

# Molecular repertoire of flowering plant male germ cells

Mohan B. Singh · Prem L. Bhalla · Scott D. Russell

Received: 8 December 2007 / Accepted: 1 January 2008 / Published online: 29 January 2008  
© Springer-Verlag 2008

**Abstract** Sperm cells—the male gametes of flowering plants—constitute the male founding lineage of angiosperms, possessing the unique capacity to fuse with the egg and central cells during double fertilization. Although it is well established that these cellular fusions are involved with initiating the development of the seedling-forming zygote and the endosperm that nourishes it, considerable information will be needed to characterize the full male molecular repertoire, which includes expressed genes of the male lineage, encoded proteins, and regulatory elements controlling male germ line identity, as well as male molecules that may mediate interactions with the female partner that initiate fertilization and development. Progress is being made using increasingly sensitive molecular methods to uncover important genes. With the pace of this discovery rapidly increasing, the likely outcome is that key molecules will be discovered within the next several years that control the founding cells of the embryo and endosperm and are involved in directing early development. Further insights into the genes and gene pathways that regulate male germ line differentiation will advance not only our fundamental understanding of these reproductive

cells, but also the nature of cell–cell recognition, membrane fusion, double fertilization, zygote activation, early plant development and may aid our understanding of factors that have contributed to the overwhelming evolutionary success of flowering plants.

**Keywords** Male germ unit · Sperm · Promoter · Expression profile · Double fertilization

## Introduction

Flowering plant sperm cells, which are the cryptic male gametes of angiosperms, are small cells delivered by pollen tubes to the female gametophyte (embryo sac) with the capacity to fuse with the egg and central cells, effecting double fertilization and initiating subsequent zygote and endosperm development. Despite their key role in sexual reproduction, the male germ cells of angiosperms are deceptively simple in appearance. Each sperm cell contains a nuclear complement of genomic DNA, as well as a limited quantity of cytoplasm, which frequently contains heritable organellar DNA in mitochondria (almost always present) and, in some species, plastids (Russell 1991, 1992). Angiosperm sperm cells are apparently moved by the pollen tube but are not independently motile. The formation of these sperm cells requires two steps that may trigger initiation of critical developmental programs. First, the division of the microspore leads to the activation of divergent cellular programs in the generative and vegetative cell, which seem to be associated with asymmetry during cytokinesis, as disruption of division asymmetry leads to failure of germ line initiation (Eady et al. 1995). Second, mitotic division of the progenitor generative cell subsequently forms the two sperm cells, which seems to

---

Communicated by Thomas Dresselhaus.

---

M. B. Singh (✉) · P. L. Bhalla  
Plant Molecular Biology and Biotechnology Laboratory,  
Australian Research Council Centre of Excellence for  
Integrative Legume Research, Faculty of Land and Food  
Resources, The University of Melbourne, Parkville,  
Victoria 3010, Australia  
e-mail: mohan@unimelb.edu.au

S. D. Russell (✉)  
Department of Botany and Microbiology,  
University of Oklahoma, Norman, OK 73019, USA  
e-mail: srussell@ou.edu

invoke a developmentally regulated cellular transition that endows the sperm cells with an ability to fuse with the egg and/or central cells. The latter phase of sperm maturation and achievement of receptivity is believed to be involved with the cell cycle and its determinants (discussed later).

An apparently unique cellular event in seed plants, compared to other eukaryotes is the migration of the generative cell in pollen from its superficial position bordering the intine, into the interior of the vegetative cell, forming a unique cell-within-a-cell structure (Russell et al. 1996). Mitotic division of the generative cell forms two sperm cells within the confines of the vegetative cell cytoplasm. As a consequence, the generative and sperm cells are entirely dependent on the vegetative cell for their nutritional requirements. However, it is unclear that this nutritional dependence reduces the metabolic requirements of male germ cells. Undoubtedly the diminutive size of male germ cells reduces the need for highly expressed housekeeping genes, but nonetheless there may be a significant complement of such genes expressed. The expressional profiles of pollen cells are described as being uniquely divergent from the profiles of other somatic cells (Becker and Feijó 2007; Nobuta et al. 2007). Inside the pollen cells, the male germ cells are likely to represent a further divergence in expression related to the unique characteristics of the male germ lineage.

Functions crucial to the fate of sperm cells are likely to be expressed uniquely in the male germ lineage, including cell recognition, fusion mechanisms, and post fusion triggers for the delivery of the hereditary material. The sperm cells themselves are transmitted to the female gametophyte via a pollen tube, which consists of an elongating protrusion of the pollen grain formed by tip growth in continuity with the largely cellulosic inner pollen wall. Until recently, it was thought that once a pollen tube crosses pre-zygotic fertilization barriers, such as self-incompatibility, and arrives at the receptive embryo sac, double fertilization will inevitably ensue. However, late acting barriers are now known, including female-expressed mutations that alter the ability of pollen tubes to complete their function within the embryo sac, and thus prevent tube attraction and discharge (Huck et al. 2003; Rotman et al. 2003). If male germ transmission is successful, one sperm cell fuses with the egg to produce the embryo, and the other fuses with the central cell to produce the nutritive endosperm. Although molecular details are yet incomplete, *in vitro* experiments suggest that sperm cells, although naturally fusogenic, discriminate and fuse only with egg and central cells—which constitute the female gametic cells—rather than with synergids, antipodals or somatic cells (Faure et al. 1994; Wang et al. 2006; Kranz and Scholten 2008; this issue). The idea that cell surface molecules in male gametes may mediate gametic recognition and fusion processes in

flowering plants seems a logical conclusion based on studies of preferential gamete fusion (Russell 1985) and is supported by analogy with animal systems.

In animals, the interaction of gametes is mediated by cell surface fertilization proteins (Vacquier 1998; Rubinstein et al. 2006). Sperms of both invertebrate and vertebrate animals carry surface proteins needed for gamete recognition, adhesion and fusion. In sea urchins fertilization protein, bindin mediates species-specific sperm adhesion with female gametes (Ulrich et al. 1999; Zigler and Lessios 2003). A vertebrate counterpart protein, fertilin, has been implicated as a key mediator of plasma membrane binding and sperm-egg fusion (Bigler et al. 2000). Compelling evidence indicates that the mouse sperm protein Izumo plays a critical role in fertilization (Inoue et al. 2005; Rubinstein et al. 2006). Izumo knockout mice produced normal appearing sperm cells in both morphology and motility, but were completely infertile (Inoue et al. 2005). The nature of interactions of male-presented molecules with the female partner is discussed by Márton and Dresselhaus (2008; this issue).

The identity of specific flowering plant surface proteins required for fertilization will eventually be determined, but currently the exact players are yet to be characterized. When and where such proteins are expressed are among the questions concerning the molecular basis of differentiation, gamete interactions, and germ cell specifications in plants.

Among factors contributing to difficulties in characterizing molecular determinants of plant sexual reproduction are the relative inaccessibility of sperm and egg cells due to their encasement in gametophyte tissues; however, cell separation technologies for generative and sperm cells have been developed for numerous plant species (Russell 1991; Uchiumi et al. 2006) leading the way to molecular approaches to studying these cells (see also Kranz and Scholten 2008; this issue). Metabolite labeling experiments using isolated male germ cells have confirmed that these cells possess transcriptional and translational capabilities independent of the outer pollen vegetative cell (Zhang et al. 1993; Blomstedt et al. 1996). Despite completed genomic sequences in such model plants as *Arabidopsis* and rice, the identification of the portion of the genome that is expressed in the male germ lineage of flowering plants has remained a challenge.

Identifying the genome- and proteome-wide expression occurring in sperm and egg cells will be essential for understanding the molecular basis of flowering plant sexual reproduction and for uncovering the wealth of genes most likely to encode proteins that interact during productive sperm–egg interactions related to fertilization. To date, advances in genomic technologies have allowed capture of some important information on the transcriptional repertoire of the male germ line.

## Transcriptional profile of male germ cells

High throughput technologies addressing gene expression at the transcript level have enabled selective profiling of generative cells (in lily) and sperm cells of different plant species (maize, *Plumbago*, and rice). These studies employed a variety of molecular approaches, including subtractive cDNA library analysis (Gou and Russell, unpublished data), expressed sequence tag (EST) identification (Engel et al. 2003; Okada et al. 2006a), and cDNA microarrays (Okada et al. 2007). Data originating from the above studies have provided snapshots of the transcriptional repertoire of these male germ line cells that reflect the existence of three classes of genes expressed in these cells: (a) housekeeping genes expressed constitutively in male germ line, pollen vegetative cell and other plant cells; (b) genes found in common with other cells and cell types, but up-regulated in the male germ cells; and (c) germ cell specific genes. Of the latter category, some germ cell specific genes share functions with somatic counterparts, whereas others are unique in sequence and function exclusively in germ line cells. Most of our information on the transcriptional repertoire of male germ cells comes from three plants: lily, *Plumbago* and maize.

### Lily generative cells

The generative cells of lily have been transcriptionally analyzed to reveal their molecular signatures. EST data

revealed that the gene-expression profile of generative cells is unlike that of other plant cells including the pollen vegetative cell, with abundant ESTs reflecting the specialized features of the male germ line cells. Interestingly, the most abundant class of genes includes those genes whose functions have yet to be annotated (Okada et al. 2007). Among the first transcripts to be reported as specific to lily generative cells were LGC1, a gene coding for a membrane protein (Xu et al. 1999b) and male germ line histone variants *gcH2A*, and *gcH3* (Xu et al. 1999a; Ueda et al. 2000). Another strongly upregulated gene was a highly conserved DNA repair gene with homology to human ERCC1 (Xu et al. 1998). Gene transcripts most abundantly transcribed among lily generative cell-expressed ESTs are presented in Table 1. Also abundant in lily generative cells are transcripts encoding all major elements of the ubiquitin pathway, such as polyubiquitin, proteasome subunit, ubiquitin-conjugating enzyme, Skp1 and Ring box protein suggesting that this pathway is highly upregulated and active (Singh et al. 2002; Okada et al. 2006a).

Transcripts for *LGCI* and other lily male gamete-specific genes are not detectable at the uninucleated microspore stage, indicating that the male germ lineage differentiates late in development, after the completion of microsporogenesis (Xu et al. 1999a, b; Ueda et al. 2000, 2005). Analysis of transgenic tobacco plants carrying an *LGCI* promoter::GUS construct (further supported by similar aflatoxin constructs) confirmed male germ line-

**Table 1** Putative function of most abundant ESTs present in lily generative cell cDNA library (from Okada et al. 2006b, by permission of Oxford University Press)

MIPS <sup>a</sup>	No. of ESTs	Similarity <sup>b</sup>	<i>E</i> value <sup>c</sup>	Functional role categories <sup>d</sup>
At4g05320	33	Polyubiquitin <sup>e</sup>	1E-116	Protein destination
At5g42190	12	Skp1	1E-61	Cell growth, cell division and DNA synthesis
No hits	11	LGC1 ( <i>Lilium longiflorum</i> )	1E-64	
No hits	10	No similarity		
At4g11240	8	Serine/threonine protein phosphatase PP1	1E-106	Cellular communication/signal transduction
No hits	7	No similarity		
At5g07610	6	Hypothetical protein ( <i>Oryza sativa</i> )	1E-19	Unclassified
No hits	5	Hypothetical protein ( <i>Oryza sativa</i> )	5E-20	
At5g27670	5	Histone gH2A.1 ( <i>Lilium longiflorum</i> )	6E-65	Cellular organization
At1g09200	5	Histone gH3 ( <i>Lilium longiflorum</i> )	1E-61	Cellular organization
At1g67580	5	Putative protein kinase ( <i>Arabidopsis thaliana</i> )	3E-27	Cellular communication/signal transduction
At3g42790	5	Nucleic acid binding protein-like ( <i>Arabidopsis thaliana</i> )	2E-38	Transcription
No hits	5	No similarity		

<sup>a</sup> The most similar *Arabidopsis* genes are listed as their MIPS codes (<1.0E-10)

<sup>b</sup> Similarity to known gene

<sup>c</sup> The best E-value of BLASTX search against GenBank nonredundant protein data base

<sup>d</sup> Annotated putative function according to *Arabidopsis* database

<sup>e</sup> Four ubiquitin genes identified as nonredundant gene according to DNA sequence were summarized in one column

specific activation of the *LGC1* promoter (Singh et al. 2003). Microarray experiments using generative cell derived ESTs revealed that only 17% of the ESTs showed detectable hybridization with sporophytic sources of mRNA. Thus, with 83% of lily generative cell ESTs showing apparent male germ cell specificity, it appears that no other plant cell type analyzed to date produces such a high ratio of cell-specific transcripts (Okada et al. 2007). Similarly, among animal systems only spermatogenic cells and testis displayed such a high proportion of cell specific transcripts (Andrews et al. 2000; Schultz et al. 2003).

### Maize sperm cells

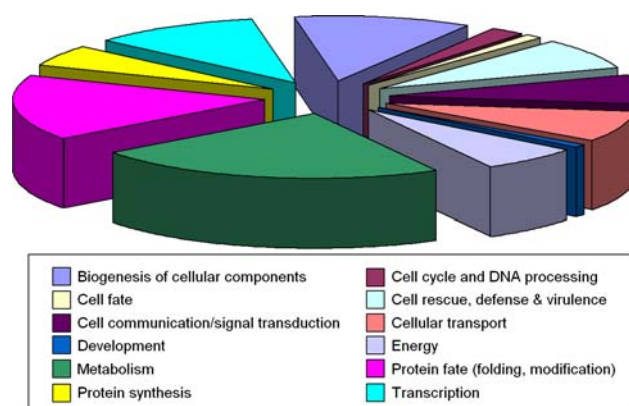
In maize, a cDNA library was constructed using sperm cells isolated by fluorescence-activated cell sorting (FACS), followed by high throughput sequencing of expressed sequence tags (Engel et al. 2003). Sequences of about 5,093 ESTs representing maize sperm-expressed transcripts are available in GenBank. A characteristic feature of the grass genomes has been a relatively high representation of transposable elements, for which the sperm lineage is no exception. Annotated retrotransposons, represent 9.46% of the maize sperm ESTs uploaded to GenBank, but relatively few annotated transposons (0.06%). In contrast, female gametophytes showed less than one-fifth the number of retrotransposons (1.69%) and more annotated transposons (1.44%) (Yang et al. 2006), reflecting potentially differential expression of such elements in the male germ lineage. Sperm-specific EST sequences from maize have been successfully used to identify sperm-specific gene candidates in *Arabidopsis* that have been expressed specifically in sperm cells (Engel et al. 2005). These products, frequently annotated as hypothetical proteins, have proven to be highly conserved.

### Plumbago sperm cells

*Plumbago zeylanica* is a flowering plant that produces two dimorphic sperm cells in the pollen and is known also to participate in preferential fertilization (Russell 1985). The sperm cell that typically fertilizes the central cell, and thus produces the endosperm is classically termed the  $S_{vn}$  (associated with vegetative nucleus); this sperm cell is recognized because it contains more mitochondria and rare plastids. The sperm cell that typically fertilizes the egg cell, forming the zygote, embryo and seedling, is known as the  $S_{ua}$ ; this second sperm cell contains abundant plastids and fewer mitochondria. To elucidate possible expressional differences in the sperm cells, a method for isolating and collecting two sperm cell types based on micropipetting

individual sperm cell type has been developed (Zhang et al. 1998). Individual cDNA libraries have been prepared from each sperm type, and sequencing of clones from each has highlighted some differences in the repertoire of genes expressed in each sperm type. Representative EST sequences of these two sperm cell types have been submitted to GenBank, including 893 sequences isolated from  $S_{ua}$  and 629 sequences of  $S_{vn}$ . The  $S_{ua}$  yielded 426 distinct sequences upon clustering and the  $S_{vn}$  proved 419 distinct sequences. Of these, 13.3% represent products with unknown function and 60.8% represented sequences with no known homology. Only 25.9% could be classified into functional categories. Of the categorized genes, the largest groups were metabolism, protein modification, transcription and biosynthesis (Fig. 1).

Of those sequences with no known homology in the GenBank protein database, the  $S_{vn}$  had slightly more unique sequences, with 62.8% having no hits, versus 58.9% in the  $S_{ua}$ . The high percentage of no hits and unclassified sequences indicates that current databases provide inadequate coverage of genes involved in angiosperm male gamete biology. Similar to lily generative cells, there was a conspicuous upregulation of ubiquitin, as well as metabolic, transcriptional, and biogenetic activity. Interestingly, ubiquitin was shown to be especially upregulated in the  $S_{vn}$ . That such a gene could be so highly upregulated in only one of the two sperm cells illustrates the likelihood of independent regulatory elements controlling expression in each of the two sperm cells of *Plumbago* (Singh et al. 2002), which has been confirmed by experimental analysis (X. P. Gou, X. P. Wei, T. Yuan, S. D. Russell, unpublished data).



**Fig. 1** Functional categorization of *Plumbago* sperm ESTs. Characteristic of male germ line cells, many putative protein products either have no significant homology or are unclassified proteins. Among functional categories, metabolism, protein modification (includes ubiquitin pathways), transcription and biogenesis are most highly upregulated. (Courtesy of XP Gou, XP Wei and S Russell)

### A conserved male germ line cell repertoire?

If the defining features of the male germ line cells, such as their unique ability to effect fertilization are the result of intrinsic genetic programming then it is expected that male germ line cells of different plants share a conserved molecular signature. Comparisons of available gene expression profiles of generative and sperm cells provide evidence of a conserved genetic program across monocot and dicot species. Shared genes comprise both known genes that are functionally annotated, as well as genes that lack currently known function. Table 2 shows selected genes that have male germ line expression in multiple plants analyzed so far.

Among genes having a known function, those encoding ubiquitin pathway-related proteins such as polyubiquitin, proteasome subunit, ubiquitin-conjugating enzyme, Skp1, and Ring box protein, are highly upregulated in male germ line cells, as well as being highly conserved molecules. Such male germ cell upregulation of the ubiquitin pathway may be universal, as the ubiquitin system also appears to play an essential role in male gametogenesis in mice and humans (Baarends et al. 1999a, b). Since many features of fertilization systems in animals are shared between plants and animals, it may be reasonable to expect that the ubiquitin proteolysis system plays a similar critical role in male gametogenesis of higher plants and in sperm-egg

fusion during fertilization. High expression of protein kinases and phosphatases in plant male germ cells, including uniquely gamete-expressed representatives of these enzymes, is also a feature shared with male gametogenesis in animal systems. A significant number of signature protein kinases and protein phosphatases were also identified as sperm-enriched genes in *Caenorhabditis elegans* (Reinke et al. 2000). Among plant ESTs, male germ specific proteins of these classes are expected to play an important role, as well.

Sperm cells also have unique expression that may relate to nuclear expression, such as histone and non-histone proteins. Cell specific expression of histone variants, particularly histone H3, is likely a common characteristic of the male germ line in flowering plants (Xu et al. 1999a; Okada et al. 2005b). At least five cell-specific variants of histone H3 are expressed in lily male germ cells (Ueda et al. 2000; Okada et al. 2006b) and among available lily generative cell ESTs, there were more than twice as many ESTs of histone *H3* genes compared to other histone components such as *H2A*, *H2B* and *H4* (Okada et al. 2005a, b). Immunocytochemical studies have further demonstrated that histone variants gH3 and gcH3 are incorporated in male germ line chromatin in a replication independent manner.

In maize sperm cells, among 1,100 ESTs reported, 20 ESTs of histone H3 products were found, in comparison

**Table 2** Selected list of plant male germ line marker genes

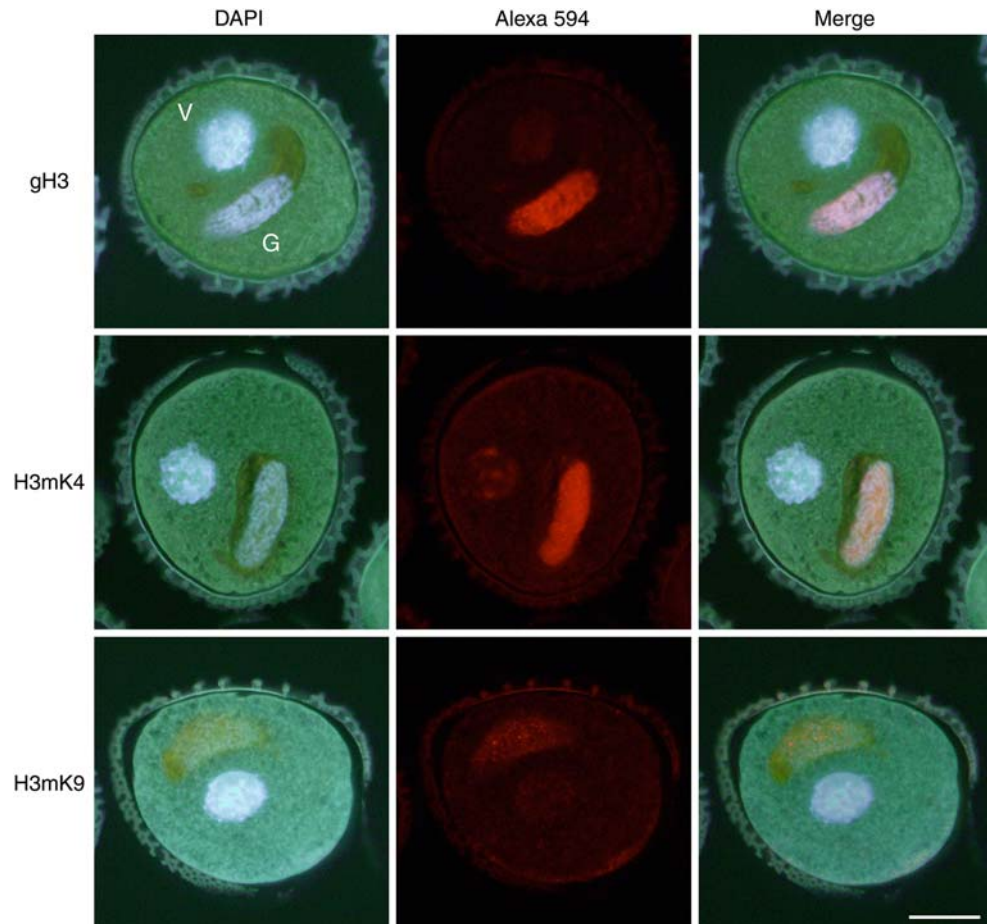
Gene	Expression	Functional annotation	Localization	Plant studied	Reference
LGC 1	GC and SC	128 aa Glycoprotein with GPI anchor	Cell membrane	<i>Lilium longiflorum</i>	Xu et al. (1999b)
GCS1	GC and SC	722 aa with transmembrane domain, essential for fertilization	Cell membrane	<i>Lilium longiflorum</i>	Mori et al. (2006)
HAP2	SC	705 aa HAP2 homologue, essential for fertilization	Cell membrane	<i>Arabidopsis thaliana</i>	von Besser et al. (2006)
DUO1	GC	298 aa with MYB domain, required for sperm cell formation	Nucleus	<i>Arabidopsis thaliana</i>	Rotman et al. (2005)
AtGEX2	GC and SC	Protein with transmembrane domains	Cell membrane	<i>Arabidopsis thaliana</i> ( <i>Zea mays</i> )	Engel et al. (2005)
gH3	GC and SC	149 aa Germ line Histone H3 variant	Nucleus	<i>Lilium longiflorum</i>	Ueda and Tanaka (1995)
gH2A	GC and SC	110 aa Germ line histone H2A variant	Nucleus	<i>Lilium longiflorum</i>	Xu et al. (1999a)
gcH3	GC and SC	111 aa germ line histone H3 variant	Nucleus	<i>Lilium longiflorum</i>	Xu et al. (1999a)
AtMGH3	GC and SC	137 aa germ line histone H3 variant	Nucleus	<i>Arabidopsis</i>	Okada et al. (2005b)

GC generative cell; SC sperm cell

From Singh and Bhalla (2007), by permission of Wiley Periodicals Inc.

Parentheses indicate source plant for original sequence

**Fig. 2** Immunocytochemical localization of variant histone gH3, showing dimethylation of H3K4 and H3K9 in lily mature pollen. Pollen sections reacted with anti-gH3 peptide, antimethyl H3K4 (H3K4Me2), anti-dimethyl H3K9 (H3K9Me2) antibody were followed by a secondary antibody labeling with fluorescent dye (Alexa594, center lane). DAPI-labeling was used to observe pollen vegetative cell nuclei and GC nuclei (left lane), with merged image in center using epifluorescence microscopy. *G* GC nucleus; *V* vegetative cell nucleus. Bar = 50  $\mu$ m (Courtesy of T Okada, after Okada et al. 2006b)



with less than five for other histone genes (Engel et al. 2003). At least three H3-variant genes are expressed in *Arabidopsis* generative cells, including the male gamete-specific variant gene *AtMGH3* as well as variant H3.3 (Okada et al. 2005a). The compaction of chromatin accompanying animal spermatogenesis involves replacement of approximately 85% of the histones with protamines (Churikov et al. 2004; Kimmins and Sassone-Corsi 2005; Ooi and Henikoff 2007). Variant forms of histones H3 and H4 also participate in these chromatin-remodeling processes in animals (Sassone-Corsi 2002). The role of plant male germ line specific histones is not as clear as in animals, since male gamete specific loss-of-function mutants of *Arabidopsis* do not show any obvious phenotype. It is, however, conceivable that male germ line histone variants play an important role in epigenetic regulation of gene expression via modifications of their lysine residues (Grewal and Moazed 2003). The lily variant gH3 histone, for instance, can only be localized in generative nucleus chromatin (Fig. 2). In addition to variant histones, generative cells also show evidence of chromatin modification, as indicated by the localization of a strong methylation signal at lysine residue position 4 (H3K4) and

position 9 (H3K9) of generative nucleus H3 histone (Okada et al. 2006b; Fig. 2). These results suggest that the genome of male gametic cells shows epigenetic modification distinct from the vegetative cell genome (Sano and Tanaka 2007), pointing to possible involvement in differential transcriptional regulation.

*LGCI* (Xu et al. 1999b) and *GCSI* (Mori et al. 2006) from lily generative cells are among the conserved male germ line specific genes whose functions could not be annotated from their sequences. *LGCI* (*Lily Generative Cell-specific 1*) encodes a small protein of 128 amino acids with a hydrophobic domain having characteristics of a GPI anchor, suggesting membrane localization of this protein (Xu et al. 1999b). *LGCI* is among genes abundantly expressed in the lily germ line, as 11 of 886 sequenced lily generative cell ESTs encode this protein (Okada et al. 2005a). Genes with similarity to *LGCI* are present in the *Arabidopsis* and rice genomes. An *LGCI* ortholog of *Arabidopsis* is expressed in generative and sperm cells only.

*GCSI* (*Generative Cell Specific-1*), another lily generative cell specific gene, was isolated by a differential display approach (Mori et al. 2006). *Arabidopsis* plants

with a gene knockout for a *GCS1* homologue gene, *HAP2*, show failure of fertilization pointing to a conserved role of the GCS1 family of proteins in gamete interactions. Sperm cells of *hap2* plants are incapable of fertilizing the egg or central cell, leading to the degeneration of the egg cell due to lack of fertilization (von Besser et al. 2006). *HAP2* encodes a 705 amino acid membrane protein that shows no similarity with genes of known function and shows no obvious functional motifs. *GCS1* and *HAP2* belong to a group of genes that can be defined as a conserved molecular signature for the plant male germ line, which is supported by the observation that its homologue in rice is one of the most highly expressed sperm cell specific genes (Gou and Russell, unpublished data).

In animal systems, gamete interactions are typically mediated by cell surface fertilization proteins (reviews by Vacquier 1998; Rubinstein et al. 2006). Sperms of both vertebrate and invertebrate animals carry surface proteins needed for gamete recognition, adhesion and fusion. For example, a sperm-expressed type 1a membrane protein is essential for sperm-egg fusion in mice (Inoue et al. 2005; Rubinstein et al. 2006). *HAP2/GSC1* protein is the first identified flowering plant sperm surface protein that is essential for fertilization.

*Arabidopsis* gene *AtGEX2*, with yet unannotated function, is expressed in a male germ line specific manner (Engel et al. 2005). *AtGEX2* protein is predicted to have six transmembrane domains. Genes homologous to *AtGEX2* have been detected in rice, maize and poplar genomes. Recent transcriptome analysis has shown that rice counterparts of *HAP2/GSC1* and *AtGEX2* correspond to some of the most up-regulated sperm cell specific transcripts. Thus, *LGC1*, *HAP2/GSC1* and *AtGEX2* are among products that could be considered as conserved male germ line transcriptional signatures in flowering plants.

An *Arabidopsis* male germ line specific gene *DUO*, encodes a novel R2R3-MYB transcription factor. The *DUO1* protein is present in the nucleus of generative and sperm cells and it promotes generative cell division by activating specific targets such as cyclin genes. An impressive demonstration is that cessation of generative cell division does not necessarily inhibit generative cells from acquiring cell fate normally reserved for sperm cells (Rotman et al. 2005). Failure of generative cell division in *Arabidopsis* mutant *cdc2a* (Nowack et al. 2006; Iwakawa et al. 2006) results in the production of only one cell, but it has fusion characteristics of a sperm rather than generative cell and has thus acquired a unicellular sperm identity. The serine/threonine protein kinase *cdc2* is a key regulator of the cell cycle, acting through cyclin-dependent phosphorylation. Loss of activity of chromatin assembly factor CAF-1 (Chen et al. 2008) causes delay and arrest of cell cycle resulting in a generative cell that acquires sperm cell

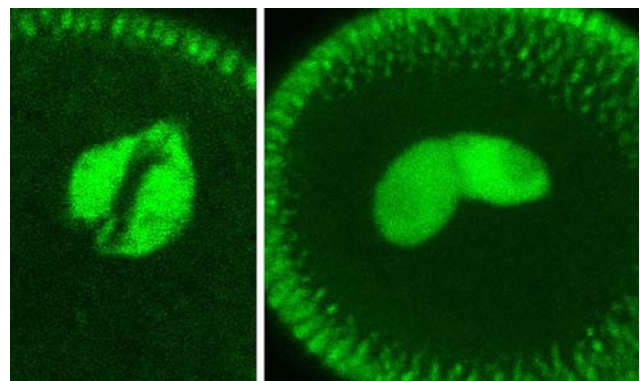
fate. In both cell cycle arrest mutants, the single gamete can effect fertilization. Intriguingly, while in *CAF-1* mutants, a single male gamete can fertilize either the egg cell or the central cell (Chen et al. 2008), the single male gamete in *cdc2* mutants fertilize only the egg cell (Nowack et al. 2006, Iwakawa et al. 2006).

### Transcriptional regulation of male germ line expressed genes

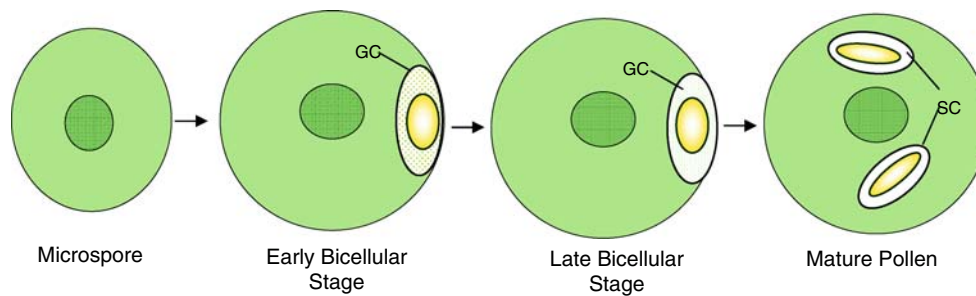
Of the categories of gene expression in male germ cells, the genes expressed specifically in the male germ line may define a molecular signature for male germ line cells. To discover the transcriptional regulation systems that control the expression of male germ cell specific genes, *LGC1* provides a model gene.

Transient transformation experiments using an *LGC1::GFP* protein reporter gene construct confirmed the male germ cell specific expression of this gene. Such promoter *LGC1* constructs do not show species specificity and are highly activated in male germ cells of bicellular lily and tobacco pollen (Singh et al. 2003). This promoter construct isolated from bicellular pollen of the monocot lily functions equally effectively in sperm cells of *Arabidopsis*, which is a dicot with tricellular pollen (Fig. 3). Thus, regulatory elements that control the specificity of male germ specific genes appear to be highly conserved and functional across broad taxonomic boundaries.

Control of the male germ unit specificity of *LGC1* promoter is linked to a 43 bp silencer fragment. If this fragment is removed, *LGC1* is expressed constitutively in all cells rather than expressing male germ line specificity,



**Fig. 3** Sperm cells of transgenic *Arabidopsis* plant carrying a fusion of lily *LGC1* promoter to GFP sequence (*pLGC1::GFP*), showing expression of green fluorescent protein. Living sperm cells were observed viewed using confocal microscopy in optical section (left) and composite image (right). Expression of GFP using a monocot promoter in a dicot plant illustrates the highly conserved nature of male germline transcriptional regulation mechanisms. (Image courtesy of X. P. Gou and S. Russell)



**Fig. 4** Schematic presentation of proposed model for the asymmetric distribution of the transcriptional repressor GRSF during pollen mitosis. A gene encoding GRSF protein (shown in green), is activated in the microspore; this repressor protein represses activation of male

suggesting regulation by suppression. A 26 kDa repressor protein named GRSF (germ-cell restrictive silencing factor) has recently been reported that shows specific binding to an 8–9 bp motif in the *LGCI* silencer region and is retained in other male germ specific genes (Haerizadeh et al. 2006). GRSF, a DNA binding repressor protein, functions by upregulated expression in all somatic and non-germ lineage nuclei, including those of the uninucleate microspore and pollen vegetative cell, with the only exception being the male germ line cells, where GRSF is absent. These data indicate that *LGCI* remains repressed in non-male germ line cells due to the binding of repressor protein GRSF. Asymmetric division of the haploid microspore leads to male germ line establishment and inactivation of GRSF in the male germ lineage concomitantly with the de-repression and transcriptional activation of *LGCI* (Fig. 4).

The core GRSF binding domain has been found in promoters of a number of male germ specific genes, including lily *gH3* (Okada et al. 2005a), lily *gH2A* (Ueda et al. 2005) and *Arabidopsis AtMGH3* (Okada et al. 2005b). Chromatin immunoprecipitation experiments reveal that GRSF binds with a silencing region not only in the lily *LGCI* promoter but also in the *gH3* promoter. These data showed that similar to *LGCI*, other coordinately expressed male germ specific genes are also kept repressed in somatic cells via binding of repressor nuclear proteins to upstream silencer elements (Haerizadeh et al. 2006).

Besides the GRSF binding motif, other shared *cis*-elements have been observed in the upstream regions of male germ specific genes (Okada et al. 2005a, b). A 9-bp motif (CCAAATTCA) is conserved between *AtMGH3* and lily *gH3* promoters. Intriguingly, this motif is conserved in four male gamete-specific genes in two different plant species, suggesting it to be a conserved *cis*-element in flowering plants.

Variant histones and post-translational chromatin modeling are common themes suspected to control chromatin organization, heterochromatic condensation and expression

germ line genes in non-germ line cells. Following asymmetric division, GRSF is absent from the smaller generative cell, leading to de-repression of gamete-specific genes. (Modified from Singh and Bhalla 2007)

in male germ cells (Xu et al. 1999a; Ueda et al. 2000). These male variants are suspected to be transcribed in a replication-independent manner, as they lack the well-conserved OCT *cis*-element in their promoter. This OCT element in histone gene promoters has been implicated in transcriptional control of replication-dependent histone synthesis in plants, which is associated with S-phase during cell proliferation (Meshi et al. 2000).

That several male germ-line specific genes of flowering plants are controlled by transcriptional mechanisms that maintain their silenced state in somatic cells and permit their up-regulation in male germ line cells may be anticipated.

#### Concluding remarks and prospects

The transcriptional repertoire of male germ line cells has revealed molecular complexity in the apparently simple lineage of sperm cells. Complete characterization of the molecular repertoire of the male germ line, however, will require more thorough profiling of genetic components at the transcriptome, proteome, protein-protein and protein-DNA levels of expression and interaction.

Although the foundation of these basic regulatory control mechanisms was established by the recent identification of a protein whose presence on the male gamete surface is essential for flowering plant fertilization, there are likely others, and much more work will be needed to decipher the function of the multitude of genes that seem to be expressed specifically in the male germ cells. The protein repertoire, with implicated roles in flowering plant sperm-egg interactions at the plasma membrane level is still unknown, but is rapidly approaching the technical capability of modern proteomic methods.

Initial identification of male germ line specific genes using tractable model systems like lily, coupled with functional disruption of their homologues in such model genomic plants as *Arabidopsis*, has turned out to be a

promising strategy for characterizing higher plant male gamete proteins that are critical for fertilization. The predicament of not knowing the surface proteins that mediate sperm-egg fusion in flowering plants is changing as newer, more sensitive approaches become available. Investigation of this problem could be advanced by applying a proteomics approach to identify genes that code proteins recovered from isolated generative and sperm cells, and will perhaps indicate functions for them. We may then be able to assess mechanisms and specific molecules of fertilization that may be conserved during the long evolution of eukaryotic organisms. Despite the completion of genomic sequences for such model plants as *Arabidopsis* and rice, and the availability of genomic microarray chips, the full repertoire of genes expressed in male germ line cells remains elusive. Further insights into the genes and gene pathways that regulate flowering male germ line differentiation will not only advance our fundamental understanding of these reproductive cells, but also the entire area of cell-cell recognition, membrane fusion and fertilization.

In animal systems, it has been reported that mature human sperm cells, despite their diminutive size, contain a complex repertoire of mRNAs and that some of these may be directly involved in coding unique proteins from male mRNA during early embryo development (Ostermeier et al. 2004). Whether plant sperm cells transfer translatable mRNAs to egg and central cell, and whether they are expressed, remains an open question. Recent data on gene expression profiles of tobacco eggs, zygotes and sperm cells show evidence for mRNAs that are present in both sperm cells and zygotes (Ning et al. 2006). While it has not been confirmed that these mRNAs in the zygote are transcribed de novo, or are contributed by sperm cytoplasm during fertilization, the possibility of male transmission of zygote-translated mRNA introduces intriguing and exciting possibilities. If a requirement for translated male mRNA is proven in zygotes, this insight is likely to change fundamental paradigms of zygote activation, control of sexual seed formation and early plant development.

**Acknowledgments** We thank Dr. Xiaoping Gou for providing unpublished data on *Plumbago* ESTs and expression of *pLGC1::GFP* constructs in *Arabidopsis*, and gratefully acknowledge research support by the Australian Research Council (ARC), The University of Melbourne and University of Oklahoma.

## References

- Andrews J, Bouffard GG, Cheadle C, Lu J, Becker KG, Oliver B (2000) Gene discovery using computational and microarray analysis of transcription in the *Drosophila melanogaster* testis. *Genome Res* 10:2030–2043
- Baarends WM, Hoogerbrugge JW, Roest HP, Ooms M, Vreeburg J, Hoeijmakers JH, Grootegoed JA (1999a) Histone ubiquitination and chromatin remodeling in mouse spermatogenesis. *Dev Biol* 207:322–333
- Baarends WM, Roest HP, Grootegoed JA (1999b) The ubiquitin system in gametogenesis. *Mol Cell Endocrinol* 25:5–16
- Becker JD, Feijo JA (2007) How many genes are needed to make a pollen tube? Lessons from transcriptomics. *Ann Bot* 100:1117–1123
- Bigler D, Takahashi Y, Chen MS, Almeida EAC, Osbourne L, White JM (2000) Sequence-specific interaction between the disintegrin domain of mouse ADAM 2 (fertilin beta) and murine eggs. Role of the alpha 6 integrin subunit. *J Biol Chem* 275:11576–11584
- Blomstedt CK, Knox RB, Singh MB (1996) Generative cells of *Lilium longiflorum* possess translatable mRNA and functional protein synthesis machinery. *Plant Mol Biol* 31:1083–1086
- Chen Z, Tan JLH, Ingouff M, Sundaresan V, Berger F (2008) Chromatin assembly factor 1 regulates the cell cycle but not cell fate during male gametogenesis in *Arabidopsis thaliana*. *Development* 135:65–73
- Churikov D, Siino J, Svetlova M, Zhang K, Gineitis A, Morton Bradbury E, Zalensky A (2004) Novel human testis-specific histone H2B encoded by the interrupted gene on the X chromosome. *Genomics* 84:745–756
- Eady C, Lindsey K, Twell D (1995) The significance of microspore division and division symmetry for vegetative cell-specific transcription and generative cell differentiation. *Plant Cell* 7:65–74
- Engel ML, Chaboud A, Dumas C, McCormick S (2003) Sperm cells of *Zea mays* have a complex complement of mRNAs. *Plant J* 34:697–707
- Engel ML, Davis RH, McCormick S (2005) Green sperm. Identification of male gamete promoters in *Arabidopsis*. *Plant Physiol* 138:2124–2133
- Faure J-E, Digonnet C, Dumas C (1994) An in vitro system for adhesion and fusion of maize gametes. *Science* 263:1598–1600
- Grewal S, Moazed D (2003) Heterochromatin and epigenetic control of gene expression. *Science* 301:798–802
- Haerizadeh F, Singh MB, Bhalla PL (2006) Transcriptional repression distinguishes somatic from germ cell lineages in a plant. *Science* 313:496–499
- Huck N, Moore JM, Federer M, Grossniklaus U (2003) The *Arabidopsis* mutant *feronia* disrupts the female gametophytic control of pollen tube reception. *Development* 130:2149–2159
- Inoue N, Ikawa M, Isotani A, Okabe M (2005) The immunoglobulin superfamily protein Izumo is required for sperm to fuse with eggs. *Nature* 434:234–238
- Iwakawa H, Shinmyo A, Sekine M (2006) *Arabidopsis* CDKA1;1, a CDC2 Homologue, controls proliferation of generative cells in male gametogenesis. *Plant J* 45:819–831
- Kimmins S, Sassone-Corsi P (2005) Chromatin remodeling and epigenetic features of germ cells. *Nature* 434:583–589
- Kranz E, Scholten S (2008) In vitro fertilization: Analysis of early post-fertilization development using cytological and molecular techniques. *Sex Plant Reprod* 21 (in press)
- Márton ML, Dresselhaus T (2008) A comparison of early molecular fertilization mechanisms in animals and flowering plants. *Sex Plant Reprod* 21 (in press)
- Meshi T, Taoka KI, Iwabuchi M (2000) Regulation of histone gene expression during the cell cycle. *Plant Mol Biol* 43:643–657
- Mori T, Kuroiwa H, Higashiyama T, Kuroiwa T (2006) GENERATIVE CELL SPECIFIC 1 is essential for angiosperm fertilization. *Nat Cell Biol* 8: 64–71
- Ning J, Peng X-B, Qu L-H, Xin H-P, Yan T-T, Sun M (2006) Differential gene expression in egg cells and zygotes suggests that the transcriptome is restructured before the first zygotic division in tobacco. *FEBS Lett* 580:1747–1752

- Nobuta K, Venu RC, Lu C, Belo A, Vemaraju K, Kulkarni K, Wang W, Pillay M, Green PJ, Wang GL, Meyers BC (2007) An expression atlas of rice mRNAs and small RNAs. *Nat Biotechnol* 25:473–477
- Nowack MK, Grini PE, Jakoby MJ, Lafos M, Koncz C, Schnittger A (2006) A positive signal from the fertilization of the egg cell sets off endosperm proliferation in angiosperm embryogenesis. *Nat Genet* 38:63–67
- Okada T, Bhalla PL, Singh MB (2005a) Transcriptional activity of male gamete-specific histone gcH3 promoter in sperm cells of *Lilium longiflorum*. *Plant Cell Physiol* 46:797–802
- Okada T, Endo M, Singh MB, Bhalla PL (2005b) Analysis of the histone H3 gene family in *Arabidopsis* and identification of the male-gamete-specific variant AtMGH3. *Plant J* 44:557–568
- Okada T, Bhalla PL, Singh MB (2006a) Expressed sequence tag analysis of *Lilium longiflorum* generative cells. *Plant Cell Physiol* 47:698–705
- Okada T, Singh MB, Bhalla PL (2006b) Histone H3 variants in male gametic cells of lily and H3 methylation in mature pollen. *Plant Mol Biol* 62:503–512
- Okada T, Singh M, Bhalla P (2007) Transcriptome profiling of *Lilium longiflorum* generative cells by cDNA microarray. *Plant Cell Rep* 26:1045–1052
- Ooi SL, Henikoff S (2007) Germline histone dynamics and epigenetics. *Curr Opin Cell Biol* 19:1–9
- Ostermeier GC, Miller D, Huntriss JD, Diamond MP, Krawetz SA (2004) Reproductive biology—delivering spermatozoan RNA to the oocyte. *Nature* 429:154
- Reinke V, Smith HE, Nance J, Wang J, Van Doren C, Begley R, Jones SJM, Davis EB, Scherer S, Ward S, Kim SK (2000) A global profile of germline gene expression in *C. elegans*. *Mol Cell* 6:605–616
- Rotman N, Durbarray A, Wardle A, Yang WC, Chaboud A, Faure JE, Berger F, Twell D (2005) A novel class of MYB factors controls sperm-cell formation in plants. *Curr Biol* 15:244–248
- Rotman N, Rozier F, Boavida L, Dumas C, Berger F, Faure J-E (2003) Female control of male gamete delivery during fertilization in *Arabidopsis thaliana*. *Curr Biol* 13:432–436
- Rubinstein E, Ziyat A, Wolf JP, Le Naour F, Boucheix C (2006) The molecular players of sperm-egg fusion in mammals: The Wnt signaling pathway and embryonic polarity; calcium waves at fertilization. *Sem Cell Devel Biol* 17:254–263
- Russell SD (1985) Preferential fertilization in *Plumbago*: Ultrastructural evidence for gamete-level recognition in an angiosperm. *Proc Natl Acad Sci USA* 82:6129–6132
- Russell SD (1991) Isolation and characterization of sperm cells in flowering plants. *Annu Rev Plant Physiol Plant Mol Biol* 42:189–204
- Russell SD (1992) Double fertilization. *Int Rev Cytol* 140:357–388
- Russell SD, Strout GW, Stramski AK, Mislan TW, Thompson RA, Schoemann LM (1996) Microgametogenesis in *Plumbago zeylanica* (Plumbaginaceae). 1. Descriptive cytology and three-dimensional organization. *Am J Bot* 83:1435–1453
- Sano Y, Tanaka I (2007) Detection of differentially expressed variant histone H3.3 in the vegetative nucleus of lily pollen. *Sex Plant Reprod* 20:27–33
- Sassone-Corsi P (2002) Unique chromatin remodeling and transcriptional regulation in spermatogenesis. *Science* 296:2176–2178
- Schultz N, Hamra FK, Garbers DL (2003) A multitude of genes expressed solely in meiotic or postmeiotic spermatogenic cells offers a myriad of contraceptive targets. *Proc Natl Acad Sci USA* 100:12201–12206
- Singh MB, Bhalla PL (2007) Control of male germ cell development in flowering plants. *Bioessays* 29:1124–1132
- Singh M, Bhalla PL, Xu HL, Singh MB (2003) Isolation and characterization of a flowering plant male gametic cell-specific promoter. *FEBS Lett* 542:47–52
- Singh MB, Xu HL, Bhalla PL, Zhang ZJ, Swoboda I, Russell SD (2002) Developmental expression of polyubiquitin genes and distribution of ubiquitinated proteins in generative and sperm cells. *Sex Plant Reprod* 14:325–329
- Uchiumi T, Komatsu S, Koshiba T, Okamoto T (2006) Isolation of gametes and central cells from *Oryza sativa* L. *Sex Plant Reprod* 19:37–45
- Ueda K, Tanaka I (1995) The appearance of male gamete-specific histones gH2B and gH3 during pollen development in *Lilium longiflorum*. *Dev Biol* 169:210–217
- Ueda K, Kinoshita Y, Xu ZJ, Ide N, Ono M, Akahori Y, Tanaka I, Inoue M (2000) Unusual core histones specifically expressed in male gametic cells of *Lilium longiflorum*. *Chromosoma* 108:491–500
- Ueda K, Suzuki M, Ono M, Ide N, Tanaka I, Inoue M (2005) Male gametic cell-specific histone gH2A gene of *Lilium longiflorum*: genomic structure and promoter activity in the generative cell. *Plant Mol Biol* 59:229–238
- Ulrich AS, Tichelaar W, Forster G, Zschornig O, Weinkauff S, Meyer HW (1999) Ultrastructural characterization of peptide-induced membrane fusion and peptide self-assembly in the lipid bilayer. *Biophys J* 77:829–841
- Vacquier VD (1998) Evolution of gamete recognition proteins. *Science* 281:1995–1998
- von Besser K, Frank AC, Johnson MA, Preuss D (2006) *Arabidopsis* HAP2 (GCS1) is a sperm-specific gene required for pollen tube guidance and fertilization. *Development* 133:4761–4769
- Wang YY, Kuang A, Russell SD, Tian HQ (2006) In vitro fertilization as a tool for investigating sexual reproduction of angiosperms. *Sex Plant Reprod* 19:103–115
- Xu H, Swoboda I, Bhalla PL, Sijbers AM, Zhao C, Ong EK, Hoeijmakers JH, Singh MB (1998) Plant homologue of human excision repair gene ERCC1 points to conservation of DNA repair mechanisms. *Plant J* 13:823–829
- Xu H, Swoboda I, Bhalla PL, Singh MB (1999a) Male gametic cell-specific expression of H2A and H3 histone genes. *Plant Mol Biol* 39:607–614
- Xu H, Swoboda I, Bhalla PL, Singh MB (1999b) Male gametic cell-specific gene expression in flowering plants. *Proc Natl Acad Sci USA* 96:2554–2558
- Yang H, Kaur N, Kiriakopoulos S, McCormick S (2006) EST generation and analyses towards identifying female gametophyte-specific genes in *Zea mays* L. *Planta* 224:1004–1014
- Zhang G, Gifford DJ, Cass DD (1993) RNA and protein-synthesis in sperm cells isolated from *Zea mays* L pollen. *Sex Plant Reprod* 6:239–243
- Zhang Z, Xu H, Singh MB, Russell SD (1998) Isolation and collection of two populations of viable sperm cells from the pollen of *Plumbago zeylanica*. *Zygote* 6:295–298
- Zigler KS, Lessios HA (2003) 250 million years of bindin evolution. *Biol Bull* 205:8–15