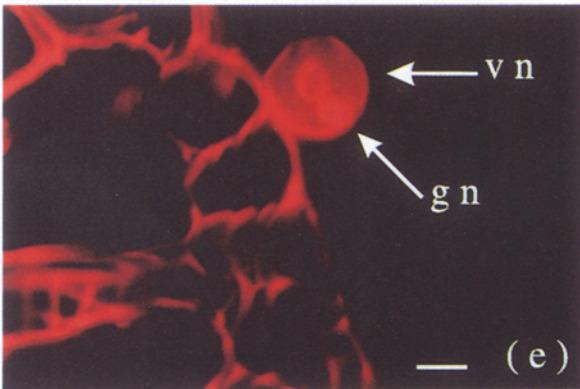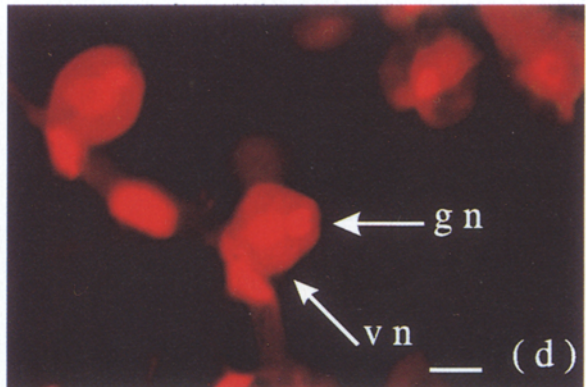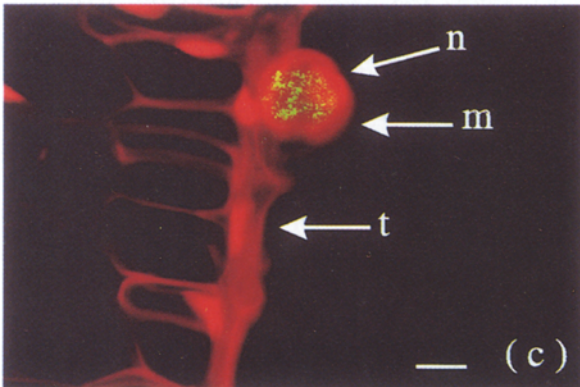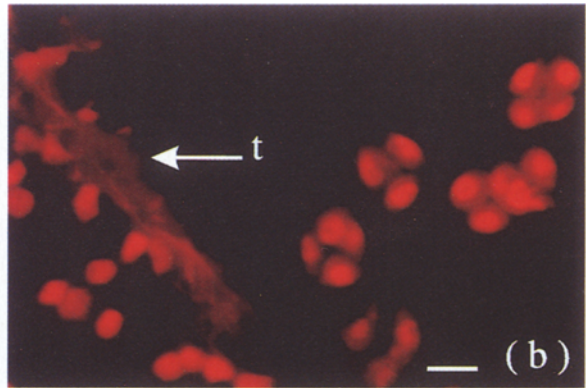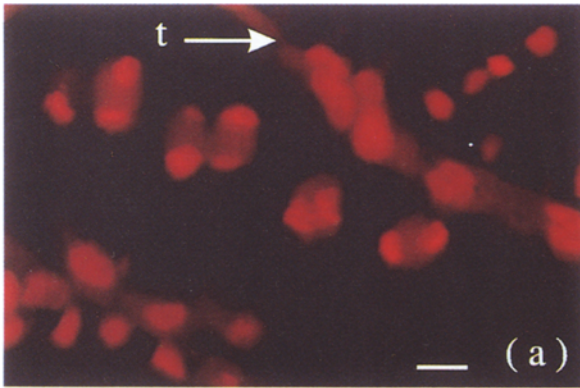
Inheritance of the NTM19 gene and presence in other plant species

There is no consensus in literature yet about the ancestry of the amphidiploid tobacco. *Nicotiana sylvestris* is accepted as an ancestor, but for the second ancestor *N. tomentosiformis* and *N. otophora* are candidates [14, 37].

To determine the inheritance of the NTM19 gene from the putative ancestors, Southern analysis was performed on these three *Nicotiana* species and on *N. tabacum*. Total DNA was digested with *Eco*RI and *Nsi*I, and probed with part of the genomic clone ranging from position −189 to +199 bp. From a restriction map of the original entire genomic clone, it was known that the probe used had to hybridize to a NTM19 *Eco*RI fragment of 4.3 kb. By sequence analysis a second *Nsi*I site was found that closely flanked the probe sequence at the *Eco*RI site and a fragment of 0.4 kb was determined to hybridize with the probe. In the left *Nsi*I panel, the lowest band in the tobacco lane was derived from the NTM19 gene and the other hybridizing bands were derived from related genes. In the tobacco lane with *Eco*RI-digested DNA (Fig. 6, right) at least 6 hybridizing bands were visible. Consequently NTM19 belongs to a small gene family. In both panels the NTM19 band in the tobacco lane was also found in the *N. tomentosiformis* lane but not in the lanes with DNA from the other *Nicotiana* species. Hence, the NTM19 gene originates from *N. tomentosiformis*.

To address the question of whether NTM19 homologues are present in other plant species, various species were chosen in a range of closely and distantly related species. Genomic DNA was
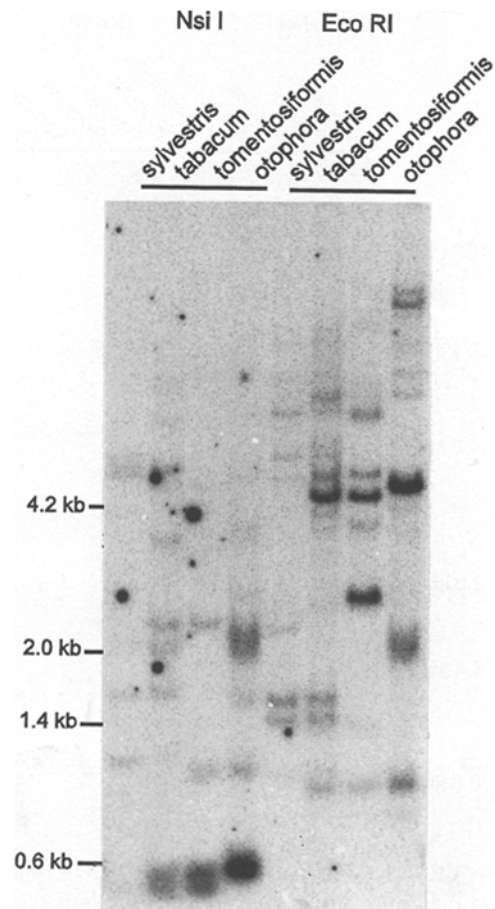

Fig. 6. Southern blot analysis of genomic DNA from tobacco and the putative ancestor species. Southern analysis of 10 µg genomic DNA from *N. sylvestris*, *N. tabacum*, *N. tomentosiformis* and *N. otophora* digested with *Nsi*I or *Eco*RI and probed with a ^{32}P-labelled NTM19 probe containing a 5′ non-coding and coding region. Sizes of DNA marker fragments are given in kb.

extracted, digested and blotted (Fig. 7). The same restriction enzymes were used as described above.

Fig. 5. Localization of TNM19 mRNA during anther development in longitudinal sections of anthers by *in situ* hybridization. a–g: combinations of fluorescence and reflection CLSM images of anther sections during the microspore stage and preceeding and successive pollen developmental stages, determined as described in Materials and methods. Abbreviation: n, nucleus; vn, vegetative nucleus; gn, generative nucleus. Expression of NTM19 analysed in anther sections with the antisense NTM19 probe: (a) pollen mother cells undergoing meiosis. (b) tetrads. (c) and (h) microspore; m, microspore with green spots which represents regions of RNA-/RNA hybridization; t, tapetum. (d) early bicellular pollen. (e) mid-bicellular pollen. (f) late bicellular pollen. (g) control section with microspore with the sense NTM19 of microspore. Bar = 10 µm. In the bicellular pollen 2 nuclei are visible by a more intensive red staining, the mature pollen shows a characteristic discus-shaped generative nucleus.
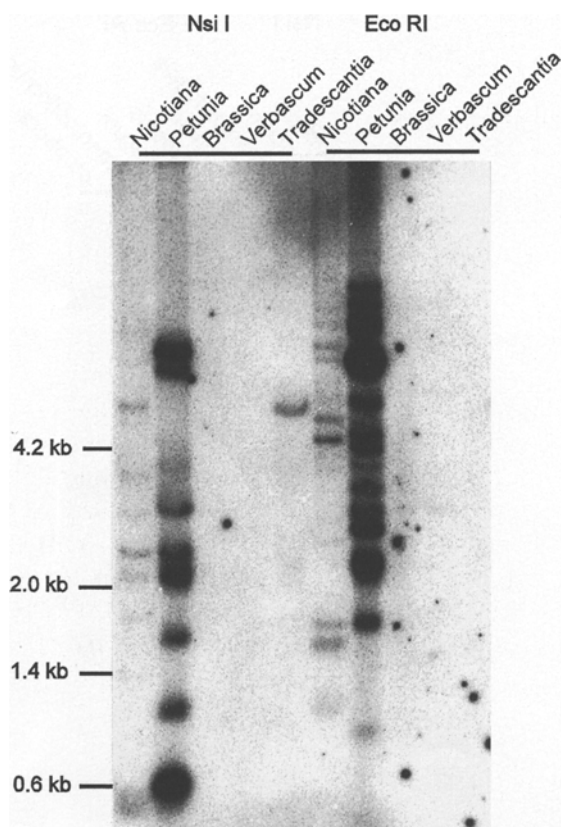
Fig. 7. Southern blot analysis of genomic DNA from tobacco and different plant species. Southern analysis of 10 μg genomic DNA from *N. tabacum, Petunia hybrida, Brassica oleracea, Verbascum thapsus* and *Tradescantia virginiana* was digested with *Nsi*I and *Eco*RI and probed with a ^{32}P-labelled NTM19 probe containing a 5′ non-coding and coding region. Sizes of DNA marker fragments are given in kb.

The *Petunia hybrida* DNA strongly hybridized with the NTM19 probe and a NTM19 homologous family of at least 8 members was seen. The *Brassica oleracea* and *Verbascum thapsus* lanes did not show any signal. For the monocot *Tradescantia virginiana* a single band was found in the *Nsi*I panel indicating a NTM19 homologue in this species.

Discussion

This paper describes the first identification of a microspore-specific gene (NTM19). The NTM19 gene is shown to be exclusively active in the male

reproductive cell of a higher plant during early pollen development. Transcripts were observed only in the unicellular microspore. There is no expression in the sporophytic tissues that represent the majority of cells in the anther, nor in any other tissue of the plant. Expression is also developmently restricted to the unicellular microspore: the gene is inactive during the preceding tetrad stage and the activity disappears at microspore mitosis.

The results show that our experimental approach using isolated microspores, instead of whole anthers, to find microspore-specific genes was effective. The successful identification of NTM19 can be accounted for by the selection of a distinct stage of development and the isolation of the microspores from the anther. All attempts described so far to circumvent the laborious isolation of microspores by using anthers as the starting material for gene identification have failed to isolate genes expressed exclusively in the microspore. In all these cases the high contribution of transcripts from the tapetum to the total amount of anther transcripts resulted in the isolation of several genes expressed mainly in the tapetum. Some authors reported microspore-specific genes like Bp4 [1] and Bp19 [2] in *Brassica napus*. But the term microspore was also used for the bicellular pollen. In contrast, our definition for microspore is for the unicellular stage of pollen development.

The gene family of NTM19 consists of at least 5 members as found in the Southern analysis of genomic DNA. The NTM19 gene itself in the amphidiploid tobacco is inherited from the ancestor *N. tomentosiformis*. The other members of the gene family also indicate similarities between *N. tabacum, N. sylvestris* and *N. tomentosiformis* but not with *N. otophora*. The pollen-specific gene NTP303 in tobacco expressed in mature and germinating pollen [41] also originated from *N. tomentosiformis*. As far as can be judged from these two genes, *N. otophora* is not an ancestor of *N. tabacum*. The NTM19 gene is not unique for the genus *Nicotiana* because it is also found in several other members of the Solanaceae. The hybridisation signal in *Petunia hybrida* is even stronger

than in tobacco and the size of the gene family in increased. In the unrelated monocot *Tradescantia* one NTM19 homologue was found. More species should be analysed to obtain better insight into the phylogenetic situation.

Since searches for homology using DNA and polypeptide databases have not resulted in any known sequence, one can only speculate on the function of the NTM19 protein sequence as deduced by computer analysis. Without the predicted signal peptide, the amino acid composition of the protein indicates an integral membrane localization. Also the presence of a signal peptide points to a membrane positioning. The putative phosphorylation of a serine residue may influence the functioning of the NTM19 protein. If the protein is as short-living as the mRNA, its functioning would be correlated to a specific microspore process.

During the microspore stage the cell prepares for an asymmetrical mitosis and the intine and exine layers are formed [31]. The protein is not expected to play a general role in symmetrical mitosis because the gene is not expressed in actively dividing cells of root tips, as was tested by northern analysis. A role in the formation of the intine layer, the exine structure pattern, or in asymmetrical mitosis is possible.

A comparison of promoter region of NTM19 with known cell cycle related motifs and pollen developmental *cis*-acting elements revealed only homology to a regulatory sequence element from the pollen-specific gene NTP303 in tobacco [40]. However the spacing of this *cis*-acting element sequence in NTM19 is quite different. Therefore it is unclear if these sequence elements have a regulatory effect on the promoter activity of the NTM19 gene.

Further research concerning the localization of the protein and a functional analysis, using an antisense approach to inactivate the gene, may give more insight into the function of NTM19. Experiments with transgenic plants to confirm the NTM19 expression pattern, observed by northern analysis and *in situ* analysis, are in progress. The NTM19 promoter importantly offers the opportunity to manipulate unicellular microspore-

specific expression of genes introduced in the plant by genetic engineering.

Finally, the approach presented in this paper to isolate a gene from highly specified cells, according to localization and developmental period, may contribute to the elucidation of the genetic control of the developmental programme of the male gametophyte.

References

1. Albani D, Robert LS, Donaldson PA, Altosaar I, Arnison PG, Fabijanski SF: Characterisation of a pollen-specific gene family from *Brassica napus* which is activated during early microspore development. Plant Mol Biol 15: 605–622 (1990).
2. 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).
3. Albani D, Sardana R, Robert LS, Altosaar I, Arnison PG, Fabijanski SF: A *Brassica napus* which shows sequence similarity to ascorbate oxidase is expressed in developing pollen. Molecular characterisation and analysis of promoter activity in transgenic tobacco plants. Plant J 2: 331–342 (1992).
4. Brander KA, Kuhlemeier C: A pollen-specific DEAD-box protein related to translation initiation factor eIF-4A from tobacco. Plant Mol Biol 27: 637–649 (1995).
5. Cathala G, Savouret JF, Mendez B, West BL, Karin M, Martial JA, Baxter JD: Laboratory methods. A method for isolation of intact, translationally active ribonucleic acid. DNA 2: 329–335 (1983).
6. Doyle JJ, Doyle JL: Isolation of plant DNA from fresh tissue. Focus (Gibco/BRL) 12: 13–15 (1990).
7. Estruch JJ, Kadwell S, Merlin E, Crossland L: Cloning

and characterization of a maize pollen-specific calcium-dependent calmodulin-independent protein kinase. Proc Natl Acad Sci USA 91: 8837–8841 (1994).

8. 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).

9. Feinburg AP, Vogelstein B: A technique for radiolabelling DNA fragments to high specificity. Anal Biochem 137: 266–267 (1985).

10. Goldberg RB, Beals TP, Sanders PM: Anther development: basic principles and practical applications. Plant Cell 5: 1217–1229 (1993).

11. Jakobsen KS, Breivold E, Hornes E: Purification of mRNA directly from crude plant tissues in 15 minutes using magnetic oligo dT microspheres. Nucl Acids Res 18: 3669 (1990).

12. Joshi CP: Putative polyadenylation signals in nuclear genes of higher plants: a compilation and analysis. Nucl Acids Res 23: 9627–9640 (1987).

13. 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).

14. Kenton A, Parokonny AS, Gleba YY, Bennett MD: Characterization of the Nicotiana tabacum L. genome by molecular cytogenetics. Mol Gen Genet 240: 159–169 (1995).

15. Koltunow AM, Truettner J, Cox KH, Wallroth M, Goldberg RB: Different temporal and spatial gene expression patterns occur during anther development. Plant Cell 2: 1201–1224 (1990).

16. Kozak M: The scanning model for translation: an update. J Cell Biol 108: 229–241 (1989).

17. Kyte J, Doolittle RF: A simple method for displaying the hydropathic character of a protein. J Mol Biol 157: 105–137 (1982).

18. Loenen WAM, Blattner RR: Lambda Charon vectors (Ch 32, 33, 34, 35) adapted for DNA cloning in recombinant-deficient hosts. Gene 26: 171–179 (1983).

19. Maliga P, Sz.-Breznowvitis A, Marton L: Streptomycin resistent plant from callus culture of haploid tobacco. Nature New Biol 244: 29–30 (1973).

20. Mascarenhas JP: Gene activity during pollen development. Annu Rev Plant Physiol Mol Biol 41: 317–338 (1990).

21. Mu J-H, Lee H-S, Kao T-H: Characterization of a pollen-expressed receptor-like kinase gene of Petunia inflata and the activity of its encoded kinase. Plant Cell 6: 709–721 (1994).

22. Pinna LA: Casein kinase 2: an 'eminence grise' in cellular regulation? Biochim Biophys Acta 1054: 267–284 (1990).

23. Reijnen WH, Vanherpen MMA, Degroot PFM, Olmedilla A, Schrauwen JAM, Weterings KAP, Wullems GJ: Cellular localization of a pollen-specific messenger RNA by in situ hybridization and confocal laser scanning microscopy. Sex Plant Reprod 4: 254–257 (1991).

24. Roberts MR, Foster GD, Blundell RP, Robinson SW, Kumar A, Draper J, Scott R: Gametophytic and sporophytic expression of an anther-specific Arabidopsis thaliana gene. Plant J 3: 111–120 (1993).

25. Sambrook J, Fritsch EF, Maniatis T: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989).

26. Sanger F, Nicklen S, Coulsen AR: DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467 (1977).

27. Schrauwen JAM, de Groot PFM, van Herpen MMA, van der Lee T, Reynen WH, Weterings KAP, Wullems GJ: Stage-related expression of messenger-RNAs during pollen development in lily and tobacco. Planta 182: 298–304 (1990).

28. Scott R, Dagless E, Hodge R, Paul W, Soufleri I, Draper J: Patterns of gene expression in developing anthers of Brassica napus. Plant Mol Biol 17: 195–207 (1991).

29. Shen JB, Hsu Fc: Brassica anther-specific genes: characterisation and in situ localization of expression. Mol Gen Genet 234: 379–389 (1992).

30. Short JM, Fernandez JM, Sorge JA, Huse WD: Lambda ZAP: a bacteriophage lambda expression vector with in vitro excision properties. Nucl Acids Res 16: 7583–7600 (1988).

31. Stanley RG, Linskens HF: Pollen. Springer-Verlag, Berlin (1974).

32. Stinson JR, Eisenberg AJ, Willing RP, Pe MP, Hanson DD, Mascarenhas JP: Genes expressed in the male gametophyte of flowering plants and their isolation. Plant Physiol 83: 442–447 (1987).

33. Theerakulpisut P, Xu HL, Singh MB, Pettitt JM, Know RB: Isolation and developmental expression of Bcp1, an anther-specific cDNA clone in Brassica campestris. Plant Cell 3: 1073–1084 (1991).

34. Triezenberg SJ: Primer extension. Preparation and analysis of RNA. In: Sharp PA (ed) Current Protocols in Molecular Biology, unit 8.4.1. Greene Publishing Associates/Wiley Interscience, New York (1992).

35. Tsuchiya T, Toriyama K, Nasrallah ME, Ejiri S: Isolation of genes abundantly expressed in rice anthers at the microspore stage. Plant Mol Biol 20: 1189–1193 (1992).

36. Tupy J, Rihova L, Zarsky V: Production of fertile tobacco pollen from microspres in suspension culture and its storage for in situ pollination. Sex Plant Reprod 4: 284–287 (1991).

37. Van BM, Neuhaus JM, Shinsi H, Ryals J, Meins F: The structure and regulation of homologous tobacco endochitinase genes of Nicotiana sylvestris and Nicotiana tomentosiformis origin. Mol Gen Genet 232: 460–469 (1992).

38. von Heijne G: Sequence Analysis in Molecular Biology: Treasure Trove or Trivial Pursuit. Academic Press, San Diego (1986).

39. Weterings K, Reijnen W, van Aarssen R, Kortstee A, Spijkers J, van Herpen M, Schrauwen J, Wullems G: Characterization of a pollen-specific cDNA clone from

Nicotiana tabacum expressed during microgametogenesis and germination. Plant Mol Biol 18: 1101–1111 (1992).

40. Weterings K: Pollen Gene Regulation. Thesis, University of Nijmegen (1994).

41. Weterings K, Reijnen W, Wijn G, van de Heuvel K, Appeldoorn N, de Kort G, van Herpen M, Schrauwen J, Wullems G: Molecular characterization of the pollen-specific genomic clone NTPg303 and *in situ* localization of expression. Sex Plant Reprod 8: 11–17 (1995).

42. 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).

43. Yie Y, Wei Z, Tien P: A simplified and reliable protocol for plasmid DNA sequencing: fast miniprep and denaturation. Nucl Acid Res 21: 361 (1992).