The Plant Journal (2003) **34**, 697–707

Sperm cells of *Zea mays* have a complex complement of mRNAs

Michele L. Engel[1], Annie Chaboud[2,3], Christian Dumas[2] and Sheila McCormick[1,*]
[1]*Plant Gene Expression Center, United States Department of Agriculture-Agriculture Research Service and UC-Berkeley, 800 Buchanan St., Albany, CA 94710, USA,*
[2]*Laboratory of Plant Reproduction and Development, Ecole Normale Supérieure de Lyon, UMR 5667-CNRS-INRA-ENS-Université Lyon1, Lyon, France, and*
[3]*Flow Cytometry Facility, Ecole Normale Supérieure de Lyon, UMR5665, CNRS-ENS, Lyon, France*

Received 6 January 2003; revised 3 March 2003; accepted 5 March 2003.
*For correspondence (fax +1 510 559 5678; e-mail sheilamc@nature.berkeley.edu).

Summary

Although double fertilization in angiosperm was discovered in 1898, we still know nothing about the proteins that mediate gamete recognition and fusion in plants. Because sperm are small and embedded within the large vegetative cell of the pollen grain, mRNAs from sperm are poorly represented in EST databases. We optimized fluorescence-activated cell sorting (FACS) in order to isolate *Zea mays* sperm free of contaminating vegetative cell cytoplasm, and constructed a cDNA library. Sequencing of over 1100 cDNAs from the unamplified library revealed that sperm have a diverse complement of mRNAs. Most transcripts were singletons; the most abundant was sequenced only 17 times. About 8% of the sequences are predicted to encode secreted or plasma membrane-localized proteins and are therefore candidates that might mediate gamete interactions. About 8% of the sequences correspond to retroposons. Plant sperm have condensed chromatin and are thought to be transcriptionally inactive. We used RT-PCR and *in situ* hybridization to determine when selected sperm mRNAs were transcribed. Sperm transcripts encoding proteins involved in general cell functions were present throughout pollen development and were more abundant in tricellular pollen than in sperm cells, suggesting that these transcripts were also present in the larger vegetative cell. However, several transcripts, which encode proteins that are most similar to hypothetical *Arabidopsis* proteins, appeared to be present exclusively in the sperm cells inside mature pollen, but were already present in unicellular microspores. This suggests that certain transcripts might be transcribed early during pollen development and later partitioned into the sperm cells.

Keywords: maize, gamete, cDNA library, retroposons, whole-mount *in situ* hybridization.

Introduction

In flowering plants, each product of male meiosis undergoes two mitotic divisions in order to form the three-celled gametophyte (pollen grain). The first mitosis yields a larger vegetative cell that later forms the pollen tube and a smaller generative cell enclosed within the vegetative cell cytoplasm. The generative cell undergoes mitosis to yield two sperm cells. Thus, the generative cell and sperm cells are plasma membrane-bound (but without cell walls) and are enclosed within the cytoplasm of the vegetative cell (reviewed in McCormick, 1993). The sperm are not motile, but are transported into the pollen tube cytoplasm via the pollen tube, which grows by tip extension through the female tissue (reviewed in Lord and Russell, 2002). The egg cell and the central cell, fusion partners for the two sperm, are located within the embryo sac in the ovule (reviewed in Grossniklaus and Schneitz, 1998). Upon arrival at the embryo sac, the pollen tube discharges its passenger sperm into one of the two synergid cells of the embryo sac. This synergid degenerates and releases the sperm (Higashiyama *et al*., 2001). One sperm then fuses with the central cell to give rise to the primary endosperm cell, while the other fuses with the egg cell to give rise to the zygote (Faure *et al*., 1994; Kranz *et al*., 1998). In animals, the initial recognition events and fusion are mediated by

© 2003 Blackwell Publishing Ltd

secreted or plasma membrane-bound proteins (reviewed in Evans, 2001). It is unknown whether plant gametes are similar to animal gametes in this respect.

In maize, it is possible to isolate sperm, eggs, and central cells (Breton *et al.*, 1995; Faure *et al.*, 1994; Kranz *et al.*, 1998). Such isolated cells have been used to show that the timing of sperm–egg and sperm–central cell fusions were different, and to show that the barrier to polyspermy is maintained *in vitro* (Faure *et al.*, 1994), and to monitor the calcium ion fluctuations occurring upon sperm–egg fusion (Antoine *et al.*, 2000). *In vivo*, the events leading to delivery of the two sperm to an embryo sac have been extensively studied (reviewed in Franklin-Tong, 1999; Higashiyama *et al.*, 2001). Genetic (Roman, 1948; Rusche *et al.*, 1997) and morphological (Russell, 1984) data indicate that the two sperm in a pollen grain can, in some cases, be distinguished and can have different fusion partner preferences (Russell, 1985).

How could differences between the sperm arise? The generative cell might unequally partition transcripts or proteins between the two sperm, or, alternatively, each sperm cell might express its own complement of RNA and proteins. Because sperm have condensed chromatin, it has, generally, been stated that the vegetative nucleus of the pollen grain is more transcriptionally active than are the generative or sperm nuclei (reviewed in McCormick, 1993). Indeed, most pollen-specific mRNAs and/or promoters that have been studied are vegetative cell-specific (McCormick, 1993; Taylor and Hepler, 1997) and are not expressed in the generative cell or sperm. However, sperm are, probably, not quiescent because transmission electron microscopy images reveal mitochondria, endoplasmic reticulum, ribosomes, vesicles, dictyosomes, and microtubules (reviewed in Mogensen, 1992). Partially purified maize sperm synthesized RNA and protein (Matthys-Rochon *et al.*, 1994; Zhang *et al.*, 1993). The transcriptional inhibitor actinomycin D had no effect on the protein profile of lily generative cells, but ^{35}S-methionine incorporation into proteins was abolished when generative cells were incubated with the translational inhibitor cycloheximide (Blomstedt *et al.*, 1996). Lily generative cells contain at least one transcript that encodes a gamete-surface protein, LGC1 (Xu *et al.*, 1999b). Together, these results suggest that the generative and the sperm cells can translate proteins from pre-existing mRNA, independent of the vegetative cell.

Sperm cells represent only a small fraction of the volume of a pollen grain. Although abundant mRNAs from sperm might be represented in cDNA libraries prepared from pollen, further analysis would be required to determine their expression profiles. Thus, the complexity of the mRNA and protein profiles within sperm cells was largely unexplored, and where and when sperm mRNAs were transcribed was largely unknown. Therefore, we obtained highly pure (fluorescence-activated cell sorted) sperm of maize and constructed a high-quality cDNA library. Limited sequencing of this library shows that sperm have a complex complement of mRNA. We used RT-PCR and *in situ* hybridization analysis to determine the expression profiles for selected cDNAs from this library. Several transcripts are apparently present in both the vegetative cell and the sperm cells. But some, specifically those that are most similar to proteins predicted as hypothetical in the *Arabidopsis* database, appear to be present exclusively in the sperm cells in mature pollen, although they are transcribed earlier during pollen development. Such transcripts may be partitioned into the sperm.

Results and discussion

Isolation of sperm and analysis of purity

Pollen was ruptured by osmotic and pH shock (Dupuis *et al.*, 1987). Centrifugation onto a Percoll™ cushion yielded a fraction enriched in sperm (most amyloplasts pellet under these conditions), but this fraction was still highly contaminated with debris from the vegetative cell cytoplasm and burst sperm (Figure 1a). The Percoll™-purified sperm were labeled with Hoechst dye and subjected to flow cytometry. Cells with preserved morphology could be sorted from debris based on their DNA staining and light-scattering properties. These highly pure sperm (Figure 1b) were competent for fusion with isolated eggs.

To assess whether the fluorescence-activated cell sorted (FACS)-purified sperm were contaminated with any vegetative cell RNA, we used RT-PCR and primers specific to *Zmc13*, a robustly expressed vegetative cell-specific transcript (Hanson *et al.*, 1989). The *Zmc13* transcript was present in the total pollen RNA, but not in the sperm RNA (Figure 2a). A serial dilution PCR using total pollen RNA indicated that we could have detected contamination of the sperm RNA by vegetative cell RNA from less than 10 pollen grains (Figure 2b). We therefore concluded that the sperm RNA was free of vegetative cell mRNA.

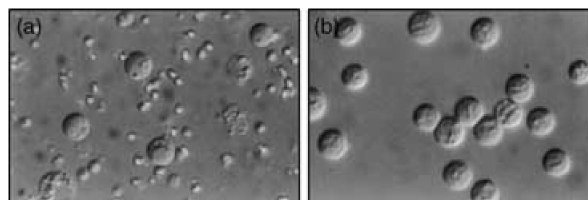

Figure 1. Purified maize sperm.
(a) Percoll-purified sperm.
(b) FACS-purified sperm; DIC microscopy (Optiphot-2 microscope, Zeiss International, Thornwood, NY, USA).

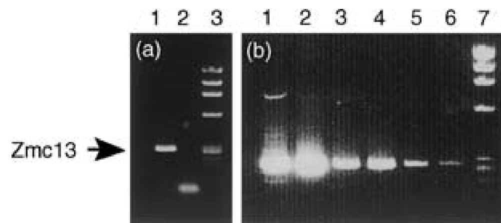

Figure 2. FACS-sorted sperm are pure.
(a) RT-PCR, using primers to the vegetative cell-specific transcript *Zmc13*, was performed on 100 ng of total pollen RNA (lane 1) or on total sperm RNA (lane 2). Lane 3, molecular weight markers.
(b) RT-PCR was performed on a dilution series of total pollen RNA with *Zmc13*-specific primers. Lane 1, 100 ng; lane 2, 50 ng; lane 3, 20 ng; lane 4, 10 ng; lane 5, 5 ng; lane 6, 1 ng; lane 7, molecular weight markers. One hundred nanograms is equivalent to the amount of RNA isolated from 800 pollen grains.

cDNA library construction and analysis

We obtained 1.8 μg of total RNA from 1 million FACS-purified sperm. This yield is similar to that from yeast cells (3 μg per 10^6 cells; http://pingu.salk.edu/~forsburg/plasmids.html), but is less than the typical yield (10–15 μg per 10^6 cells; http://www.medizin.uni-tuebingen.de/med_virologie/exp/deutsch/Useful_remarks.html) from many types of animal cells. We used 800 ng of this RNA to construct a size-selected (300 bp to 4 kb) cDNA library. It was possible that the sperm cell RNA complement would be of limited complexity. To ascertain if this was true, we sequenced from the unamplified library. Table S1 shows a summary of our sequencing results to date, and includes accession numbers, functional category, cluster size and best BLASTP match. This project is ongoing and updated information can be found at http://www.pgec.usda.gov/McCormick/mclab.html, or by searching GenBank (http://www.ncbi.nlm.nih.gov/) with the query '*Zea mays* sperm cell'.

Maize sperm cells have diverse message populations. Currently, the most prevalent transcript (encoding a protein similar to a low-temperature/salt stress-induced protein) was sequenced only 17 times. In other organisms, such as humans (Miller, 2000; Ostermeier *et al.*, 2002), *Drosophila* (Andrews *et al.*, 2000), and the ascidian *Ciona intestinalis* (Inaba *et al.*, 2002), the sperm or testes also contain diverse transcripts. In *Drosophila*, 16% of the cDNA sequences did not align with predicted or known genes, indicating that cDNA libraries from gametic tissues are an excellent resource for genome annotation.

Of the maize sperm cDNAs categorized for this paper, approximately 19% had no database matches and had only short open reading frames. Further sequencing of these inserts might reveal that the sequenced portions correspond to the 5′ or 3′ UTRs of genes not yet in EST databases. Alternatively, some of these transcripts might encode non-coding RNAs. Non-coding RNAs are polyadenylated transcripts that are involved in regulating many processes, including transcription and chromosomal silencing (MacIntosh *et al.*, 2001). These transcripts were not analyzed further. The remaining transcripts had obvious open reading frames. About 30% of these cDNAs were full length, or were judged to be so, based on size of insert and/or alignment with the closest match in the databases. The full-length cDNAs range in size from 500 bp to 2.5 kb. Other cDNAs correspond to internal regions of known or predicted proteins, indicating that the corresponding mRNAs were truncated.

Interestingly, 5% of the sperm cDNAs match genes in the *Arabidopsis* genome that are annotated (http://www.arabidopsis.org/info/agilinks.html) as encoding hypothetical proteins, that is, they are not represented by ESTs from any tissue. These candidates might correspond to rare mRNAs, but that is unlikely given the limited amount of sequencing we did. Thus, these cDNAs likely correspond to transcripts found only in male gametes. About 4% of the cDNAs have long open reading frames, of 90 amino acids or more, but do not have significant homology (1e-8 or greater) to anything in the databases (as of November 2002). Such cDNAs might correspond to maize-specific or monocot-specific genes that have little resemblance to anything in the *Arabidopsis thaliana* genome sequence. We searched with the 5′-end sequences of the maize sperm cDNAs, and this might have reduced our ability to find matches in *Arabidopsis*. Compositional gradients in rice genes reduced the probability of finding a match between rice and *Arabidopsis*, especially when the query sequence was near the 5′ end of the gene (Wong *et al.*, 2002). Furthermore, BLAST scores from alignments between genomic and cDNA sequences were affected by the presence of introns in the genomic sequence. Some sperm cDNAs had low-significance BLAST scores against the rice genome sequence. The significance of these matches may increase as the rice genome is better annotated and as these sperm cDNAs are completely sequenced.

The most abundant class of sperm cDNAs (approximately 8% of the transcripts) encodes components of retroelements. Retroelements are very abundant in the maize genome (Wessler, 2001), and the PREM-1 and PREM-2 retroposons are highly expressed during early stages of maize pollen development (Turcich and Mascarenhas, 1994; Turcich *et al.*, 1996). Retroposons are also highly expressed during animal spermatogenesis (Miller *et al.*, 1999; Pasyukova *et al.*, 1997), and reverse transcriptase activity has been reported in mature mouse spermatozoa (Giordano *et al.*, 2000).

Numerous transcripts encode proteins likely to be involved in DNA synthesis or chromosome structure. The DNA in plant sperm is highly condensed, relative to that in the vegetative cell nucleus (reviewed in Southworth, 1992). In lily, the generative cell synthesizes variant histones that

are probably used to condense the sperm DNA (Ueda *et al.*, 2000); however, the transcripts for these variant histones were present only in generative cells and were not found in lily sperm (Xu *et al.*, 1999a). In maize sperm, we found cDNAs for histones H2A, H2B, H3 and H4, but not for any variant histones that might represent gamete-specific isoforms. Maize sperm might not contain transcripts for the variant histones, or, perhaps, maize sperm use an alternative method for chromosome condensation. Interestingly, several maize sperm cDNAs correspond to Hmgd1, a member of the high-mobility group proteins that play roles in chromosome condensation (Thomas and Travers, 2001), transcription, replication and recombination (Grasser *et al.*, 1997). Several maize sperm cDNAs encode SMT3 (Mannen *et al.*, 1996), a ubiquitin-like protein first identified as a suppressor of a MIF2 mutant in yeast. MIF2 is similar to CENP-C; thus, SMT3 may be involved in the function and/or structure of the kinetochore.

Many maize sperm cDNAs predicted to encode cytoplasmic proteins are similar to genes expected to be widely expressed (e.g. kinases, phosphatases, clathrin assembly carrier protein, secretory carrier membrane protein, various GTP-binding proteins, ATPases, nucleotide translocators, ADP ribosylation factors, CAAX prenyl protease, and actin depolymerization factor). Sperms are not quiescent (reviewed in Mogensen, 1992), and these transcripts are likely required for general cellular metabolism. However, there were also a few surprises. Although transmission electron microscopy indicates that maize sperm lack plastids (Wagner and Dumas, 1989), sperm contain transcripts for two chloroplast enzymes: phytoene desaturase, which is involved in the plastid-localized carotenoid pathway (Li *et al.*, 1996); and neoXanthin cleavage enzyme, which is involved in abscisic acid biosynthesis (Qin and Zeevaart, 1999). Maize sperm are reported to have few nuclear pores (Southworth *et al.*, 1988), but sperm contain transcripts for a Sec13-like protein, implicated in nuclear pore formation (Siniossoglou *et al.*, 1996). A role for these transcripts in sperm is currently unknown.

For gamete interactions, surface-localized proteins are of special interest because they might be involved in sperm–egg or sperm–central cell fusion. About 8% of the cDNAs with obvious open reading frames are known or predicted to encode secreted or plasma-membrane-localized proteins. This percentage is close to that expected in any cell (Jacobs *et al.*, 1997). Most sperm cDNAs in this class correspond to proteins of unknown function. For example, *Zmsp041* is predicted to encode a plasma-membrane-localized protein that is similar to *MtN3*, a *Medicago truncatula* gene that is induced upon nodulation (Gamas *et al.*, 1996). However, the annotation 'similar to MtN3' has been assigned to 17 different *Arabidopsis* genes (Ward, 2001; http://www.cbs.umn.edu/labs/jward), and the membrane topologies predicted for the different proteins are quite variable. Although much is known about animal fertilization, the proteins involved in the actual fusion event are still unknown. However, the proteins involved in vesicle and viral fusions are well characterized, and if these are used as a model, then proteins mediating cell fusion are likely to have a single transmembrane domain (reviewed in Bentz and Mittal, 2000). The *Zmsp943* transcript is most similar to an *Arabidopsis* transcript that encodes a protein predicted to have a single transmembrane domain. The *Arabidopsis* protein has homology to the attachment domain from reovirus (BLASTP).

Zmsp192 encodes a cysteine-rich secreted protein similar to that of the defensin-like ZmES1-4 proteins that are expressed in the egg, central cell, and synergids of maize (Cordts *et al.*, 2001). In animals, β-defensins expressed in the epididymis are thought to play an antimicrobial role during sperm maturation or function (Hammami-Hamza *et al.*, 2001; Li *et al.*, 2001). The antimicrobial protein Sob3 has been implicated in sperm binding to the zona pellucida (Hammami-Hamza *et al.*, 2001; Martin Ruiz *et al.*, 1998). As the *ZmES1-4* genes are downregulated upon fertilization, a defensive role for these proteins was proposed (Cordts *et al.*, 2001). A more tantalizing possibility is that gamete defensins might serve as ligands for gamete receptors because a defensin-like protein is a ligand for the S receptor kinase that mediates self-incompatibility in *Brassica* (Schopfer *et al.*, 1999).

Several other maize sperm cDNAs are interesting because they encode proteins similar to some known to play roles in animal sperm–egg interactions. *Zmsp051* encodes a Rab-like GTPase, and an acrosome-associated Rab3A GTPase in rat sperm has been proposed to play a role in membrane fusion (Iida *et al.*, 1999). Another, *Zmsp131,* encodes glucose-6-phosphate isomerase (GPI). Three forms of GPI have been reported in maize (Lal and Sachs, 1995; Salamini *et al.*, 1972). Because GPI catalyzes the second step in glycolysis, it is likely expressed in all the cells. In mammals, GPI has a different role outside the cell: it is secreted (in an unknown way) and serves as a nerve growth factor and cytokine (Jeffery *et al.*, 2000). An antibody against GPI causes agglutination of mouse sperm (Yakirevich and Naot, 2000). *Zmsp541* encodes a lipid kinase that phosphorylates inositol lipids. Intriguingly, phosphorylated inositol lipids may play a role in the sperm-induced calcium elevation seen in animal eggs (Jones and Nixon, 2000).

Numerous transcripts encode members of the ubiquitin and ubiquitin-like protein degradation pathways (del Pozo and Estelle, 2000; Muller *et al.*, 2001). Recently, it has become clear that ubiquitin-like proteins can also bind to target proteins and thereby regulate their function without targeting them for degradation (for reviews, del Pozo and Estelle, 2000; Pozo *et al.*, 1998). Ubiquitination appears to be involved in fertilization in at least one animal system; in

the ascidian *Halocynthia roretzi*, ubiquitination of the egg coat allows sperm to degrade and penetrate the outer layer (Sawada *et al.*, 2002; Sutovsky *et al.*, 2001).

Expression analysis

We reasoned that cDNAs that were sperm-specific might not be present in existing EST databases. We accordingly chose four cDNAs that were significantly similar to predicted *Arabidopsis* proteins that were not represented in the *Arabidopsis* EST database. *Zmsp271*, *Zmsp943*, and *Zmsp443* are most similar to hypothetical *Arabidopsis* proteins. *Zmsp041*, the MtN3-like cDNA, is most similar to a transcript isolated from microspore mother cells of lily, but its best match in *Arabidopsis* does not have an EST. To analyze their expression, semi-quantitative RT-PCR was performed on random-primed cDNA synthesized from total RNA from various tissues. As an internal control for RNA concentration, we used GPI (*Zmsp131*) because it was found in the sperm cDNA library and was expected to be expressed in all tissues. The results suggested that, of these, only *Zmsp943* appeared sperm cell-specific (Figure 3a). *Zmsp041* and *Zmsp443* might be expressed predominantly or only in the sperm within the mature pollen grain because the transcripts appeared to be much more abundant in the sperm RNA lane than in the total pollen lane, but both transcripts were also detectable in other tissues. The *Zmsp271* transcript (Figure 3a, lower right panel, middle band) was expressed weakly in all tested tissues but showed a slight enrichment in the sperm relative to whole pollen (Figure 3a, lower right panel, lanes 3 and 4). We suppose that the *Zmsp943* transcript was not visible in the total pollen sample (Figure 3a, lane 3) because it contains RNA from only 1600 sperm cells whereas the sperm sample (Figure 3a, lane 4) contains RNA from 55 600 FACS-purified sperm cells.

Because sperm were thought to be relatively transcriptionally inactive (reviewed in McCormick, 1993), we wanted to determine if the transcripts were present in sperm because they had been transcribed in the progenitor generative cell. To determine when transcripts in sperm were first transcribed, we performed RT-PCR on cDNAs made from unicellular microspores, from bicellular and tricellular pollen, and from FACS-purified sperm cells (Figure 3b). We examined the expression patterns of the cDNAs shown in Figure 3(a), as well as those of four other cDNAs corresponding to hypothetical or unknown proteins (*Zmsp444, Zmsp842, Zmsp034,* and *Zmsp721*). Unlike the previous experiment (Figure 3a) where 100 ng of RNA from each tissue was used, in the experiment shown in Figure 3(b), we used RNA from 800 unicellular microspores, from 800 bicellular pollen grains (i.e. 800 vegetative cells and 800 generative cells), or from 800 tricellular pollen grains (i.e. 800 vegetative cells and 1600 sperm cells). We knew the

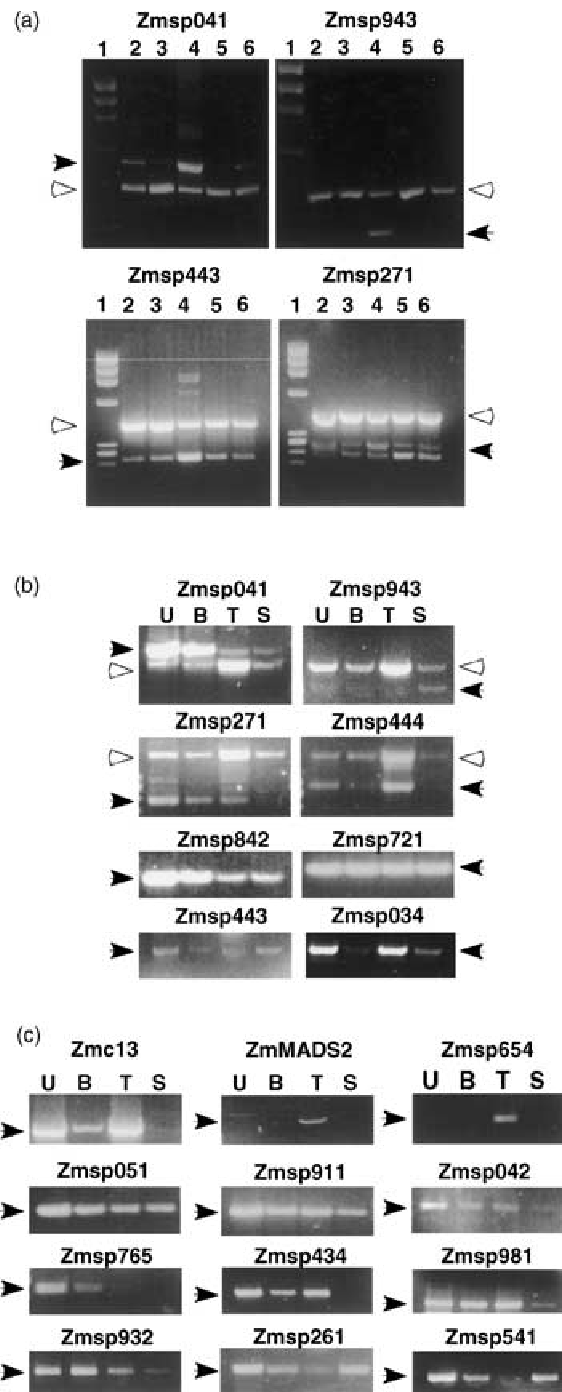

Figure 3. Expression profiles for sperm transcripts.
(a) Analysis of tissue expression. Lane 1, molecular weight marker; lane 2, silk; lane 3, pollen; lane 4, sperm; lane 5, seedling; lane 6, root.
(b) Analysis of the expression of transcripts encoding hypothetical or unknown proteins during pollen maturation. U, unicellular microspore; B, bicellular, T, tricellular pollen; or S, FACS-purified sperm cells.
(c) Analysis of the expression patterns of transcripts encoding proteins with predicted function or with known expression patterns. Symbols are the same as in (b). RT-PCR was performed on random-primed cDNA using transcript-specific primers. The arrows point to the transcripts of interest. The open arrowheads indicate the glucose-6-phosphate isomerase transcript, used as a control.

RNA yield from 1 million FACS-purified sperm cells and from a known quantity of mature pollen, assuming 100% recovery. Using these values, we deduced that at most 3% of the RNA isolated from a mature pollen grain came from the sperm cells inside it. In some cases, primers specific to *GPI* were also included in the PCR reaction; the GPI band is indicated with an open arrow. Surprisingly, all the transcripts were already present in the unicellular microspores, as well as in the FACS-purified sperm and mature pollen (Figure 3b).

We wondered if the presence of sperm transcripts at the unicellular microspore stage would be typical only for transcripts corresponding to hypothetical proteins. First, we looked at the expression of two transcripts with known expression profiles, *ZmMADS2* and *Zmc13*. *ZmMADS2* is present only in mature pollen (Heuer *et al.*, 2000). *Zmc13* is vegetative cell-specific in the mature pollen grain, but is also present at the unicellular microspore stage (Hamilton *et al.*, 1989). The *Zmc13* expression profile (Figure 3c) was identical to that reported (Hamilton *et al.*, 1989). As *ZmMADS2* was detected only in mature pollen (Figure 3c), we conclude that amplification of certain sperm transcripts from the unicellular microspore and bicellular pollen RNAs was not an artifact of the RT-PCR. We determined the expression patterns of several other sperm transcripts, some of which we expected to be expressed at relatively equal levels in all cells, such as *Zmsp981 (*vacuolar ATPase), *Zmsp932* (snrnp-F), and *Zmsp131* (GPI). We also examined several transcripts predicted to encode proteins of proposed function, but with unknown expression patterns: *Zmsp765 (*transcriptional activator), *Zmsp654 (*protein kinase), *Zmsp434 (*kinase activator), *Zmsp911 (*sec13-like), *Zmsp541 (*phosphoinositide kinase), *Zmsp051* (RAS-like GTPase), *Zmsp042 (*glucan synthase), and a transcript similar to a Zn finger-binding protein (*Zmsp261*). Figure 3(c) shows that these transcripts show varied expression levels, and that all except *Zmsp654* were present at the unicellular microspore stage.

The RT-PCR analysis shown in Figure 3(b,c) also provided information about the sperm specificity of the transcripts in the mature pollen grain. When amplification appeared to be nearly equivalent in the tricellular pollen and in the FACS-purified sperm, we concluded that the transcript was likely present only in the sperm and not in the vegetative cell. For example, for *Zmsp041* (upper band), *Zmsp943*, *Zmsp842, Zmsp443*, and *Zmsp721*, the amount of product in the sperm cell lane (S) is approximately equal to that seen in the mature pollen lane (T), suggesting that the transcript is located exclusively in the sperm cells (Figure 3b). Note that *Zmsp943* is detectable in the T lane (Figure 3b) when gel conditions are optimized (compare the different intensities of the GPI signal in Figure 3a, lane 3 and Figure 3b, lane T).

For others, such as *Zmsp271, Zmsp034*, and *Zmsp444*, the transcript appears to be enriched in the vegetative cell relative to the sperm cells, suggesting that there are transcripts in both the vegetative cell and the sperm cells (Figure 3b). Among the transcripts that encode proteins of known function, some transcripts, such as *GPI* (Figure 3b), sec13-like (*Zmsp911*), and vacuolar ATPase (*Zmsp981*), are clearly more abundant in the tricellular pollen fraction than in the sperm cell fraction, indicating that the transcript is present in both the vegetative cell and the sperm cells (Figure 3c). As we saw for two of the hypothetical protein transcripts (*Zmsp271* and *Zmsp444*) in Figure 3(b), *Zmsp654, Zmsp932, Zmsp434,* and *Zmsp042* did not amplify well in the sperm RNA fraction under the conditions used; we therefore suppose that they must be rare transcripts in sperm cells. Of the transcripts with proposed or known functions, only *Zmsp051* has a pattern similar to that of some of the hypothetical protein transcripts (*Zmsp842*, *Zmsp721*, *Zmsp443*, and *Zmsp041*), examined in Figure 3(b), in that it appears to be sperm-specific in the mature pollen grain. The *Zmsp541* and *Zmsp261* transcripts appear more abundant in FACS-purified sperm than in mature pollen; one explanation is that their transcription might have been induced by stress during sperm cell isolation (Figure 3c).

Whole-mount in situ *hybridization*

To determine if transcripts that appeared to be sperm-specific in mature pollen (Figure 3c) were indeed so, we used whole-mount *in situ* hybridization. Antisense probes were first tested on isolated sperm cells (data not shown), and probes that gave a strong signal were then used for whole-mount *in situ* hybridization. The results with *Zmsp041* are shown. Unlike hybridization signals for vegetative cell-expressed transcripts such as *ZmMADS2* (Heuer *et al.*, 2000), the sperm cell hybridization signal on whole mounts (Figure 4) is, at first, not obvious because sperm occupy only a small area of the total pollen grain volume and have a limited cytoplasm. The sperm nuclei and the vegetative cell nucleus are clearly visible in pollen after DAPI staining (Figure 4a,d,g). When pollen was hybridized with the antisense probe for *Zmsp041*, the deposited substrate was localized exclusively to the cytoplasm of the two sperm (Figure 4b, enlargement Figure 4c). In this pollen grain, the two sperm are oriented in different directions; the upper cell is perpendicular to the plane of focus and hybridization is detected only in the thin band of cytoplasm on either side of the nucleus (more clearly seen in Figure 4c). The lower sperm cell is parallel to the plane of focus, and hybridization can be detected throughout its cytoplasm (Figure 4b,c). The vegetative cytoplasm shows no hybridization above background. When pollen was hybridized with the sense probe, no signal was detected over the sperm (Figure 4e, enlargement Figure 4f). Control pollen (Figure 4h) was not hybridized with RNA but was,

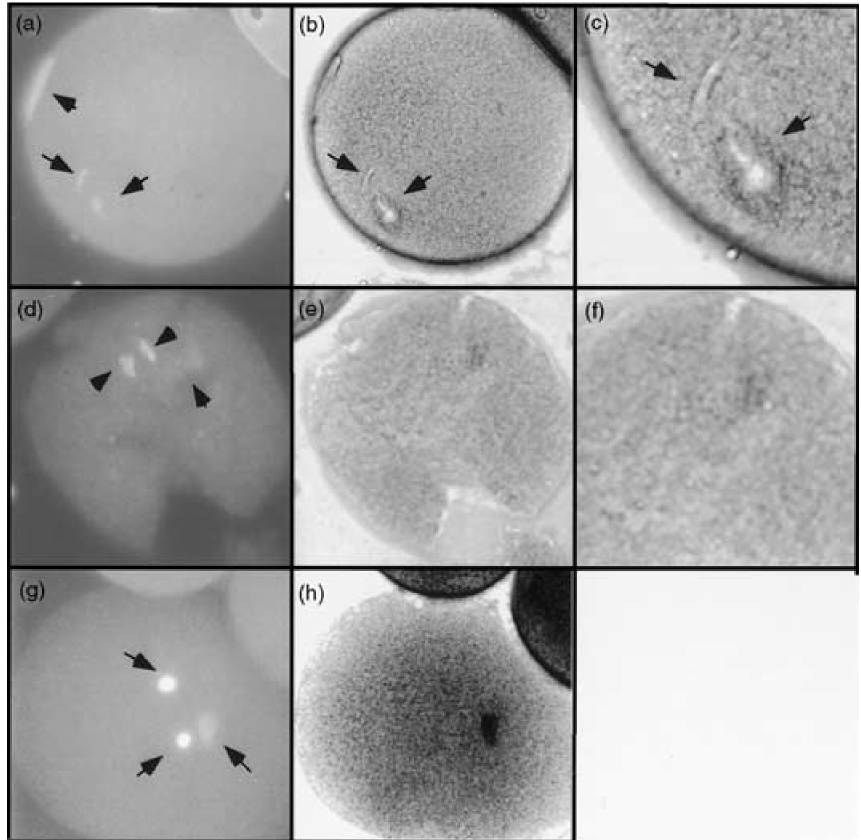

Figure 4. Whole-mount *in situ* hybridization with *Zmsp041*.
(a) A pollen grain stained with DAPI. The two sperm cell nuclei and the vegetative cell nucleus are visible. The sperm cells are the bright dots, and the vegetative cell nucleus is the diffusely staining body. The arrows indicate the nuclei.
(b) The same pollen grain as in (a) hybridized with *Zmsp041* antisense probe. The cytoplasm of the two sperm is visible following detection of the digoxigenin-labeled RNA. The arrows indicate the two sperm cells.
(c) An enlarged image of the two sperm from (b); arrows indicate the sperm cells.
(d) DAPI staining of a pollen grain; the arrows indicate the three nuclei.
(e) A pollen grain hybridized with the *Zmsp041* sense probe.
(f) An enlargement of the sperm cell containing region of the pollen grain from (e).
(g) DAPI staining of a control pollen grain. The two sperm nuclei are the bright dots, and the vegetative cell nucleus is the diffusely staining body. The arrows indicate the nuclei.
(h) The control pollen grain that was not hybridized with RNA. The same pollen grain as seen in (g).

otherwise, processed identically. When no probe or sense probe was used, there was some background over the vegetative nucleus, but neither the sperm cytoplasm nor the vegetative cytoplasm showed any signal (Figure 4e,f,h). RT-PCR showed that *Zmsp041* is also expressed in some other tissues (Figure 3a), but whole mounts indicate that within the pollen grain, *Zmsp041* is sperm cell-specific. Also, note that the *Zmsp041* RNA is detected in the sperm inside the intact pollen grains, and thus its transcript was not induced *de novo* during FACS purification.

Conclusions

The maize sperm library has greatly increased our knowledge about plant sperm gene expression. Before this work, only a few transcripts were known to be present in plant sperm. The *LGC1* transcript is present in the generative and sperm cells of lily (Xu *et al.*, 1999b), and the transcript for a homolog of the excision repair enzyme ERCC1 is in the generative cells (Xu *et al.*, 1999a). A transcript encoding a polygalacturonase was isolated from a cDNA library prepared from partially purified *Nicotiana tabacum* sperm (Xu *et al.*, 2002). In *Plumbago zeylanica*, a species with dimorphic sperm, a transcript for a polyubiquitin gene localized in the generative cell, but in only one of the two sperm cells of a pollen grain (Singh *et al.*, 2002). In addition, there are 60 ESTs from a cDNA library prepared from rice sperm (GenBank accession numbers BE225314–BE225323; BF475189–BF475237; F. Chen, Sichuan University); currently, less than half are similar to the maize sperm cDNAs. We believe that the limited overlap between the sequences from the rice and the maize sperm libraries is a result of the shallow sequencing of both libraries. Sequencing of the maize sperm cDNA library demonstrates that sperm cells have a diverse RNA complement, including a large number of transcripts that encode hypothetical proteins and proteins of unknown function. Thus, this cDNA library is a useful resource for genome annotation.

The *in situ* hybridization and RT-PCR results show that *Zmsp041* is sperm cell-specific in mature pollen. Several other transcripts (*Zmsp943*, *Zmsp842, Zmsp443*, and *Zmsp721*) have the same RT-PCR expression pattern as that of *Zmsp041* and may also be sperm cell-specific in mature pollen. Interestingly, the RT-PCR analysis of transcript expression during pollen development indicates that some of these transcripts are synthesized early in pollen development, as early as the unicellular microspore stage. What can explain the early presence of transcripts that later appear to be sperm-specific? Perhaps, such transcripts are extremely labile and are newly transcribed in the sperm. As we sequenced 17 isolates of a low-temperature/salt-induced cDNA, and as the FACS procedure is performed

at 4°C, this might support the notion of some transcriptional activity by the FACS-purified sperm. However, in our sequencing, we identified only a few cDNAs for proteins involved in transcription. For example, *Zmsp593* encodes a protein similar to RUSH-like (ring finger/helicase-like) transcription factors (for reviews, Chilton *et al.*, 2000; Devine *et al.*, 1999), *Zmsp765* encodes a protein identical to a pollen-expressed transcriptional activator (GenBank CAD22882.1), *Zmsp3144* encodes a basic helix-loop-helix (BHLH) protein and *Zmsp1389* encodes a MCM1, Agamous, Deficiens and Serum response factor (SRF) (MADS) box-like protein. It is possible that some of the cDNAs corresponding to unknown proteins encode as yet unidentified transcription factors. An alternative explanation is transcript partitioning, whereby certain transcripts destined for the sperm are synthesized at the unicellular microspore stage and are then progressively partitioned until they are localized specifically in the sperm cells. Partitioning of RNA transcripts is very well studied in *Drosophila melanogaster* and in *Xenopus* oocytes and embryos (for reviews, Johnstone and Lasko, 2001; Kloc *et al.*, 2001). For example, the bicoid transcript is synthesized in the nurse cells and transported into the oocyte where it is restricted to the anterior of the oocyte (Salles *et al.*, 1994). The bicoid transcript is not translated during oogenesis, but during embryogenesis, the transcript receives additions to its polyA tail at which point it becomes competent for translation (Salles *et al.*, 1994). A similar mechanism may be occurring during sperm development. The transcripts may be synthesized in progenitor cells, but may not be competent for translation until later.

Experimental procedures

Sperm isolation

Sperm were released from *Zea mays* (A188) pollen by a pH and osmotic shock, using a protocol modified from that of Dupuis *et al.* (1987), and were filtered through 50 and 20 μm filters. All steps were performed at 4°C unless otherwise noted. The filtrate was loaded onto a cushion of 10% Percoll™ (Amersham Pharmacia, Piscataway, NJ, USA) in 15% sucrose, 10 mM 3-(*N*-Morpholino) propanesulfonic acid (MOPS), pH 7.5, and centrifuged at 4500 ***g*** for 30 min. The sperm were collected from the Percoll™ per filtrate interface, washed with a filter-sterilized sperm isolation buffer (0.52 M mannitol, 10 mM MOPS, pH 7.5), and were concentrated by centrifugation at 300 ***g*** for 20 min. The sperm were labeled with 20 $\mu g\ ml^{-1}$ of Hoechst 33342 (Molecular Probes, Eugene, OR, USA) for 30 min in the dark at 25°C, washed with the cold sperm isolation buffer to remove residual Hoechst dye, concentrated by centrifugation at 300 ***g*** for 20 min, and were re-suspended to a density of 10^7 cells ml^{-1}. Cell sorting was performed on a FACSTAR+ flow cytometer (Becton Dickinson, Le Pont De Claix, France) equipped with a double laser. The 50 mW argon laser was tuned to 488 nm, and the INNOVA 300 ion multilined/UV laser was tuned to UV. Hoechst fluorescence was monitored with a DF 424/44 filter. Using the Exclusion sort mode, cells were sorted into 1.5 ml polypropylene tubes (low-binding microtubes, Polylabo, Fontenay Sous Bois, France) containing 50 μl of the sperm isolation buffer. To preserve cell viability, the sperm isolation buffer was also used as the dilution buffer, the sheath fluid and the recovery buffer. Data acquisition and analysis were performed with 'CellQuest' software (Becton Dickinson), with amplifiers set to linear mode for the scatter channels (FSC and SSC) and the fluorescence channel (FL4). The enriched fractions of isolated sperm were analyzed by collecting dotplots of forward scatter (FSC) versus DNA fluorescence emission (FL4), in order to determine the region containing cells. This region was used as a sorting gate (cell gate), cell purity was verified by microscopic observations, and then aliquots were collected (*c.* 200 000 sorted cells per tube and per 30 min). Sorted cells were immediately concentrated by centrifugation at 300 ***g*** for 20 min. The cell pellet was frozen in liquid nitrogen, and stored at −80°C until further analysis.

Library construction and sequencing

RNA was isolated from 1×10^6 FACS-purified sperm using a Qiagen RNAeasy kit (Qiagen, Valencia, CA, USA), and quantified with Ribogreen (Molecular Probes, Eugene, OR, USA). A Clontech SMART cDNA kit (Clontech, Palo Alto, CA, USA) was used for cDNA synthesis. The SMART kit employs a directional PCR-based protocol; the resulting PCR-amplified DNA has unique *Sfi*I sites. *Sfi*I-cut lambda arms were attached, and the ligations were packaged with Stratagene Gigapack III Gold packaging reagents (Stratagene, La Jolla, CA, USA). The titer of the unamplified library was 2.3×10^6 Pfu. Individual lambda clones were converted to plasmid using the Clontech Cre *loxP* system. Inserts were sequenced from the 5′ end using the Clontech 5′-end sequencing primer and ABI reagents. cDNA inserts of 500 bp or longer were preferentially sequenced. In order to eliminate prevalent transcripts from later sequencing runs (i.e. *Zmsp3325* to *Zmsp3516*), the unamplified library was plated and hybridized with ^{32}P-labeled transcripts corresponding to the most prevalent transcripts from the sequencing to date (i.e. components of retroposons, a low-temperature/salt stress-induced protein, histone H3, STM3, a nitrilase-associated protein, Hmgd1, Yptm3, and a chaperone). The plaques that did not hybridize were chosen for sequencing.

Sequence analysis

To determine sequence identity, each cDNA was searched against TBLASTX, BLASTP, and TBLASTN. A transcript was assigned a match to a protein of known identity if the probability of similarity score in BLAST was e-9 or smaller. All sequences have been deposited into GenBank: accession numbers BM659994–BM660223; BM675879; BM675880; CB278079–CB278388; and CB179149–CB179876. Sub-cellular location was predicted with PSORT (http://www.psort.nibb.ac.jp). Sequencing is ongoing; updated information can be viewed at the McCormick lab web site http://www.pgec.usda.gov/McCormick/mclab.html, or at NCBI (http://www.ncbi.nlm.nih.gov/) using the query '*Zea mays* sperm cell'.

RT-PCR analysis

RNA was isolated from *Z. mays* (B73) tissues essentially as described (Bedinger and Edgerton, 1990), except that tissues were ground with a mortar and pestle in the presence of liquid nitrogen. The RNA was treated with DNase I. The *Zmc13* transcript was amplified with 100 ng of cDNA prepared with random-primed total RNA from pollen or FACS-purified sperm, using *Zmc13*-specific primers. For analysis of multiple tissue expression patterns, several micrograms of RNA from each tissue was used to synthesize

random-primed cDNA. The cDNA was then divided, and 100 ng of random-primed cDNA from each tissue was amplified with *glucose-6-phosphate isomerase* (*GPI*) primers and transcript-specific primers or with transcript-specific primers alone. All PCR reactions were carried out for 30 cycles. Unicellular microspores and bicellular pollen were partially purified as described (Bedinger and Edgerton, 1990), and several micrograms of RNA were used to make random-primed cDNA. The sperm and bicellular cDNA reactions were supplemented with tRNA to a final concentration of 100 ng μl^{-1} of RNA to stabilize the cDNA. The sequences of the transcript-specific primers are available upon request.

Whole-pollen in situ *hybridization*

The pollen (A188) was incubated for 1 h in a solution of 12% sucrose, 0.03% $CaCl_2$, and 0.01% boric acid (Torres *et al.*, 1995). The pollen was fixed and processed as described (Heuer *et al.*, 2000), except that, before the proteinase K treatment, pollen was incubated with a combination of 0.5% cellulase, 0.5% macerozyme R10, 0.5% macerase in 100 mM Tris–HCl, pH 7.5, 50 mM EDTA, for 20 min at 25°C, and after proteinase K treatment, the pollen was incubated in dimethyl sulfoxide for 30 min at 42°C, to partially clear the cytoplasm (Kato, 1998). Digoxigenin-labeled probe was hybridized and detected as described (Heuer *et al.*, 2000). The pollen was mounted with 1 $\mu g\ ml^{-1}$ of 4′,6-diamidine-2′ phenylindole dihydrochloride (DAPI) and squashed with a glass coverslip. Images were captured using a Zeiss Axiophot microscope with a SPOT camera (Diagnostic Instruments, Inc., Sterling Heights, MI, USA).

Acknowledgements

This work was supported by United States Department of Agriculture CRIS 5335-21000-011-00D. M.E. was partially supported by an ARS-headquarters-funded postdoctoral fellowship. Continued support for EST sequencing is provided by NSF Plant Genome grant DBI-0211742. A.C and C.D. thank Annie Thomas at the flow cytometry facility for her help in setting up the sperm purification procedure and Sandrine Mourandian for collecting FACS-purified sperm. We thank Maria Mouchess and Stephanie Kirakopolos for their assistance in cDNA sequencing. M.M. was funded by a UC-LEADS summer internship. We thank Weihua Tang and Paul Herzmark for helpful comments on the manuscript, and S.M. thanks William Brown for encouragement and for useful discussions.

Supplementary Material

The following material is available from http://www.blackwellpublishing.com/products/journals/suppmat/TPJ/TPJ1761/TPJ1761sm.htm

Table S1 *Zea mays* sperm cDNAs are grouped by functional category, as determined by their best BLASTP scores. Only scores of 1e-8 or lower were considered significant. The GenBank number and the *e*-value for the best BLASTP alignment are listed. No Signif. Sim. means no significant similarity.

References

Andrews, J., Bouffard, G.G., Cheadle, C., Lu, J., Becker, K.G. and Oliver, B. (2000) Gene discovery using computational and microarray analysis of transcription in the *Drosophila melanogaster* testis. *Genome Res.* **10**, 2030–2043.

Antoine, A.F., Faure, J.E., Cordeiro, S., Dumas, C., Rougier, M. and Feijo, J.A. (2000) A calcium influx is triggered and propagates in the zygote as a wavefront during *in vitro* fertilization of flowering plants. *Proc. Natl. Acad. Sci. USA*, **97**, 10643–10648.

Bedinger, P.A. and Edgerton, M.D. (1990) Developmental staging of maize microspores reveals a transition in developing microspore proteins. *Plant Physiol.* **92**, 474–479.

Bentz, J. and Mittal, A. (2000) Deployment of membrane fusion protein domains during fusion. *Cell. Biol. Int.* **24**, 819–838.

Blomstedt, C.K., Knox, R.B. and Singh, M.B. (1996) Generative cells of *Lilium longiflorum* possess translatable mRNA and functional protein synthesis machinery. *Plant. Mol. Biol.* **31**, 1083–1086.

Breton, C., Faure, J.E. and Dumas, C. (1995) From *in vitro* fertilization to early embryogenesis in maize. *Protoplasma*, **187**, 3–12.

Chilton, B.S., Hewetson, A., Devine, J., Hendrix, E. and Mansharamani, M. (2000) Uteroglobin gene transcription: what's the RUSH? *Ann. NY Acad. Sci.* **923**, 166–180.

Cordts, S., Bantin, J., Wittich, P.E., Kranz, E., Lorz, H. and Dresselhaus, T. (2001) *ZmES* genes encode peptides with structural homology to defensins and are specifically expressed in the female gametophyte of maize. *Plant J.* **25**, 103–114.

Devine, J.H., Hewetson, A., Lee, V.H. and Chilton, B.S. (1999) After chromatin is switched-on can it be RUSHed? *Mol. Cell. Endocrinol.* **151**, 49–56.

Dupuis, I., Roeckel, P., Matthys-Rochon, E. and Dumas, C. (1987) Procedure to isolate viable sperm cells from corn (*Zea mays* L.) pollen grains. *Plant Physiol.* **85**, 876–878.

Evans, J.P. (2001) Fertilin beta and other ADAMs as integrin ligands: insights into cell adhesion and fertilization. *Bioessays*, **23**, 628–639.

Faure, J.E., Digonnet, C. and Dumas, C. (1994) An *in vitro* system for adhesion and fusion of maize gametes. *Science*, **263**, 1598–1600.

Franklin-Tong, V.E. (1999) Signaling in pollination. *Curr. Opin. Plant Biol.* **2**, 490–495.

Gamas, P., Niebel, F., Lescure, N. and Cullimore, J. (1996) Use of a subtractive hybridization approach to identify new *Medicago truncatula* genes induced during root nodule development. *Mol. Plant Microbe Interact.* **9**, 233–242.

Giordano, R., Magnano, A.R., Zaccagnini, G., Pittoggi, C., Moscufo, N., Lorenzini, R. and Spadafora, C. (2000) Reverse transcriptase activity in mature spermatozoa of mouse. *J. Cell Biol.* **148**, 1107–1113.

Grasser, K.D., Grimm, R. and Igloi, G.L. (1997) Purification and cDNA cloning of maize HMGd reveal a novel plant chromosomal HMG-box protein with sequence similarity to HMGa. *Gene*, **190**, 303–307.

Grossniklaus, U. and Schneitz, K. (1998) The molecular and genetic basis of ovule and megagametophyte development. *Semin. Cell Dev. Biol.* **9**, 227–238.

Hamilton, D.A., Bashe, D.M., Stinson, J.R. and Mascarenhas, J.P. (1989) Characterization of a pollen-specific genomic clone from maize. *Sex. Plant Reprod.* **2**, 208–212.

Hammami-Hamza, S., Doussau, M., Bernard, J., Rogier, E., Duquenne, C., Richard, Y., Lefevre, A. and Finaz, C. (2001) Cloning and sequencing of *SOB3*, a human gene coding for a sperm protein homologous to an antimicrobial protein and potentially involved in zona pellucida binding. *Mol. Hum. Reprod.* **7**, 625–632.

Hanson, D.D., Hamilton, D.A., Travis, J.L., Bashe, D.M. and Mascarenhas, J.P. (1989) Characterization of a pollen-specific cDNA clone from *Zea mays* and its expression. *Plant Cell*, **1**, 173–179.

Heuer, S., Lortz, H. and Dresselhaus, T. (2000) The MADS box gene ZmMADS2 is specifically expressed in maize pollen and during maize pollen tube growth. *Sex. Plant Reprod.* **13**, 21–27.

Higashiyama, T., Yabe, S., Sasaki, N., Nishimura, Y., Miyagishima, S., Kuroiwa, H. and Kuroiwa, T. (2001) Pollen tube attraction by the synergid cell. *Science*, **293**, 1480–1483.

Iida, H., Yoshinaga, Y., Tanaka, S., Toshimori, K. and Mori, T. (1999) Identification of Rab3A GTPase as an acrosome-associated small GTP-binding protein in rat sperm. *Dev. Biol.* **211**, 144–155.

Inaba, K., Padma, P., Satouh, Y., Shin, I.T., Kohara, Y., Satoh, N. and Satou, Y. (2002) EST analysis of gene expression in testis of the ascidian *Ciona intestinalis*. *Mol. Reprod. Dev.* **62**, 431–445.

Jacobs, K.A., Collins-Racie, L.A., Colbert, M. ***et al.*** (1997) A genetic selection for isolating cDNAs encoding secreted proteins. *Gene*, **198**, 289–296.

Jeffery, C.J., Bahnson, B.J., Chien, W., Ringe, D. and Petsko, G.A. (2000) Crystal structure of rabbit phosphoglucose isomerase, a glycolytic enzyme that moonlights as neuroleukin, autocrine motility factor, and differentiation mediator. *Biochemistry*, **39**, 955–964.

Johnstone, O. and Lasko, P. (2001) Translational regulation and RNA localization in *Drosophila* oocytes and embryos. *Annu. Rev. Genet.* **35**, 365–406.

Jones, K.T. and Nixon, V.L. (2000) Sperm-induced Ca(2+) oscillations in mouse oocytes and eggs can be mimicked by photolysis of caged inositol 1,4,5-trisphosphate: evidence to support a continuous low level production of inositol 1,4,5-trisphosphate during mammalian fertilization. *Dev. Biol.* **225**, 1–12.

Kato, A. (1998) Hematoxylin procedure for staining mature pollen grains in maize with dimethylsulfoxide as a clearing agent. *Biotechnol. Histochem.* **73**, 1–5.

Kloc, M., Bilinski, S., Chan, A.P., Allen, L.H., Zearfoss, N.R. and Etkin, L.D. (2001) RNA localization and germ cell determination in *Xenopus*. *Int. Rev. Cytol.* **203**, 63–91.

Kranz, E., von Wiegen, P., Quader, H. and Lorz, H. (1998) Endosperm development after fusion of isolated, single maize sperm and central cells *in vitro*. *Plant Cell*, **10**, 511–524.

Lal, S.K. and Sachs, M.M. (1995) Cloning and characterization of an anaerobically induced cDNA encoding glucose-6-phosphate isomerase from maize. *Plant Physiol.* **108**, 1295–1296.

Li, Z.H., Matthews, P.D., Burr, B. and Wurtzel, E.T. (1996) Cloning and characterization of a maize cDNA encoding phytoene desaturase, an enzyme of the carotenoid biosynthetic pathway. *Plant Mol. Biol.* **30**, 269–279.

Li, P., Chan, H.C., He, B., So, S.C., Chung, Y.W., Shang, Q., Zhang, Y.D. and Zhang, Y.L. (2001) An antimicrobial peptide gene found in the male reproductive system of rats. *Science*, **291**, 1783–1785.

Lord, E.M. and Russell, S.D. (2002) The mechanisms of pollination and fertilization in plants. *Annu. Rev. Cell. Dev. Biol.* **18**, 81–105.

MacIntosh, G.C., Wilkerson, C. and Green, P.J. (2001) Identification and analysis of *Arabidopsis* expressed sequence tags characteristic of non-coding RNAs. *Plant Physiol.* **127**, 765–776.

Mannen, H., Tseng, H.M., Cho, C.L. and Li, S.S. (1996) Cloning and expression of human homolog HSMT3 to yeast SMT3 suppressor of MIF2 mutations in a centromere protein gene. *Biochem. Biophys. Res. Commun.* **222**, 178–180.

Martin Ruiz, C., Duquenne, C., Treton, D., Lefevre, A. and Finaz, C. (1998) SOB3, a human sperm protein involved in zona pellucida binding: physiological and biochemical analysis, purification. *Mol. Reprod. Dev.* **49**, 286–297.

Matthys-Rochon, E., Mol, R., Heizmann, P. and Dumas, C. (1994) Isolation and microinjection of active sperm nuclei into egg cells and central cells of isolated maize embryo sacs. *Zygote*, **2**, 29–35.

McCormick, S. (1993) Male gametophyte development. *Plant Cell*, **5**, 1265–1275.

Miller, D. (2000) Analysis and significance of messenger RNA in human ejaculated spermatozoa. *Mol. Reprod. Dev.* **56**, 259–264.

Miller, D., Briggs, D., Snowden, H., Hamlington, J., Rollinson, S., Lilford, R. and Krawetz, S.A. (1999) A complex population of RNAs exists in human ejaculate spermatozoa: implications for understanding molecular aspects of spermiogenesis. *Gene*, **237**, 385–392.

Mogensen, H.L. (1992) The male germ unit: concept, composition, and significance. In *Sexual Reproduction in Flowering Plants (*Russell, S.D. and Dumas, C., eds). San Diego: Academic Press, Inc., pp. 129–145.

Muller, S., Hoege, C., Pyrowolakis, G. and Jentsch, S. (2001) SUMO, ubiquitin's mysterious cousin. *Nat. Rev. Mol. Cell. Biol.* **2**, 202–210.

Ostermeier, G.C., Dix, D.J., Miller, D., Khatri, P. and Krawetz, S.A. (2002) Spermatozoal RNA profiles of normal fertile men. *Lancet*, **360**, 772–777.

Pasyukova, E., Nuzhdin, S., Li, W. and Flavell, A.J. (1997) Germ line transposition of the copia retrotransposon in *Drosophila melanogaster* is restricted to males by tissue-specific control of copia RNA levels. *Mol. Gen. Genet.* **255**, 115–124.

del Pozo, J.C. and Estelle, M. (2000) F-box proteins and protein degradation: an emerging theme in cellular regulation. *Plant Mol. Biol.* **44**, 123–128.

Pozo, J.C., Timpte, C., Tan, S., Callis, J. and Estelle, M. (1998) The ubiquitin-related protein RUB1 and auxin response in *Arabidopsis*. *Science*, **280**, 1760–1763.

Qin, X. and Zeevaart, J.A. (1999) The 9-cis-epoxycarotenoid cleavage reaction is the key regulatory step of abscisic acid biosynthesis in water-stressed bean. *Proc. Natl. Acad. Sci. USA*, **96**, 15354–15361.

Roman, H. (1948) Directed fertilization in maize. *Proc. Natl. Acad. Sci. USA*, **34**, 46–52.

Rusche, M.L., Mogensen, H.L., Shi, L., Keim, P., Rougier, M., Chaboud, A. and Dumas, C. (1997) B chromosome behavior in maize pollen as determined by a molecular probe. *Genetics*, **147**, 1915–1921.

Russell, S.D. (1984) Ultrastructure of the sperm of Plumbago zeylanica. *Planta*, **162**, 385–391.

Russell, S.D. (1985) Preferential fertilization in *Plumbago*: ultrastructural evidence for gamete-level recognition in an angiosperm. *Proc. Natl. Acad. Sci. USA*, **82**, 6129–6132.

Salamini, F., Tsai, C.Y. and Neslon, O.E. (1972) Multiple forms of glucose-phosphate isomerase in maize. *Plant Physiol.* **50**, 256–261.

Salles, F.J., Lieberfarb, M.E., Wreden, C., Gergen, J.P. and Strickland, S. (1994) Coordinate initiation of Drosophila development by regulated polyadenylation of maternal messenger RNAs. *Science*, **266**, 1996–1999.

Sawada, H., Sakai, N., Abe, Y., Tanaka, E., Takahashi, Y., Fujino, J., Kodama, E., Takizawa, S. and Yokosawa, H. (2002) Extracellular ubiquitination and proteasome-mediated degradation of the ascidian sperm receptor. *Proc. Natl. Acad. Sci. USA*, **99**, 1223–1228.

Schopfer, C.R., Nasrallah, M.E. and Nasrallah, J.B. (1999) The male determinant of self-incompatibility in *Brassica*. *Science*, **286**, 1697–1700.

Singh, M.B., Xu, H., Bhalla, P.L., Zhang, Z., Swoboda, I. and Russell, S.D. (2002) Developmental expression of polyubiquitin genes and distribution of ubiquitinated proteins in generative and sperm cells. *Sex. Plant Reprod.* **14**, 325–329.

Siniossoglou, S., Wimmer, C., Rieger, M., Doye, V., Tekotte, H., Weise, C., Emig, S., Segref, A. and Hurt, E.C. (1996) A novel complex of nucleoporins, which includes Sec13p and a Sec13p homolog, is essential for normal nuclear pores. *Cell*, **84**, 265–275.

Southworth, D. (1992) *Freeze Fracture of Male Reproductive Cells*. San Diego: Academic Press, Inc.

Southworth, D., Platt-Aloia, K.A. and Thomson, W.W. (1988) Freeze fracture of sperm and vegetative cells in *Zea mays* pollen. *J. Ultrastruct. Mol. Struc. Res.* **101**, 165–172.

Sutovsky, P., Moreno, R., Ramalho-Santos, J., Dominko, T., Thompson, W.E. and Schatten, G. (2001) A putative, ubiquitin-dependent mechanism for the recognition and elimination of defective spermatozoa in the mammalian epididymis. *J. Cell. Sci.* **114**, 1665–1675.

Taylor, L.P. and Hepler, P.K. (1997) Pollen germination and tube growth. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **48**, 461–491.

Thomas, J.O. and Travers, A.A. (2001) HMG1 and 2, and related 'architectural' DNA-binding proteins. *Trends Biochem. Sci.* **26**, 167–174.

Torres, M., Rigau, J., Puigdomenech, P. and Stiefel, V. (1995) Specific distribution of mRNAs in maize growing pollen tubes observed by whole-mount *in situ* hybridization with non-radioactive probes. *Plant J.* **8**, 317–321.

Turcich, M.P. and Mascarenhas, J.P. (1994) PREM-1 a putative maize retroelement has LTR (long terminal repeat) sequences that are preferentially transcribed in pollen. *Sex. Plant Reprod.* **7**, 2–11.

Turcich, M.P., Bokhari-Riza, A., Hamilton, D.A., He, C., Messier, W., Stewart, C. and Mascarenhas, J.P. (1996) PREM-2, a copia-type retroelement in maize is expressed preferentially in early microspores. *Sex. Plant Reprod.* **9**, 65–74.

Ueda, K., Kinoshita, Y., Xu, Z.J., Ide, N., Ono, M., Akahori, Y., Tanaka, I. and Inoue, M. (2000) Unusual core histones specifically expressed in male gametic cells of *Lilium longiflorum*. *Chromosoma*, **108**, 491–500.

Wagner, V.T. and Dumas, C. (1989) Morphometric analysis of isolated *Zea mays* sperm. *J. Cell Sci.* **93**, 179–184.

Ward, J.M. (2001) Identification of novel families of membrane proteins from the model plant *Arabidopsis thaliana*. *Bioinformatics*, **17**, 560–563.

Wessler, S.R. (2001) Plant transposable elements: a hard act to follow. *Plant Physiol.* **125**, 149–151.

Wong, G.K., Wang, J., Tao, L., Tan, J., Zhang, J., Passey, D.A. and Yu, J. (2002) Compositional gradients in Gramineae genes. *Genome Res.* **12**, 851–856.

Xu, H., Swoboda, I., Bhalla, P.L. and Singh, M.B. (1999a) Male gametic cell-specific expression of H2A and H3 histone genes. *Plant Mol. Biol.* **39**, 607–614.

Xu, H., Swoboda, I., Bhalla, P.L. and Singh, M.B. (1999b) Male gametic cell-specific gene expression in flowering plants. *Proc. Natl. Acad. Sci. USA*, **96**, 2554–2558.

Xu, H., Weterings, K., Vriezen, W., Feron, R., Xue, Y., Derksen, J. and Mariani, C. (2002) Isolation and characterization of male-germ-cell transcripts in *Nicotiana tabacum*. *Sex. Plant Reprod.* **14**, 339–346.

Yakirevich, E. and Naot, Y. (2000) Cloning of a glucose phosphate isomerase/neuroleukin-like sperm antigen involved in sperm agglutination. *Biol. Reprod.* **62**, 1016–1023.

Zhang, G., Gifford, D.J. and Cass, D.D. (1993) RNA and protein synthesis in sperm cells isolated from *Zea mays* L. *Pollen. Sex. Plant Reprod.* **6**, 239–243.

All sequences have been deposited into GenBank: accession numbers BM659994–BM660223; BM675879; BM675880; CB278079–CB278388; and CB179149–CB179876.