ISSN 0582-9879 Acta Biochimica et Biophysica Sinica 2004, **36**(7): 467–476 **CN** 31-1300/Q

Isolation and Partial Characterization of a Novel Pollen-specific cDNA with Multiple Polyadenylation Sites from Wheat

Yong-Feng JIN* and Teng-Fei BIAN

Institute of Biochemistry, Zhejiang University, Hangzhou 310029, China

Abstract A novel pollen-specific full-length cDNA clone *PSG076* was isolated using suppression subtractive hybridization and 5′/3′ RACE techniques. *PSG076* was shown to exhibit multi-site polyadenylation by sequencing the 3′ ends of the cDNAs. At least six transcripts with different length were produced from the single gene based on different poly(A) tail attachment sites. However, polyadenylation consensus sequence AAUAAA was not seen at the 3′-untranslated sequence. *PSG076* contained a 299 bp 5′ untranslated region and an open reading frame of 663 bp encoding a 221 amino acid peptide with pI of 4.31. A blast search revealed that this sequence did not show a significant similarity to any genes deposited in the public database. Southern blot indicated that *PSG076* was a single copy gene. Northern blot and RT-PCR analysis indicated that *PSG076* transcripts showed specific expression in mature pollen, and weak or undetectable signals in uninucleate microspore, immature seed, stem, young leave, root and ovary. Further analysis of the expression pattern in gametophyte showed that *PSG076* transcripts were undetectable in uninucleate, binucleate microspore and pollen at early stage, and were first detectable and increased rapidly at middle and late stages of pollen development with the maximum level in mature pollen and also expressed in germinating pollen *in vivo*, suggesting that *PSG076* might play a role in pollen germination and pollen tube growth in addition to its function in maturation. The evidences gathered in this work indicated that the six different transcripts from the single gene were differentially expressed during pollen development.

Key words wheat (*Triticum aestivum* L.); pollen-specific cDNA; suppression subtractive hybridization; multiple-site polyadenylation; differential expression

Wheat is the foremost staple food crop that offers both calories and proteins to a large global population. Wheat (hexaploid AABBDD genome, 16 billion bp) is a genetically complex, self-pollinating plant with bisexual flowers that produce short-lived pollen. Very little is known about the molecular biology of its gametophyte development despite a longstanding interest in hybrid seeds. Most of our information is from the studies on a few model and crop plants such as *Arabidopsis*, tobacco, vegetable crops, and maize [1–10]. There is very little information on gene concerned specifically with pollen and anther development in wheat with the exception of partial EST clones for which no functional information is available. Only one example was described about wheat gene function during gametophyte development that three apparently homologous genes (*TAA1a*, *TAA1b* and *TAA1c*) were characterized and expressed only at specific stages of pollen development as the microspore wall thickened during the progression of free microspores into vacuolated-microspores. However, *TAA1* expresses specifically within the sporophytic tapetum cells but not gametophyte cells [11]. The study of wheat anther and pollen development is immensely important not only for the understanding of gene regulation in the sexual reproduction of wheat, but also because of its potential application in agriculture such as male-sterility.

Most eukaryotic mRNAs are characterized by the presence of a poly(A) tail and a series of adenine residues are attached to the 3′ untranslated region post transcription.

Received: April 15, 2004 Accepted: May 25, 2004

This work is supported by a grant from the National Natural Science Foundation of China (No. 30370759)

*Corresponding author: Tel, 86-571-86971414; Fax, 86-571-86971501; E-mail, jinyf@zju.edu.cn

Poly(A) tail plays an important role in many aspects of RNA function and metabolism, particularly in mRNA translation and turnover [12]. A significant proportion of mRNAs display multiple polyadenylation sites [13]. In human, at least 22% of mRNAs undergo alternative polyadenylation [14], the effect of which may vary in a differentiation or development regulated fashion [12]. A number of genes in plants have been found to contain multiple poly(A) attachment sites yielding different transcripts from a single gene [15–20]. Two transcripts produced from the same gene by alternative splicing of the *Arabidopsis* U1-70K pre-mRNA are differentially expressed in different tissues [15]. A cell wall invertase gene (*Incw-1*) in maize is expressed as two transcripts with different length in the 3′ untranslated region. Sugars differentially modulate an unusual mode of the two transcripts through the 3′ untranslated region [16]. Alternative polyadenylation generates three low-pI α-amylase mRNAs with differential expression in barley in various tissues [17]. However, the four different transcripts generated off subunit A gene of the vacuolar type H^+-ATPase of *Arabidopsis* are not significantly differentially regulated in response to salt, chilling, or etiolation in axenically grown seedlings [18]. The function of multiple-site polyadenylation is still not clear. It is possible that varying length of 3′ untranslated region may have a regulatory role in mRNA transport, stability, and/or translational efficiency of the transcripts [15].

In this paper, we described the isolation of the novel pollen-specific gene *PSG076* from wheat via suppression subtractive hybridization and 5′/3′ RACE techniques. *PSG076* was pollen-specifically expressed, and its transcripts were first detectable and increased rapidly at early-middle stages of pollen development with the maximum level in mature pollen. The gene expression product had no significant sequence similarity to proteins deposited in the public database. *PSG076* was shown to exhibit multi-site polyadenylation by sequencing the 3′ ends of the cDNAs. At least six different length transcripts were produced from the single gene based on different poly(A) tail attachment sites. In this study we explored the possibility that differentially polyadenylation transcripts of *PSG076* could be regulated in pollen development.

Materials and Methods

Materials

Seeds of the winter cultivar Ferdinand were germinated on soil in pots for one month in a growth chamber. Temperature was 15 °C during the day and 12 °C during night, with a day-length of 16 h, 60% humidity and a light intensity of 10,000 lux. Anthers were collected to obtain microspore and pollen of different developmental stages. The determination of developmental stage and isolation of microspore and pollen were carried out as described previously [21]. All tissues were frozen immediately in liquid nitrogen and stored at –80 °C.

DNA and RNA isolation

Total RNA was isolated from different wheat tissues using Qiagen RNeasy plant mini kit according to the manufacturer's protocol. Genomic DNA for Southern blot was isolated as described (http://www.protocol-online.org/prot/Molecular_Biology/DNA/DNA_Extraction_Purification). RNA was stored at –80 °C and genomic DNA was stored at 4 °C. Plasmid DNA was purified using Qiagen plasmid isolation kit.

Suppression subtractive hybridization (SSH)

Suppression subtractive hybridization [22] was done using the PCR-select cDNA subtraction kit (BD Biosciences Clontech, USA) according to the manual. Subtracted cDNAs were cloned into the T/A vector pCR2.1 (Invitrogen, USA). Primary analysis of clones in order to identify those carrying inserts included the colony PCR, where minute amount of bacteria was used instead of DNA templates. The PCR mixture (50 μL) consisted of 0.2 μM of M13 reverse and forward primers each, 0.2 mM of each dNTP, 2 mM $MgCl_2$ and 1 U of *Taq* polymerase in an appropriate buffer. PCR parameters were: 94 °C for 2 min, 30 cycles of 94 °C for 1 min, 50 °C for 1 min, 72 °C for 2 min 30 s and final extension at 72 °C for 5 min. The differential screening was carried out by reverse Northern hybridization (RNH). DNA fragments obtained by colony PCR were dotted onto nylon membranes and fixed with 0.4 M NaOH for 20 min. Comparative RNH was performed with two or more identical replicas under stringent condition (7% SDS, 1% bovine serum albumin, 1 mM EDTA, 0.25 M Na_2HPO_4, pH 7.2) with 100 μg/ml denatured salmon sperm DNA at 65 °C overnight with equivalent amount of labeled cDNA probes of nearly equal specific activity for each experiment. Membranes were washed twice at 65 °C in 0.1×SSC for 20–30 min.

Generation of full-length cDNA

Full-length cDNA clones were obtained using 5′/3′ RACE cDNA synthesis kit. First strand cDNA was synthesized from total RNA (0.5 μg) isolated from mature

pollen using SuperScript II reverse transcriptase primed with oligo(dT)12–18 ancher primer according to the manufacturer's instructions (Invitrogen). The 5′ and 3′ RACE specific primers were designed according to cDNA sequence isolated by SSH and EST. The 5′ RACE primers were 5′-GACTGCTCGAGCTAATGTTGCCGGT-3′ and 5′-GGCTCCGCTTCCGGTTTGCTGCTTGCCT-3′, and the 3′-RACE primer was 5′-AGAGTCGGGGAG-CCGGCAGCGGA-3′. The DNA products were purified by using Qiagen PCR purification kit and cloned into pCR2.1 vector (Invitrogen) according to the manufacturer's instructions. Isolation of recombinant clones was carried out using standard procedures [23].

Sequencing

Sequencing of selected clones was done using ALF semiautomatic DNA sequencer. Homology search for sequences of selected clones was performed using basic local alignment sequence tool (BLAST) at http://www.ncbi.nlm.nih.gov/blast.

Analysis of gene expression by RT-PCR

Total RNA was reverse transcribed using SuperScript II RT and the resulting single-stranded cDNA product was treated with DNase at 37 °C for 30 min. PCR amplification was carried out using cDNA from 5 ng or 10 ng of total RNA template in each reaction. Amplification conditions were 35 cycles of 94 °C for 30 s, 65 °C for 30 s and 72 °C for 30 s, followed by one cycle of 72 °C for 10 min. Wheat glyceraldehyde-3-phosphate dehydrogenase gene (*GAPDH*) transcripts were amplified as an external control. The specific primers for *PSG076* and *GAPDH* gene were shown in Table 1.

Primer pairs for RT-PCR for differential expression were designed with one primer upstream coding region and another anchored in the poly(A) tail of the transcript (Table 1). In order to determine the optimal annealing temperature, PCR was conducted at the annealing temperature of 35, 40, 45, 50, and 55 °C, respectively, to determine the appropriate annealing temperature of PCR. The PCR was allowed to proceed in each case 30 s at 94 °C, 1 min at 50 °C, 30 s at 72 °C for 35 cycles, and extended at 72 °C for 10 min. PCR products were then electrophoresised on 2.0% agarose gels and transferred to nylon membrane for Southern hybridization with the cDNA probe.

RNA gel blot analyses

15 μg total RNA each sample from different tissues of wheat plants was subjected to electrophoresis on 1.2% agarose (*W*/*V*)/10% formaldehyde (*V*/*V*) denaturing gels, transferred to Hybond-N^+ membrane (Amersham), and UV cross-linked. 450 bp of the gene-specific probe from the

Table 1 Primer sets used for the RT-PCR analysis of wheat cDNA *PSG076*

cDNA clone	Primer	Sequence	Predicted product size (bp)
PSG076-A	Forward	5′-AGAGTCGGGGAGCCGGCAGCGGA-3′	212
	Reverse	5′-GACTGCTCGAGCTAATGTTGCCGGT-3′	
PSG076-B	Forward	5′-GAAGCGCCGAAGGGCGGCATGTT-3′	117
	Reverse	5′-GACGGTCCGGACACGATGA-3′	
3′-end transcript-1	Forward	5′-GATGCCCTGCTGCTCT-3′	291
	Reverse	5′-TTTTTTTTTCAAATTCA-3′	
3′-end transcript-2	Forward	5′-GATGCCCTGCTGCTCT-3′	306
	Reverse	5′-TTTTTTTTTGATTCACA-3′	
3′-end transcript-3	Forward	5′-GATGCCCTGCTGCTCT-3′	315
	Reverse	5′-TTTTTTTTTCAAATGGA-3′	
3′-end transcript-4	Forward	5′-GATGCCCTGCTGCTCT-3′	371
	Reverse	5′-TTTTTTTTTACATGAATT-3′	
3′-end transcript-5	Forward	5′-GATGCCCTGCTGCTCT-3′	380
	Reverse	5′-TTTTTTTTTACAGTACC-3′	
3′-end transcript-6	Forward	5′-GATGCCCTGCTGCTCT-3′	405
	Reverse	5′-TTTTTTTTTAGATAAATG-3′	
GAPDH	Forward	5′-CAACGCTAGCTGCACCACTAACT-3′	350
	Reverse	5′-GACTCCTCCTTGATAGCAGCCTT-3′	

partial coding and whole 3′ untranslated regions was generated using PCR and labeled using random primed DNA labeling kit from Invitrogen according to the manufacturer's instructions. [α-^{32}P]dCTP labeled DNA probes were purified using Sephadex G50 before hybridization. High SDS buffer [7% SDS (*W*/*V*), 1% bovine serum albumin, 1 mM EDTA, 0.25 M Na_2HPO_4, pH 7.2] and 100 μg/ml denatured salmon sperm DNA were used for hybridization at 65 °C. Membrane stringent washing was carried out twice in 2×SSC, 0.1% SDS at 65 °C for 30 min and twice with 0.1×SSC, 0.1% SDS at 65 °C for 30 min. Membranes were exposed to Kodak films using intensifiers at –80 °C for 5–10 days.

Southern hybridization

6 μg genomic DNA was digested with a variety of restriction endonucleases under the conditions recommended by the enzyme supplier. Fragments were resolved on 0.8% agarose gels and transferred to a Hybond-N^+ membrane (Amersham). Probe labeling was carried out as RNA hybridization. DNA gel blot was hybridized and washed using the same condition as employed for RNA gel blot analysis.

Results

Isolation and identification of pollen-specific cDNA clones from wheat

Our primary aim is to isolate microspore-specific genes from wheat using SSH [22]. For the isolation of microspore-specific genes using SSH, four cDNA pools were synthesized by solid-phase RT-PCR from the mRNA of the freshly isolated wheat unicellular microspores (stage A), or the microspores that had been starved for 2 and 4 days (stages B_2 and B_4, respectively), or the mature pollen (stage C). SSH resulted in two subtracted pools: a forward pool (S), where freshly isolated microspores (A) and mature pollen (C) cDNAs were subtracted against cDNAs of starved microspores (B_2 and B_4), and a reverse pool (R), where mixture of B_2 and B_4 cDNAs were subtracted against A and C cDNAs. More than 850 EST clones were generated by cloning the cDNAs of the forward pool (S). Two rounds of macroarray RNH were carried out to confirm their microspore specificity. Total 8 clones were identified, which showed signals only when the membrane was hybridized with RNA prepared from unicellular microspores (A). Preliminary trials showed that the expression of some cDNAs in wheat tissues was too low to be detected by RNA gel blotting with total RNA. We identified the spatial and temporal expression pattern by RT-PCR. Using the individual pairs of gene-specific primers, RT-PCR was performed on mRNA samples from a variety of different wheat tissues. WSM076 mRNA showed low levels in microspore and immature seed and at high levels in mature pollen, but was undetectable in other tissues, suggesting that this clone might be reproductive-organ or male gametophyte specific. These results are not consistent with that of reverse Northern hybridization, possibly because the reverse Northern hybridization is more non-specific especially for the non-rich or low-level mRNA. However, WSM076 is noticeable because it may be pollen-specific.

Isolation of a full-length cDNA

Therefore WSM076 was subjected to more detailed investigation. At first its full-length cDNA (named *PSG076*) was isolated using RACE-PCR technology based on t he WSM076 cDNA fragment. The GC content is as high as 77%. *PSG076* (GenBank accession No. AY451239) contained a 299-bp 5′ untranslated region, and an open reading frame of 795bp. We think this is close to full-length, because a stop codon TAA is located in frame 46 nucleotides upstream from the ATG initiation codon, and the cDNA size nearly agrees with Northern blot analysis data. The predicted protein encoded by *PSG076* is supposed to have a pI of 4.31 and lacks a signal peptide. Computer searches were performed through comparing the nucleotide and deduced amino acid sequence of the *PSG076* gene with sequences in GenBank databases. No significant similarity was found between *PSG076* and any known DNA and proteins deposited in the GenBank database. It does, however, show very limited sequence similarity to a small number of proteins, including hypothetical protein (No. At2g03630) from *Arabidopsis* (40% identity over 60 aa length). In addition, this gene was completely homology with known EST cDNA clones isolated cDNA library from wheat pre-anthesis spike or spike at the flowering date. All ESTs have same nucleotide sequences in corresponding region, although the 3′ extension and polyadenylation sites were different. However, *PSG076* lacks the AUG context consensus sequence c(a/c)(A/G)(A/C)c*AUG*GCG of monocots, which is thought to be important for AUG codon recognition [24].

No putative polyadenylation consensus sequence AATAAA was found at the 3′-untranslated sequence of *PSG076* gene. But some AATAAA-like motifs with 5 out of 6 bases match (AATCAA, AAAAAA, AATAGA), or 4 out of 6 bases match (AAATTA) were found in the sequences upstream of the polyadenylation sites in *PSG076*

gene. Although some plant genes do have the unaltered AATAAA motif, plants seem to be more divergent in this motif and other sequences up and downstream of the poly (A) cleavage site might compensate for the lack of the AATAAA sequence [25].

Specific expression pattern

Because the 3′-UTR was generally the most divergent region, about 450 bp 3′ non-coding and coding regions of *PSG076* were used as gene-specific probes for Northern hybridization. Preliminary RNA gel blot analysis about organ-specific expression pattern of *PSG076* showed a high expression of about 1.8 kb transcript in mature pollen, and weak or undetectable signals in microspore, immature seed, stem, young leave, root and ovary (Fig 1). Because *PSG076* was only expressed in pollen but not other tissues based on Northern hybridization using 3′-end gene-specific probes, we were interested in the gametophytic expression pattern of *PSG076* transcripts. To determine the gametophytic expression pattern of *PSG076* transcripts, total RNA was extracted from microspore and developing pollen at eight developmental stages (one day one stage). *PSG076* transcripts were undetectable in uninucleate and binucleate microspore and pollen at early stage, and were first detectable and increased rapidly at middle and late stages of pollen development with the maximum level in mature pollen. This result was consistent with RT-PCR analysis (Fig. 2). Therefore, *PSG076* showed a typical gametophytic pattern of expression similar to late pollen-specific genes in tobacco and its transcript was accumulated after mitosis [26]. Because wheat pollen maturation and germination experiment in vitro is some difficult, we check them *in vivo*. *PSG076* showed relatively high expression in ovary with stigma and pollen and a slight expression until in immature seed one week after anthesis but undetectable in immature seed two week after anthesis (Fig. 2), which suggests that *PSG076* may also play a role during the germination of wheat pollen and pollination [27].

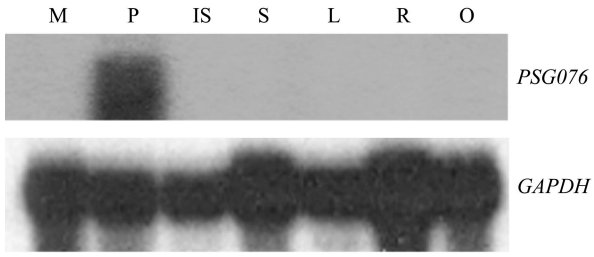

Fig. 1 Northern blot of the *PSG076* cDNA in wheat

15 μg total RNA each sample from microspores (M), mature pollens (P), immature seeds (IS), stems (S), leaves (L), roots (R), ovaries (O) were hybridized with the *PSG076* 3′-end 450 bp probes.

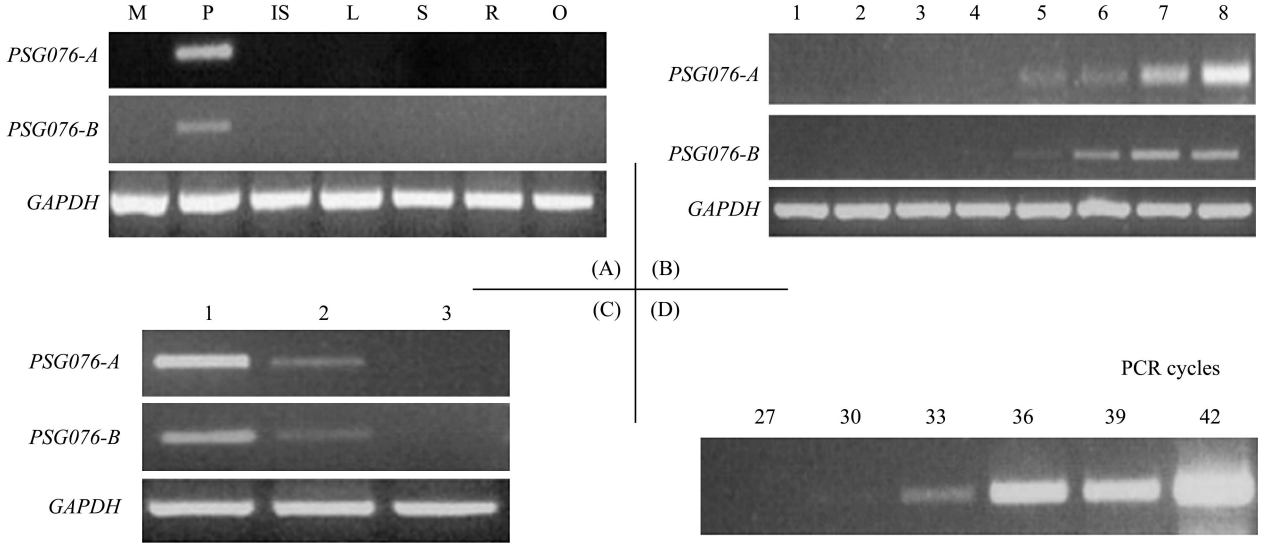

Fig. 2 Specific expression of the cDNA *PSG076* in wheat by RT-PCR analysis

(A) Tissue-specific expression analysis. M, microspores; P, mature pollens; IS, immature seeds; S, stems; L, leaves; R, roots; O, ovaries. (B) RT-PCR analysis during the development of male gametophytes *in vivo* from unicellular microspore to mature tri-cellular pollen (1–8 days). (C) RT-PCR analysis during pollen germination *in vivo*. 1, 2 and 3, ovary with germinating pollen or immature seed 1, 7 and 14 days after anthesis. (D) All amplifications were in the linear range between 27 and 41 cycles and 35 cycles were used as standard. 10 ng (A) and 5 ng (B, C and D) of total RNAs were reverse transcribed for RT-PCR analysis.

Genomic DNA gel blot analysis

Genomic DNA gel blot analysis was performed to determine the copy number of this gene in the wheat genome. About 450 bp 3′ non-coding and coding regions of *PSG076* were used as gene-specific probes for Southern hybridization to estimate the number of *PSG076*-related genes in the wheat genome. The *PSG076* 3′ probe did not include restriction sites of any of enzymes utilized (*Bam*HI, *Hin*dIII, *Pst*I, *Sac*I). After hybridization and washing at high stringency, we observed only one band with the 3′ probe (Fig. 3). In addition, polymerase chain reaction failed to detect an intron length polymorphism (data not shown). Therefore, we conclude that *PSG076* gene should be a single copy gene.

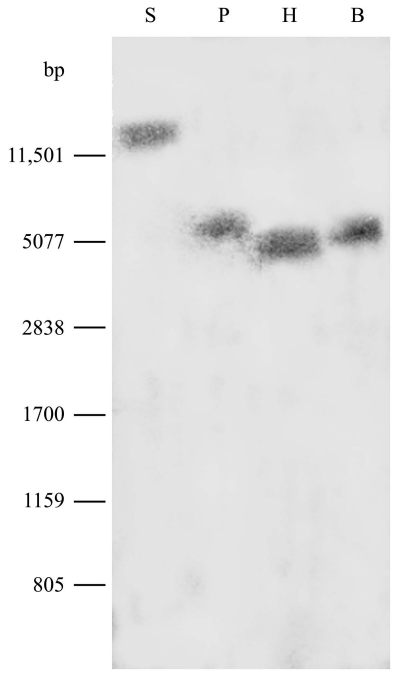

Fig. 3 Southern hybridization analysis of wheat genomic DNA with probes from *PSG076*

DNA samples were digested with different restriction enzymes. S, *Sac*I; P, *Pst*I; H, *Hin*dIII; B, *Bam*HI.

Differentially polyadenylated transcripts in pollen development

Sequencing the 3′ ends of 3′ RACE PCR and ESTs indicated that at least six alternative polyadenylation sites were present (Fig. 4), which was conferred by the fact that several bands were generated in 3′ RACE PCR. The 3′ RACE products using a universal amplification primer annealed to poly(A)$^+$ tail and 5′ specific primer led to at

```
076-1  TGATAAACACCGTCAATGGATCGCCGGCGGCAAAAACCGGCTAGGAGGGACGCTCATCGTGTCCGGACCGTCACCGGCAACATTAGCTCGAGCAGTCGGAAATGGCGGCCAGAGACAGGA
076-2  TGATAAACACCGTCAATGGATCGCCGGCGGCAAAAACCGGCTAGGAGGGACGCTCATCGTGTCCGGACCGTCACCGGCAACATTAGCTCGAGCAGTCGGAAATGGCGGCCAGAGACAGGA
076-3  TGATAAACACCGTCAATGGATCGCCGGCGGCAAAAACCGGCTAGGAGGGACGCTCATCGTGTCCGGACCGTCACCGGCAACATTAGCTCGAGCAGTCGGAAATGGCGGCCAGAGACAGGA
076-4  TGATAAACACCGTCAATGGATCGCCGGCGGCAAAAACCGGCTAGGAGGGACGCTCATCGTGTCCGGACCGTCACCGGCAACATTAGCTCGAGCAGTCGGAAATGGCGGCCAGAGACAGGA
076-5  TGATAAACACCGTCAATGGATCGCCGGCGGCAAAAACCGGCTAGGAGGGACGCTCATCGTGTCCGGACCGTCACCGGCAACATTAGCTCGAGCAGTCGGAAATGGCGGCCAGAGACAGGA
076-6  TGATAAACACCGTCAATGGATCGCCGGCGGCAAAAACCGGCTAGGAGGGACGCTCATCGTGTCCGGACCGTCACCGGCAACATTAGCTCGAGCAGTCGGAAATGGCGGCCAGAGACAGGA
       >
076-1  AAAGGAAGTGAGAAATTATGGATCGATTCGAAGGACAAAAAACACAATCAAAACCAGCTTTTAAATTTTCTTTGCCCCAGAGTGGTGTCTTGTAATAGATTGCTTCCCTTTTTGGTCACA
076-2  AAAGGAAGTGAGAAATTATGGATCGATTCGAAGGACAAAAAACACAATCAAAACCAGCTTTTAAATTTTCTTTGCCCCAGAGTGGTGTCTTGTAATAGATTGCTTCCCTTTTTGGTCACA
076-3  AAAGGAAGTGAGAAATTATGGATCGATTCGAAGGACAAAAAACACAATCAAAACCAGCTTTTAAATTTTCTTTGCCCCAGAGTGGTGTCTTGTAATAGATTGCTTCCCTTTTTGGTCACA
076-4  AAAGGAAGTGAGAAATTATGGATCGATTCGAAGGACAAAAAACACAATCAAAACCAGCTTTTAAATTTTCTTTGCCCCAGAGTGGTGTCTTGTAATAGATTGCTTCCCTTTTTGGTCACA
076-5  AAAGGAAGTGAGAAATTATGGATCGATTCGAAGGACAAAAAACACAATCAAAACCAGCTTTTAAATTTTCTTTGCCCCAGAGTGGTGTCTTGTAATAGATTGCTTCCCTTTTTGGTCACA
076-6  AAAGGAAGTGAGAAATTATGGATCGATTCGAAGGACAAAAAACACAATCAAAACCAGCTTTTAAATTTTCTTTGCCCCAGAGTGGTGTCTTGTAATAGATTGCTTCCCTTTTTGGTCACA

076-1  AATGTCGTCCACAACGATGAATTTGA----------
076-2  AATGTCGTCCACAACGATGAATTTGAGATTTCTGTGAATC----------
076-3  AATGTCGTCCACAACGATGAATTTGAGATTTCTGTGAATCATCCATTTGA----------
076-4  AATGTCGTCCACAACGATGAATTTGAGATTTCTGTGAATCATCCATTTGACATTTCAGATCTTCTGAATGTATATAACCAGAGGAAAAACAGGTTAAATTCATGTAA----------
076-5  AATGTCGTCCACAACGATGAATTTGAGATTTCTGTGAATCATCCATTTGACATTTCAGATCTTCTGAATGTATATAACCAGAGGAAAAACAGGTTAAATTCATGTAAGGTACTGTA----
076-6  AATGTCGTCCACAACGATGAATTTGAGATTTCTGTGAATCATCCATTTGACATTTCAGATCTTCTGAATGTATATAACCAGAGGAAAAACAGGTTAAATTCATGTAAGGTACTGTATCTA

076-1  ----------(A)n
076-2  ----------(A)n
076-3  ----------(A)n
076-4  ----------(A)n
076-5  ----------(A)n
076-6  CCTAACTAGATGCAGACATTTATCT----------(A)n
```

Fig. 4 Alignment of the 3′ untranslated region nucleotide sequences of *PSG076* cDNAs and RACE products from pollens and EST from anthers

">" indicated the coding region stop codon; possible degenerated poly(A)$^+$ signals AAATTA, AAAAAA, AATCAA, AATAGA were underlined. transcript 076-1, RACE products; transcript 076-2, RACE products; transcript 076-3, EST (BJ244814); transcript 076-4, RACE products; transcript 076-5, EST (BQ171145); transcript 076-6, EST (BJ244444).

least three major products that were detectable in both ethidium-bromide-stained gels and hybridization. The 3′-untranslated sequence extended 280–400 bases downstream from the stop codon TGA (Fig. 4). One apparent polyadenylation variability might be due to artifacts from the cDNA synthesis, like an imprecise annealing of the oligo-dT primer happening in an AU-rich 3′ terminus of the mRNA. However, sequence analysis ruled out the possibility because there was no long stretch of A residue in this region. Several examples of multiple polyadenylated genes have also emerged in *Arabidopsis* [15,18], tobacco [20], barley [17], and maize [16]. Since the *PSG076* gene produced at least six different transcripts, experiments were designed to determine if differentially polyadenylated transcripts could be regulated during the pollen development.

Having established *PSG076* expression pattern in pollen development, an assay was designed to examine the response of the six individual length transcripts in pollen development. The goal was to determine if the six transcripts were differentially expressed in pollen development. A PCR based approach was taken to evaluate transcript levels for the six different *PSG076* mRNAs. Primer pairs were designed (Table 1) that would only amplify one of the six transcripts. This was accomplished by anchoring the downstream primer in the poly(A) tail. The upstream primer was located in the coding region of the *PSG076* cDNA. Anchoring the downstream primer in the poly(A) tail prevented that oligonucleotide from priming at any other site since 9 out of 17–18 nucleotides were T at the 5′ end. The transcript-1 amplification product, which is the smallest and nested within the other five, acted as a probe to detect all six transcripts. This nesting was accomplished to avoid differences in signal during hybridization based on different probe length and labeling efficiency [18]. PCR amplifications in different annealing temperature with the appropriate primer pairs were performed prior to the expression analysis to ensure that the primer pairs amplified the correct transcript only. *PSG076* transcripts were amplified and transferred to membranes. Hybridization was performed at high stringency with the transcript-1 amplification product as probe. The annealing temperatures were determined when only a single band of the expected size was amplified using primer pairs and hybridized.

The differential expression of *PSG076* transcripts was evaluated in pollen development. cDNAs were generated from pollen in different stages. The results of this experiment indicated that all six transcripts showed increases in pollen development with different expression level and pattern. 076-1, 076-2, 076-4 transcripts showed strong increase in pollen development, whereas 076-1,076-2 transcripts expressed earlier than 076-4. 076-1 transcripts were first detectable on the 6th day before anthesis and increased rapidly at middle stages of pollen development, while 076-4 transcripts were first detectable on the 4th day before anthesis. 076-3, 076-5, 076-6 transcripts only showed slight expression in pollen development, mainly in mature pollen (Fig. 5). In summation, the six *PSG076* transcripts may show differential expression during pollen development.

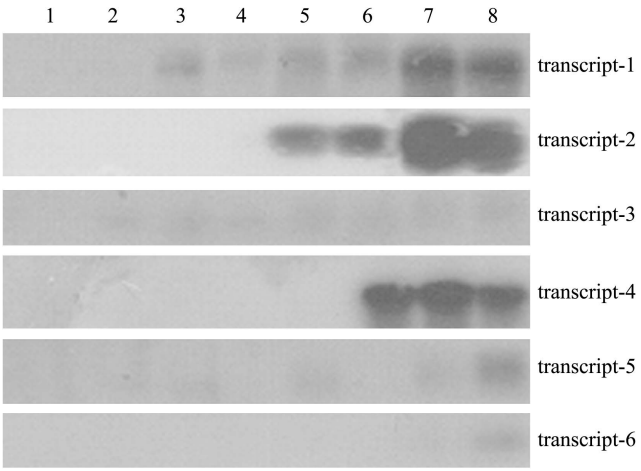

Fig. 5 Differential expression of *PSG076* transcripts during pollen development

Expression pattern of the *PSG076* cDNA by RT-PCR analysis during the development of male gametophytes *in vivo* from unicellular microspore to mature tri-cellular pollen. 1–8, mature tri-cellular pollen in 1–8 days respectively. Primer pairs for RT-PCR for differential expression were used in a PCR to amplify a cDNA fragment of the corresponding gene after reverse transcription of 5 ng of total RNA. PCR products were then electrophoresised and transferred to nylon membrane for Southern hybridization.

Discussion

In angiosperm plants, sexual reproduction requires the formation of male and female gametophytes, which produce male and female gametes [26]. Genes expressed in anthers for male gametogenesis can be divided into two groups (so-called early and late genes) based on their expression timing for convenience. The early genes become active soon after the completion of meiosis and the levels of their transcripts are reduced or undetectable in mature pollen [27]. Early genes are either expressed in both the sporophyte (e.g. tapetum) and the gametophyte (microspores and/or young pollen [28,29], or only in the tapetum [11,29,30], in the microspores and/or young pollen only [31]. Tobacco *NTM19* gene [31] and the *Brassica napus BP4* and *BP19* [32,33], *BnM3.4* gene [34],

which fall into this early gene category, have been isolated. The genes in the second group start to express after microspore mitosis and the expression increases until maturity. These late genes have been isolated from a number of plant species including the *Arabidopsis* [5], tobacco [6–8], maize [9], tomato [10], Chinese cabbage [35], sunflower [36,37]. Most of these gene products are accumulated abundantly in pollen grains and are involved in pollen maturation or germination [27]. Most of our information on the molecular biology of male gametophyte development is from the studies on a few model and crop plants such as *Arabidopsis*, tobacco, vegetable crops, and maize [1–10]. Wheat is one of most important cereal crops in the world, however, the molecular biology of wheat male gametophyte development remain poorly understood.

In this study, a novel pollen-specific cDNA *PSG076* was isolated and the search for homology of this gene in DNA and polypeptide databases have not resulted in any known sequences. Northern blot and RT-PCR analysis indicated *PSG076* transcripts showed pollen-specific. *PSG076* transcripts were undetectable in uninucleate and binucleate microspore and pollen at early stage, and were first detectable and increased rapidly at middle and late stages of pollen development with the maximum level in mature pollen and were also expressed at high level in germina-ting pollen *in vivo*, suggesting that *PSG076* might play a role in pollen germination and pollen tube growth in addition to its function in maturation. *PSG076* may be potentially important for application in agriculture such as male-sterility. However, further experimental work has to be carried out to identify its function. Searches for homology using DNA and polypeptide databases indicated *PSG076* did not show a significant similarity to any genes depo-sited in the public database, including dicot *Arabidopsis* and monocot rice. Additionally, no ESTs from other species except wheat and maize show high homology with *PSG076*, which indicates that *PSG076* protein is possibly species-specific and only functions in wheat and closely related species.

Temporal and spatial regulation of anther- or pollen-specific gene expression is primarily controlled at the transcriptional level [2]. This was confirmed by transgenic approaches using chimeric gene fusions between the 5′ sequence of an anther-specific gene and either the β-glucuronidase, diphtheria toxin, or barnase gene [38]. However, many stages of gamete development in plant and animal species proceed almost without transcriptional activity and depend mainly upon translation of presynthesized mRNAs. Also, early stages of pollen germination and growth in various plant species are associated with a massive activation of protein synthesis, which is apparently independent of transcription [27]. Thus, in these species, post-transcriptional control of gene expression is very important for gamete development. Despite the importance of translation of presynthesized mRNAs in the contribution of sexual reproduction, little attention has been paid to elucidate the mechanisms underlying post-transcriptional regulation of pollen gene expression [39,40]. Many mRNA species from different plants can be modulated in their translation efficiency by signals encoded in the 5′- or 3′-UTR during pollen development or germination [40]. The presence of regulatory elements in the 5′ or 3′ end of the gene could also be a potential mechanism by which this protein is post transcriptionally regulated in pollen gene expression, including mRNA stability and mRNA translation efficiency. Also of significant interest are the recent data from rice cell cultures that implicate the 3′ UTR as a major determinant of mRNA stability of the α-amylase3 gene [41]. The presence of multi site polyadenylation, relatively long 5′ untranslated region and a poor AUG context suggests that *PSG076* could be the subject of posttranscriptional regulation [24,25].

Most of our knowledge of the multi-site polyadenylation has come from studies in animals and viruses [12]. There are only a few examples about multi-site polyadenylation and expression regulation in plants. Several examples have recently emerged where differentially polyadenylated genes have transcripts that are post transcriptionally regulated based on sugar sensing. Specific α-amylase genes from barley produced alternative mRNAs with a 17-base 3′ extension (extension 1) or a 17-base extension beyond this (extension 2), and the extended sequences could play a functional role in α-amylase expression [17]. The long transcript of the *Arabidopsis* U1-70K pre-mRNA is more abundant in flowers, suspension cultures, and leaves, whereas a relatively high level of the short transcript is detected in roots [15]. But four different transcripts generated off subunit A gene of the vacuolar type H^+-ATPase of *Arabidopsis* are not significantly differentially regulated in response to salt, chilling, or etiolation in axenically grown seedlings [18]. No multi-site polyadenylation and their differential expression have been reported in pollen development. However, some evidence indicates that alternative polyadenylation is involved in regulation of gene expression during gamete development and now widely recognized as mechanisms of translational regulation in animals [42,43]. For example, mouse lamin B3 expression appears restricted to spermatogenic cells, whose gene structure indicates that lamin B3 is generated by differential splicing and alternative polyadenylation from lamin

B2 [42]. A mouse RING-finger protein (XYbp) gene has been identified to participate in spermatogenesis, which generates a ubiquitously expressed transcript of 4.2 kb and a testis-specific one of 2.8 kb, processed by an alternative polyadenylation mechanism from a non-canonical polyadenylation signal [43]. In this study, expression analysis indicates that the six different transcripts generated from the single gene are differentially regulated during pollen development. Multi site polyadenylation appears to be a mechanism for posttranscriptional control of gene expression in pollen development.

Acknowledgements

We are grateful to Alisher Touraev, Alisher S. Tashpulatov and Erwin Heberle-Bors (Institute of Microbiology and Genetics, Vienna University) for their advices and technical assistances.

References

1 Chaudhury AM. Nuclear genes controlling male fertility. Plant Cell, 1993, 5: 1277–1283
2 Goldberg RB, Beals TP, Sanders PM. Anther development: Basic principles and practical application. Plant Cell, 1993, 5: 1217–1229
3 McCormick S. Male gametophyte development. Plant Cell, 1993, 5: 1265–1275
4 Yang WC, Sundaresan V. Genetics of gametophyte biogenesis in *Arabidopsis*. Curr Opin Plant Biol, 2000, 3: 53–57
5 Gupta R, Ting JT, Sokolov LN, Johnson SA, Luan S. A tumor suppressor homolog, AtPTEN1, is essential for pollen development in *Arabidopsis*. Plant Cell, 2002, 14: 2495–2507
6 Tebbutt SJ, Rogers HJ, Lonsdale DM. Characterization of a tobacco gene encoding a pollen-specific polygalacturonase. Plant Mol Biol, 1994, 25: 283–297
7 Brander KA, Kuhlemeier C. A pollen-specific DEAD-box protein related to translation initiation factor eIF-4A from tobacco. Plant Mol Biol, 1995, 27: 637–649
8 Steiner C, Bauer J, Amrhein N, Bucher M. Two novel genes are differentially expressed during early germination of the male gametophyte of *Nicotiana tabacum*. Biochim Biophys Acta, 2003, 1625: 123–133
9 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, 1994, 91: 8837–8841
10 Muschietti J, Dircks L, Vancanneyt G, McCormick S. LAT52 protein is essential for tomato pollen development: Pollen expressing antisense LAT52 RNA hydrates and germinates abnormally and cannot achieve fertilization. Plant J, 1994, 6: 321–338
11 Wang A, Xia Q, Xie W, Dumonceaux T, Zou J, Datla R, Selvaraj G. Male gametophyte development in bread wheat (*Triticum aestivum* L.): Molecular, cellular, and biochemical analyses of a sporophytic contribution to pollen wall ontogeny. Plant J, 2002, 30: 613–623
12 Edwalds-Gilbert G, Veraldi KL, Milcarek C. Alternative poly(A) site selection in complex transcription units: Means to an end? Nucleic Acids Res, 1997, 25: 2547–2561
13 Gautheret D, Poirot O, Lopez F, Audic S, Claverie JM. Alternate polyadenylation in human mRNAs: A large scale analysis by EST clustering. Genome Res, 1998, 8: 524–530
14 Beaudoing E, Freier S, Wyatt JR, Claverie JM, Gautheret D. Patterns of variant polyadenylation signal usage in human genes. Genome Res, 2000, 10: 1001–1010
15 Golovkin M, Reddy AS. Structure and expression of a plant U1 snRNP 70K gene: Alternative splicing of U1 snRNP 70K pre-mRNAs produces two different transcripts. Plant Cell, 1996, 8: 1421–1435
16 Cheng WH, Taliercio EW, Chourey PS. Sugars modulate an unusual mode of control of the cell-wall invertase gene (Incw1) through its 3′untranslated region in a cell suspension culture of maize. Proc Natl Acad Sci USA, 1999, 96: 10512–10517
17 Skadsen RW, Knauer NS. Alternative polyadenylation generates three low-pI alpha-amylase mRNAs with differential expression in barley. FEBS Lett, 1995, 361: 220–224
18 Manen JF, Simon P. A possible explanation for the multiple polyadenylation sites in transcripts coding for a winged-bean leghemoglobin. Planta, 1993, 191: 289–292
19 Magnotta SM, Gogarten JP. Multi site polyadenylation and transcriptional response to stress of a vacuolar type H^+-ATPase subunit A gene in *Arabidopsis thaliana*. BMC Plant Biol, 2002, 2(1): 3
20 Klahre U, Hemmings-Mieszczak M, Filipowicz W. Extreme heterogeneity of polyadenylation sites in mRNAs encoding chloroplast RNA-binding proteins in *Nicotiana plumbaginifolia*. Plant Mol Biol, 1995, 28: 569–574
21 Indrianto A, Barinova I, Touraev A, Heberle-Bors E. Tracking individual wheat microspores *in vitro*: Identification of embryogenic microspores and body axis formation in the embryo. Planta, 2001, 212: 163–174
22 Diatchenko L, Lau YF, Campbell AP, Chenchik A, Moqadam F, Huang B, Lukyanov S *et al*. Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proc Natl Acad Sci USA, 1996, 93: 6025–6030
23 Sambrook J, Fritsch EF, Maniatis T. Molecular Cloning: A Laboratory Manual. 2nd ed. New York: Cold Spring Harbor Laboratory Press, 1989
24 Joshi CP, Zhou H, Huang X, Chiang VL. Context sequences of translation initiation codon in plants. Plant Mol Biol, 1997, 35: 993–1001
25 Rothnie HM. Plant mRNA 3′-end formation. Plant Mol Biol, 1996, 32: 43–61
26 Dickinson HG. The regulation of alternation of generation in flowering plants. Biol Rev, 1994, 69: 419–442
27 Mascarenhas JP. Molecular mechanisms of pollen tube growth and differentiation. Plant Cell, 1993, 5: 1303–1314
28 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, 1993, 3: 111–120
29 Charbonnel-Campaa L, Lauga B, Combes D. Isolation of a type 2 metallothionein-like gene preferentially expressed in the tapetum in *Zea mays*. Gene, 2000, 254: 199–208
30 Kapoor S, Kobayashi A, Takatsuji H. Silencing of the tapetum-specific zinc finger gene *TAZ1* causes premature degeneration of tapetum and pollen abortion in *Petunia*. Plant Cell, 2002, 14: 2353–2367
31 Oldenhof MT, de Groot PF, Visser JH, Schrauwen JA, Wullems GJ. Isolation and characterization of a microspore-specific gene from tobacco. Plant Mol Biol, 1996, 31: 213–225
32 Albani D, Robert LS, Donaldson PA, Altosaar I, Arnison PG, Fabijanski SF. Characterization of a pollen-specific gene family from *Brassica napus* which is activated during early microspore development. Plant Mol Biol, 1990, 15: 605–622
33 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, 1991, 16: 501–513

34 Fourgoux-Nicol A, Drouaud J, Haouazine N, Pelletier G, Guerche P. Isolation of rapeseed genes expressed early and specifically during development of the male gametophyte. Plant Mol Biol, 1999, 40: 857–872

35 Park BS, Park YD, Kim HU, Jin YM, Kim HI. BAN103, a pollen-preferential gene, from Chinese cabbage and its promoter activity. Mol Cells, 2002, 14: 150–157

36 Baltz R, Domon C, Pillay DT, Steinmetz A. Characterization of a pollen-specific cDNA from sunflower encoding a zinc finger protein. Plant J, 1992, 2: 713–721

37 Evrard JL, Nguyen I, Bergdoll M, Mutterer J, Steinmetz A, Lambert AM. A novel pollen-specific α-tubulin in sunflower: Structure and characterization. Plant Mol Biol, 2002, 49: 611–620

38 Mariani C, de Beuckeleer M, Truettner J, Leemans J, Goldberg RB. Induction of male-sterility in plants by a chimaeric ribonuclease gene. Nature, 1990, 347: 737–741

39 op den Camp RGL, Kuhlemeier C. Phosphorylation of tobacco eukaryotic translation initiation factor 4A upon pollen tube germination. Nucl Acids Res, 1998, 26: 2058–2062

40 Hulzink RJM, de Groot PFM, Croes AF, Quaedvlieg W, Twell D, Wullems GJ, van Herpen MMA. The 5′-untranslated region of the *ntp303* gene strongly enhances translation during pollen tube growth, but not during pollen maturation. Plant Physiol, 2002, 129: 342–353

41 Chan MT, Yu SM. The 3′ untranslated region of a rice α-amylase gene functions as a sugar-dependent mRNA stability determinant. Proc Natl Acad Sci USA, 1998, 95: 6543–6547

42 Furukawa K, Hotta Y. cDNA cloning of a germ cell specific lamin B3 from mouse spermatocytes and analysis of its function by ectopic expression in somatic cells. EMBO J, 1993, 12: 97–106

43 Parraga M, del Mazo J. XYbp, a novel RING-finger protein, is a component of the XY body of spermatocytes and centrosomes. Mech Dev, 2000, 90: 95–101

Edited by
Zong-Yang WANG