Molecular mechanisms of gene regulation during *Drosophila* spermatogenesis

Helen White-Cooper

School of Biosciences, Cardiff University, Museum Avenue, Cardiff CF10 3AX, UK

Correspondence should be addressed to H White-Cooper; Email: white-cooperh@cf.ac.uk

Abstract

The differentiation of sperm from morphologically unremarkable cells into highly specialised free-living, motile cells requires the co-ordinated action of a very large number of gene products. The expression of these products must be regulated in a developmental context to ensure normal cellular differentiation. Many genes essential for spermatogenesis are not used elsewhere in the animal, or are expressed elsewhere, but using a different transcription regulation module. Spermatogenesis is thus a good system for elucidating the principles of tissue-specific gene expression, as well as being interesting in its own right. Here, I discuss the regulation of gene expression during spermatogenesis in *Drosophila*, focussing on the processes underlying the expression of testis-specific genes in the male germline.

Reproduction (2010) 139 11–21

The anatomy and cell biology of *Drosophila* testes

The *Drosophila* testis comprises a blind-ended tube made of muscle and pigment cells filled with male germline cells and somatic support cells; see Fuller (1993) for a detailed description of spermatogenesis in *Drosophila*. The extreme tip of the tube has a region of thicker basal lamina, underlying a tight cluster of about 20 post-mitotic ‘hub’ cells, which constitute a signalling centre responsible for maintaining normal stem cells. Arranged in a rosette around the hub are two stem cell populations. Each mitotic division of a germline stem cell (GSC) regenerates a GSC, and produces a spermatogonium. Each mitotic division of a (somatic) cyst progenitor cell (CPC) regenerates a CPC, and usually produces a cyst cell, although CPC divisions can also generate new hub cells (Voog et al. 2008). Two cyst cells encapsulate each spermatogonium, to generate a cyst, and have a function analogous to the Sertoli cells of the mammalian testis. The cyst cells are now post-mitotic, while the encysted spermatogonium proceeds through four mitotic divisions to generate 16 primary spermatocytes. As in mammalian spermatogenesis, the spermatogonial divisions are characterised by incomplete cytokinesis, and stable cytoplasmic bridges link sister spermatocytes. The primary spermatocytes rapidly progress through pre-meiotic S-phase, and enter an extended G2 cell cycle phase lasting more than 3 days during which their volume increases substantially. Progress through the two meiotic divisions lead to 64 round spermatids; again incomplete cytokinesis results in cytoplasmic bridges connecting sister cells. Spermatid elongation results in extreme intracellular asymmetry. Each spermatid has the nucleus at the caudal (head) end, while flagellar elongation occurs at the rostral (tail) end. As the 64 spermatids from each single spermatogonium are still encysted, and interconnected, they polarise in concert, so that the cyst itself is asymmetric. During spermatid elongation differences between the two cyst cells become apparent. The head cyst cell covers the caudal end of the spermatids, while the tail cyst cell encapsulates the elongating tails. The cyst orients with the spermatid heads pointing toward the base of the testis, while the tails elongate towards the apical testis tip. Finally, the fully elongated spermatids undergo individualisation, during which time they extrude excess cytoplasm. Individualisation generates mature sperm which coil at the base of the testis before release into the seminal vesicle.

Gene expression in the *Drosophila* testis

**The germinal proliferation centre**

The apical region of the testis, comprising hub, stem cells (GSCs and CPCs) and mitotic spermatogonial cysts is collectively referred to as the germinal proliferation centre. Genes specifically expressed here are likely to be involved in regulation of stem cell behaviour; promoting stem cell fate in the cells attached to the hub and promoting differentiation of cells displaced from the hub. Several signalling pathways have been implicated in these processes, including the activation of JAK–STAT in both populations of stem cells in response to the ligand...
Upd secreted from the hub (Kiger et al. 2001, Tulina & Matunis 2001), and activation of the Dpp pathway in GSCs and spermatogonia (Bunt & Hime 2004, Kawase et al. 2004). These pathways have been recently reviewed elsewhere (Fuller & Spradling 2007). Some progress has been made in determining the transcriptional changes in target cells resulting from these signalling events. Microarray comparisons of samples enriched for GSCs (and CPCs) with samples enriched for spermatogonial cysts have revealed relatively few genes that express specifically in the hub, or stem cells, but has produced a short list of genes that are potential targets of the JAK–STAT signal pathway in stem cells (Terry et al. 2006). One potential target, zinc finger homeodomain-1 (zfh-1), a transcriptional repressor, is necessary and sufficient to maintain CPCs, presumably by repressing expression of genes associated with cyst cell differentiation. Intriguingly, persistent expression of zfh-1 in cyst cells prevented differentiation of both the cyst cells themselves, and non-autonomously blocked differentiation of the germline cells (Leatherman & DiNardo 2008). This indicates that the signalling in the apical region is more complicated than previously thought, with as yet unknown mechanisms acting to co-ordinate the behaviours of the two stem cell populations.

Primary spermatocytes activate a specialised transcriptional programme

Completion of the mitotic divisions and entry in the spermatocyte stage marks a dramatic turning point for the male germline cells. I am not aware of any genes expressed in male GSCs or spermatogonia that are not expressed in at least one other cell type in the animal; it is in the spermatocytes that the testis-specific gene expression is activated. Studies carried out in the 1960s and 1970s involving incorporation of ³H-uridine into nascent transcripts indicated that there is post-meiotic transcription in mammalian spermatids (Monesi 1965, Kierszenbaum & Tres 1975). Similar radiographic studies of Drosophila melanogaster showed no incorporation of ³H-uridine into transcripts from the mature primary spermatocyte stage onwards (Olivieri & Olivieri 1965, Gould-Somero & Holland 1974). Thus, while post-meiotic transcription is well known and widespread in mammalian round spermatids (Schultz et al. 2003, Potrzebowski et al. 2008), it was until recently believed to be essentially absent from Drosophila spermatogenesis. This predicts that transcripts for proteins needed at any stage in spermiogenesis will be transcribed in primary spermatocytes, and stored until needed. Consistent with this, transcripts for proteins known to be required specifically during spermiogenesis, including for example the sperm tail protein Don juan, and many others, are synthesised in the primary spermatocyte nuclei, and stored in a translationally repressed state for several days, until late in spermiogenesis. Similarly in vertebrates, while there is post-meiotic transcription, this shuts down on chromosome compaction, and late stages of spermiogenesis rely on translation of stored mRNA (Steger 2001).

The genes expressed by primary spermatocytes

Microarray analysis comparing different adult tissues revealed that ~50% of the genes in the genome are expressed in testes and 8% of the transcripts detected in adults are testis-specific, while a further 5% are testis-enriched (Andrews et al. 2000, Parisi et al. 2004, Chintapalli et al. 2007). Thus, about 25% of all the genes expressed in the testes are testis-specific or testis-enriched in expression compared with other tissues. These and similar studies reveal which genes are expressed in testes, but not the roles of the encoded proteins. Whole sperm proteomics has identified around 350 protein components of mature sperm (Dorus et al. 2006). About 50% of the identified sperm proteins are testis-enriched or -specific in their transcription. Reassuringly, known sperm-specific proteins, for example β2tubulin and cytoplasmic dynein are readily identified in these data sets. Testis-specific or enriched genes can be broadly grouped into two categories: those genes with obvious paralogues expressed in other tissues (and sometimes also testes), and those without such paralogues. A short list of genes in these two classes is listed in Table 1 to give a flavour of what types of genes are expressed in Drosophila primary spermatocyte. Genes with many different gene ontology classifications appear in these lists, including metabolism, cytoskeleton, chromosome organisation etc. However, perhaps the most telling statistic from classification of testis-specific genes, and sperm proteins, by gene ontology searches is that the largest category is ‘no functional prediction’. This is particularly striking when the ontology analysis is restricted to testis-specific sperm-proteome genes present as a single copy in the fly genome; very few of these genes have associated functional predictions. Even among those with a functional prediction the prediction has rarely been tested, and most of the genes remain uncharacterised.

Expression of genes on the X chromosome

There is good empirical evidence that the X chromosome is not a favourable genomic location for genes functioning primarily in the male, and that there is a general pattern of escape from the X chromosome of male-biased and testis-biased genes. Where a pair of genes have been created by a duplication event, the autosomal copy is much more likely to have testis-biased expression than the X-linked copy (Vibranovski et al. 2009). Potential evolutionary forces driving this include...
Table 1 Examples of testis-specific genes.

<table>
<thead>
<tr>
<th>Testis-specific gene</th>
<th>Paralog(s) expressed in other tissues</th>
<th>Function</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>βtub85D/β2h</td>
<td>βtub56D&lt;br&gt;βtub60D&lt;br&gt;βtub97EF</td>
<td>β-Tubulin. Meiotic spindle and axoneme structural component</td>
<td>Kemphues et al. (1979)</td>
</tr>
<tr>
<td>izo</td>
<td>Mart&lt;br&gt;dbh&lt;br&gt;Dpy-30L1&lt;br&gt;aly&lt;br&gt;Porin2&lt;br&gt;boule</td>
<td>Mitochondrial fusion, Thioredoxin, Chromatin structure, Transcriptional regulation, Mitochondrial function, Translational control</td>
<td>Hwa et al. (2002)&lt;br&gt;Svensson et al. (2003)&lt;br&gt;Vardanyan et al. (2008)&lt;br&gt;White-Cooper et al. (2000)</td>
</tr>
<tr>
<td>CG15287</td>
<td>Unknown&lt;br&gt;tub85D&lt;br&gt;tub97EF&lt;br&gt;bD&lt;br&gt;in&lt;br&gt;tub56D&lt;br&gt;ub60D&lt;br&gt;ub130&lt;br&gt;ub130d&lt;br&gt;ub60D&lt;br&gt;ub97EF&lt;br&gt;ub97EF</td>
<td>Unknown</td>
<td>_a</td>
</tr>
<tr>
<td>CG5062</td>
<td>Unknown&lt;br&gt;bub60D&lt;br&gt;bub97EF&lt;br&gt;ub56D&lt;br&gt;ub130&lt;br&gt;ub130d&lt;br&gt;ub60D&lt;br&gt;ub97EF</td>
<td>Unknown</td>
<td>_a</td>
</tr>
<tr>
<td>CG15891</td>
<td>Unknown&lt;br&gt;bub60D&lt;br&gt;bub97EF&lt;br&gt;ub56D&lt;br&gt;ub130&lt;br&gt;ub130d&lt;br&gt;ub60D&lt;br&gt;ub97EF</td>
<td>Unknown</td>
<td>_a</td>
</tr>
<tr>
<td>CG15198</td>
<td>Unknown&lt;br&gt;bub60D&lt;br&gt;bub97EF&lt;br&gt;ub56D&lt;br&gt;ub130&lt;br&gt;ub130d&lt;br&gt;ub60D&lt;br&gt;ub97EF</td>
<td>Unknown</td>
<td>_a</td>
</tr>
</tbody>
</table>

*aThese genes were selected as examples from the Drosophila sperm proteome dataset (Dorus et al. 2006) that were testis-specific in adults as deduced from FlyAtlas (Chintapalli et al. 2007), and confirmed as single copy in the Drosophila melanogaster genome by BLAST searching.

sexual antagonism: because males are haploid for the X chromosome the X-linked alleles spend more time in females than males, thus variants beneficial to males but detrimental to females will be selected against. Inactivation of the X chromosome during spermatogenesis could also provide strong selection against X-linkage of spermatogenic genes. These two forces combine in the sexual antagonism driving X inactivation (SAXI) hypothesis (Wu & Xu 2003). Inactivation of the X chromosome during meiotic prophase in vertebrate spermatocytes is well known (Khil et al. 2004), and it has long been argued that the X chromosome is inactivated in early Drosophila primary spermatocytes (Lifschytz 1972). Genome scale analysis of sex-specific gene expression has confirmed that the X chromosome is relatively depleted for male-specific genes compared to the autosomes (Parisi et al. 2003, 2004). This observation holds true both for the large class of testis-enriched genes, and the much smaller class of male-soma enriched genes. Only 43 out of the 381 proteins found as components of sperm by mass spectroscopy are encoded on the X chromosome (Dorus et al. 2006), again supporting under-representation of sperm genes on the X. Intriguingly, though, there are plenty of X-linked genes that are specifically expressed or enriched in testes, indicating that some promoters do work well in this location. Of the top 50 testis-enriched genes when compared to whole adult samples (www.flyatlas.org), 13 are located on the X chromosome. The X chromosome does appear to be a favourable location for expanding tandem clusters of structural genes, such as the Sdic (Ranz et al. 2003) and X-linked tektin gene families (Dorus et al. 2008). A thorough recent analysis of expression of transgene insertions on the X chromosome compared to the autosomes, has shown that at least one testis-specific promoter (ocnus) does function much more efficiently when autonomously inserted, although expression was detected from the X-linked inserts (Hense et al. 2007). This would imply that there is lower expression from the X-chromosome, although the generality of this observation for other promoters has not been tested.

The meiotic arrest loci: regulators of gene expression in primary spermatocytes

Although there must be co-ordination of varied cellular events during spermatid differentiation, genetic analysis reveals that most morphological events are independently regulated. For example, spermatid elongation involves flagellar axoneme synthesis, mitochondrial fusion and elongation of the mitochondrial derivate, and polarised growth of the plasma membrane. Spermatids mutant for fws, a subunit of the conserved oligomeric Golgi complex, or syntaxin 5, both important for endoplasmic reticulum–Golgi trafficking, initiate axoneme and mitochondrial elongation, but polarised cell growth, and thus cyst elongation, fails in these males (Xu et al. 2002, Farkas et al. 2003). In spermatids mutant for izo, a mitofusin, mitochondrial fusion fails; however, spermatid elongation occurs (Hales & Fuller 1997). Spermatid differentiation is also, surprisingly, not dependent on completion of the meiotic divisions. Spermatocytes mutant for the cell cycle activator twnie fail to undergo either meiotic division, but progress to spermatid differentiation as 4N, 16-cell cysts (White-Cooper et al. 1993).

Despite the independence of particular morphological events, genetic analysis has also revealed aspects of how the spermiogenic programme is co-ordinately regulated (Lin et al. 1996). A class of ‘meiotic arrest’ mutants have been discovered in which spermatocytes arrest development, rather than continuing into the meiotic divisions or spermiogenesis. Testes from male mutants for any of the meiotic arrest loci contain only stages of spermatogenesis up to and including mature primary spermatocytes. The primary spermatocytes in these testes do not enter...
the meiotic divisions neither do they initiate spermatid differentiation. The basal regions of the meiotic arrest mutant testes typically contain degenerating cells (Fig. 1). Morphologically these testes resemble that seen in testis biopsies of sterile human patients with meiosis I maturation arrest azoospermia (Meyer et al. 1992, Lin et al. 1996).

**The meiotic arrest loci fall into two distinct phenotypic classes**

Examination of the nuclear structures of the first meiotic arrest mutants to be discovered revealed that can, mia and sa arrest with partially condensed chromosomes that resemble normal prophase I structures. In contrast, the chromatin morphology was significantly disrupted in aly mutant spermatocytes, and the chromosomes appeared fuzzy (Lin et al. 1996). To understand how the meiotic arrest mutants affect progression of both meiosis and spermatid differentiation we examined the expression of several genes known to be important for these processes (White-Cooper et al. 1998). Genes important for primary spermatocyte function per se were expressed in mutant spermatocytes, revealing that the meiotic arrest genes are not global activators of transcription in primary spermatocytes. Strikingly, we found that mRNAs for genes known to have a role in spermiogenesis were expressed at very low levels can, mia and sa mutant primary spermatocytes, and were undetectable in aly mutant primary spermatocytes. aly also behaved differently from the other mutants with respect to cell cycle transcripts: aly is required for twine transcription, but can, mia and sa are required post-transcriptionally for Twine protein production. Thus, we concluded that the meiotic arrest genes are important for transcription in primary spermatocytes, primarily of genes whose products act during spermiogenesis. We further concluded that the meiotic arrest class could be subdivided into ‘aly-class’ and ‘can-class’ (White-Cooper et al. 1998), and this classification has been used with more recently discovered meiotic arrest loci (Table 2).

aly-class mutants dramatically affect the transcription of more than 1000 genes in testes, while can-class mutants have a somewhat more restricted set of targets. About 150 genes are expressed normally in can but not in aly (H White-Cooper, unpublished observations from microarray analysis). Furthermore, null mutations in aly-class genes typically abolish expression of target genes, while target gene expression is dramatically attenuated, but still detected, in testes null for can-class genes (Fig. 2).

Figure 1 Wild-type and meiotic arrest mutant Drosophila testes. When examined by phase contrast microscopy after gentle squashing, all stages of spermatogenesis are visible in a single wild-type testis, including motile sperm being released from the seminal vesicle at the testis base. In contrast, only stages up to and including mature primary spermatocytes are visible in meiotic arrest mutant testes (sa). Mutant testes are significantly smaller than wild-type. Both testes were photographed at the same magnification, and are oriented with their apical tips towards the top of the figure.

The can-class gene products form a testis-specific TFIID-paralogous complex

The multi-subunit basal transcription factor complex, TFIID, comprises TATA-binding protein (TBP) and up to 12 TBP-associated factors (TAFs), and plays a critical role in facilitating the interaction between gene promoter regions and RNA polymerase II (Pol II). Cloning of can showed that it encodes a testis-specific parologue of the ubiquitously expressed TAF, dTAF5 (formerly dTAF80; Hiller et al. 2001). Cloning of mia, sa, rye and nht revealed that they too encode tissue-specific TAFs, leading to the hypothesis that there is a testis-specific version of the ubiquitously required TFIID. This testis-TFIID includes the testis TAFs (tTAFs) encoded by the can-class meiotic arrest loci, complexed with a TAF1-2, a TAF1 splice isoform that is strongly testis enriched (Chen et al. 2005, Metcalf & Wassarman 2007). Presumably to interact with target gene promoters, this complex would also contain TBP (or one of the several TBP-related proteins), although it is possible that the DNA-binding activity usually conferred by TBP could be provided by the two AT-hook motifs found in TAF1-2. While a postulated testis–TFIID complex is consistent with the available data, it is important to note that certain TAFs are components not only of TFIID but also associate with histone acetyltransferase and Polycomb group (PcG) complexes, making functional predictions more problematic.
The discovery of the tTAFs led seductively to a model for their function, whereby they simply substitute for the generally expressed TFIID, and function as basal transcription factors for transcriptional activation, specifically at the promoters of spermiogenesis genes.

A simple substitution of tTAF subunits into a general TFIID complex predicts that the tTAF proteins would primarily co-localise with the masses of euchromatin in primary spermatocyte nuclei. However, the majority of tTAF protein, along with TAF1-2, is localised not to euchromatin, but rather to a subcompartment of the nucleolus (Chen et al. 2005, Metcalf & Wassarman 2007). These authors found that a minority of the tTAF staining was associated with the chromosomal masses in primary spermatocytes. Components of the generally expressed TFIID complex were either not detected in primary spermatocytes, or were chromatin associated, and excluded from the nucleolus (Metcalf & Wassarman 2007). Intriguingly, Pc, Polyhomeotic and dRing, all components of the Polycomb repression complex (PRC1), were also found to localise primarily in the nucleolus of primary spermatocytes, in a pattern coincident with that of the tTAFs (Chen et al. 2005).

This nucleolar localisation of PRC1 depends on the normal activity of the tTAFs, as in tTAF mutant primary spermatocytes Pc localised to chromatin but not to the nucleolus (Chen et al. 2005). These findings led to a hypothesis that the function of tTAFs in activating the testis-specific gene expression could be attributed to them being repressors of a repressor (specifically PRC1). In this scenario, tTAFs would sequester PRC1 away from testis-specific target promoters, thus de-repressing them. Testing this by chromatin immunoprecipitation on testis preparations revealed that tTAFs are bound to promoters of their target genes in primary spermatocytes (these being the only cells in the sample expressing the precipitated proteins), while in wild-type testes Pc was no more enriched at tTAF target promoters than at the promoters of non-target genes or intergenic regions. In contrast, and also in agreement with the ‘repressor of a repressor’ hypothesis, Pc was enriched at tTAF target promoters in tTAF mutant testes. The ‘tTAFs as an activator’ and ‘repressor of a

Table 2 Meiotic arrest loci.

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Gene full name</th>
<th>Molecular function</th>
<th>Classification</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>can</td>
<td>cannonball</td>
<td>TAF5</td>
<td>can-class</td>
<td>Hiller et al. (2001)</td>
</tr>
<tr>
<td>mia</td>
<td>meiosis I arrest</td>
<td>TAF6</td>
<td>can-class</td>
<td>Hiller et al. (2004)</td>
</tr>
<tr>
<td>sa</td>
<td>spermatocyte arrest</td>
<td>TAF8</td>
<td>can-class</td>
<td>Hiller et al. (2004)</td>
</tr>
<tr>
<td>nht</td>
<td>no hitter</td>
<td>TAF4</td>
<td>can-class</td>
<td>Hiller et al. (2004)</td>
</tr>
<tr>
<td>mip40</td>
<td>Myb interacting</td>
<td>Unknown (homolog of C. elegans lin-37)</td>
<td>can-class</td>
<td>Beall et al. (2007)</td>
</tr>
<tr>
<td>aly</td>
<td>protein, 40 kDa</td>
<td>Unknown (homolog of C. elegans lin-9)</td>
<td>aly-class</td>
<td>White-Cooper et al. (2000)</td>
</tr>
<tr>
<td>comr</td>
<td>cookie monster</td>
<td>Unknown</td>
<td>aly-class</td>
<td>Jiang &amp; White-Cooper (2003)</td>
</tr>
<tr>
<td>tomb</td>
<td>tombola</td>
<td>DNA binding (homolog of C. elegans lin-54)</td>
<td>aly-class</td>
<td>Jiang et al. (2007)</td>
</tr>
<tr>
<td>topi</td>
<td>matotopelli</td>
<td>DNA binding</td>
<td>aly-class</td>
<td>Perezgazga et al. (2004)</td>
</tr>
<tr>
<td>achi-vis</td>
<td>achintya and vismay</td>
<td>DNA binding</td>
<td>aly-class</td>
<td>Ayyar et al. (2003) and Wang &amp; Mann (2003)</td>
</tr>
</tbody>
</table>

Figure 2 In situ hybridisation revealing gene expression defects in meiotic arrest mutants. hml (A) is expressed at similar levels in both wild-type and comr (B) primary spermatocytes. CG1014 (C) is expressed in wild-type primary spermatocytes, with the transcript persisting into spermiogenesis. In contrast, the expression of CG1014 is not detected in achi-vis mutant testes (D).
Repressor’ hypotheses are not mutually exclusive; it is likely that tTAFs both remove a repressor (Pc) and act as an activator, for example by recruiting the Trithorax group complex.

The aly-class gene products form testis-specific Myb-MuvB-paralogous complex

When aly was cloned it was clear that the gene was evolutionarily conserved from plants to animals (but not to fungi; White-Cooper et al. 2000). The only aly homologue to have been studied in other systems was the Caenorhabditis elegans gene lin-9, which falls into the SynMuVB genetic pathway that negatively regulates vulval induction. At that time the mechanism by which lin-9 regulates cell fate in C. elegans was not determined (Beitel et al. 2000). Genome sequencing and phylogenetic analysis revealed that aly is one of two Drosophila paralogues of the lin-9 gene, the other being mip130. Clues to the potential role of aly come from more recent analysis of its homologues. Mip130 protein was purified from ovary extracts, in a complex with Drosophila Myb, CAF1 and two other previously unknown proteins, Mip120 and Mip40 (Beall et al. 2002). This complex has since been re-purified under slightly different conditions to reveal more complex subunits, and is known as the MybMuVB (MMB) or dREAM complex (Korenjak et al. 2004, Lewis et al. 2004). The additional subunits include Drosophila Rbl (Retinoblastoma (Rb) homologue), E2F2, Dp and dLin-52. Myb, E2F2, Dp and Mip120 are all known DNA binding proteins, and the function of the dREAM/MMB complex appears to predominantly be to repress gene expression (Lewis et al. 2004). Excitingly, cloning of the C. elegans SynMuVB pathway genes has revealed a similar list of genes, whose products form the DRM complex (Harrison et al. 2006). The core DRM complex appears to comprise Lin-35 (Rb), Efl-1 (E2F2), Dpl-1 (Dp), Lin-53 (CAF1), Lin-37 (Mip40), Lin-52 and Lin-54, in addition to the aly/mip130 homologue, Lin-9. A similar complex, named LINC (Schmit et al. 2006). This reveals that Aly and Comr are in at least two distinct complexes, one with Mip40, but lacking Achi/Vis, the other containing Achi/Vis.

The tissue-specific dREAM versus tMAC subunits are all putative DNA binding proteins. Use of testis-expressed paralogues of core complex components, along with different DNA binding proteins in testes versus somatic tissues would ensure that spermiogenesis genes are activated specifically in testes. dREAM and DRM complexes have both predominantly been linked to a role in transcriptional repression rather than activation, although an activatory role has been shown for the human complex LINC (Schmit et al. 2007). This raises the question of whether the aly-class meiotic arrest genes act directly as transcriptional activators, or as repressors of a repressor. Supporting a potential activatory role is the observation that all the aly-class proteins co-localise with euchromatin in primary spermatocytes, and that this localisation is essential for their function (White-Cooper et al. 2000, Jiang & White-Cooper 2003, Wang & Mann 2003, Jiang et al. 2007). Direct evidence for Achi/Vis acting as transcriptional activators in testes

Table 3 Homologous complexes across metazoa compared to testis meiotic arrest complex (tMAC) the Drosophila testis-specific complex. Common synonyms are separated by /

<table>
<thead>
<tr>
<th>H. sapiens</th>
<th>C. elegans</th>
<th>D. melanogaster</th>
<th>D. melanogaster</th>
</tr>
</thead>
<tbody>
<tr>
<td>DREAM/LINC</td>
<td>MuvB/DRM</td>
<td>dREAM/MybMuv</td>
<td>tMAC</td>
</tr>
<tr>
<td>RBL2/p130</td>
<td>Lin-35</td>
<td>RBF1 or RBF2</td>
<td>Aly</td>
</tr>
<tr>
<td>E2F4 or E2F5</td>
<td>Efl-1</td>
<td>E2F2</td>
<td>Comr</td>
</tr>
<tr>
<td>Dpl-1</td>
<td>Dp</td>
<td></td>
<td>Topi</td>
</tr>
<tr>
<td>RIPPB4</td>
<td>Lin-53</td>
<td>CAF1/p55</td>
<td>Aly</td>
</tr>
<tr>
<td>LIN9</td>
<td>Lin-9</td>
<td>Mip130</td>
<td></td>
</tr>
<tr>
<td>LIN37</td>
<td>Lin-37</td>
<td>Mip40</td>
<td></td>
</tr>
<tr>
<td>LIN52</td>
<td>Lin-52</td>
<td>dLin52</td>
<td></td>
</tr>
<tr>
<td>LIN54</td>
<td>Lin-54</td>
<td>Mip120</td>
<td>Tomb</td>
</tr>
<tr>
<td>MYB-B/MYBL2</td>
<td></td>
<td>Myb</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comr Topi</td>
<td></td>
</tr>
</tbody>
</table>
has been provided by expressing fusions of Achi/Vis with strong transactivation (VP16) and strong repression (EnR) domains. Expression of Achi-VP16 fusion proteins rescued the achi+ vis mutant phenotype, while expression of Achi-EnR did not (Wang et al. 2008).

**Cross-talk between the tTAFs and tMAC?**

Diagnostic RNA in situ hybridisations have been used to successfully subdivide aly-class and can-class meiotic arrest mutants, and this correlated with proteins encoded aly-class genes being tMAC subunits and can-class genes being tTAFs. However mip40 does not fit neatly into this categorisation. Mip40 protein is clearly a tMAC component, however mip40 mutants are apparently can-class (Beall et al. 2007). This might indicate that Mip40 is critical for mediating interactions between the two complexes, and clearly warrants further investigation.

**Testis-specific promoters**

Testis-specific genes are activated only in testes and kept silent in other tissues of the fly. It is somewhat surprising that the promoter elements conferring testis specificity that have been identified to date in Drosophila are relatively small (Table 4). One of the first to be studied was the testis-specific β-tubulin isoform, β2tubulin (βTubb85D). Astonishingly, a fragment consisting of only 53 bp of promoter region, plus the first 71 bp of the 5'UTR was sufficient to confer testis-specific expression on reporter genes. The UTR requirement was further refined to 23 bp. Within the promoter a 14 bp motif, β2UE1 was shown to be critical for testis-specific expression (Michiels et al. 1989). This testis-specific reporter gene expression depends on the normal function of the meiotic arrest genes, just as expression of the endogenous gene does (Hiller et al. 2001). Sequences related to the β2UE1 have been found upstream of several other testis-specific transcriptional start sites (Yang et al. 1995, Nurminsky et al. 1998), but this sequence is not found in many other testis promoters, so cannot be considered a signature sequence for testis-specific expression. The minimal sequences needed for testis-specific expression (and typically also translational control) of several other genes are listed in Table 3. These short promoters presumably contain a landing site for the testis-specific transcriptional control machinery, tTAFs and tMAC, outlined above. The fact that testis-specific control elements are so small may be important in allowing new gene duplicates to be expressed in testes. For example, the testis-specific gene Sdic evolved from duplication and fusion of the genes AnnX and Cdic, encoding an annexin and a cytoplasmic dynein intermediate chain respectively. The promoter for the newly generated Sdic gene was derived from coding regions from AnnX and intronic sequence from Cdic (Ranz et al. 2003).

At a genomic organisation scale, testis-specific genes have been found to cluster significantly in the genome (Boutanaev et al. 2002, Spellman & Rubin 2002). The clustering of testis-specific genes probably relates to higher order chromatin structure in primary spermatocytes and/or in other cells types. The cluster genomic regions are silenced in the soma, and have a compact chromatin structure (Kalmykova et al. 2005, Shevelyov et al. 2009). However there is relatively little variation in expression levels in testes of transgenes inserted into random genomic (autosomal, see below) positions, so while chromatin domain organisation may facilitate testis-specific expression, it cannot be critical for ensuring normal gene expression levels. It is however fair to say that the relative importance of genomic context versus promoter sequences for testis-specific gene expression is complex, and remains unresolved.

**Transcriptional activity in elongating spermatids**

As discussed above, it had been widely accepted that there is no post-meiotic transcription in Drosophila testes. Transcripts for spermiogenesis products would be synthesised in primary spermatocytes, and stored (safe from degradation) until needed. Upon translation these transcripts would become destabilised, and so the time at which a transcript disappears during spermiogenesis correlates with the time that the protein is produced. This phenomenon is useful in analysis of the function of

---

**Table 4 Testis specific promoters are small. Minimal fragments tested that confer testis-specific expression on reporter genes in assays of transgenic flies.**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Flanking 5’ (bp)</th>
<th>5’UTR (bp)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>β2tubulin</td>
<td>53</td>
<td>23</td>
<td>Michiels et al. (1989)</td>
</tr>
<tr>
<td>gdl</td>
<td>328</td>
<td>132</td>
<td>Schulz et al. (1990)</td>
</tr>
<tr>
<td>janB</td>
<td>174</td>
<td>107</td>
<td>Yancostas &amp; Lepesant (1990)</td>
</tr>
<tr>
<td>mst87F</td>
<td>102</td>
<td>201</td>
<td>Kuhn et al. (1988)</td>
</tr>
<tr>
<td>djl</td>
<td>11</td>
<td>95</td>
<td>Hempel et al. (2006)</td>
</tr>
<tr>
<td>dj</td>
<td>23</td>
<td>115</td>
<td>Blumer et al. (2002)</td>
</tr>
<tr>
<td>mst36Fb</td>
<td>145</td>
<td>10</td>
<td>Di Cara et al. (2006)</td>
</tr>
<tr>
<td>Mst35Ba</td>
<td>280</td>
<td>N/A</td>
<td>Jayaramaiah Raja &amp; Renkwitz-Pohl (2005)</td>
</tr>
<tr>
<td>Mst77F</td>
<td>278</td>
<td>N/A</td>
<td>Jayaramaiah Raja &amp; Renkwitz-Pohl (2005)</td>
</tr>
<tr>
<td>dhod</td>
<td>54</td>
<td>35</td>
<td>Yang et al. (1995)</td>
</tr>
</tbody>
</table>

**www.reproduction-online.org**
uncharacterised genes in testes; simply by analysing the transcript accumulation/disappearance profile in normal testes one can infer approximately the stage at which the protein product is produced. In my lab, we have carried out many RNA in situ hybridisations for this purpose (see www.fly-ted.org). The vast majority of genes whose transcripts we detected in spermatids were also detected at equally high levels in primary spermatocytes (529/553 genes), as expected. However, we discovered an entirely unexpected class of gene as rare exceptions to this rule (Barreau et al. 2008). For this exceptional class the staining pattern was most consistent with transcription during a particular stage of spermatid elongation. These transcripts, collectively referred to as ‘comets and cups’ were detected at very low levels in primary spermatocytes, were not detected in early spermatids, and were detected at high levels specifically at the elongating ends of mid-late elongation stage spermatids (Fig. 3). We confirmed this post-meiotic transcription by developing a single cyst quantitative RT-PCR protocol. As a general pattern we found that single isolated cysts of primary spermatocytes contained detectable comet or cup transcripts, cysts of early elongation spermatids did not have detectable comet or cup transcripts, and cysts in mid-late elongation had high levels of the transcripts. This data mirrors the pattern seen with RNA in situ hybridisation.

Highly transcriptionally active primary spermatocyte nuclei have a diameter of ~17 μm; nuclei in onion stage spermatids are spherical, and about 6 μm in diameter, while the needle shaped nuclei of mature spermatids are ~9 μm in length, with a maximum diameter of 0.3 μm (Tokuyasu 1974). For comparison, interphase nuclei at late syncytial blastoderm stage are ~10 μm in diameter. Spermiogenesis, thus, involves a decrease in nuclear volume of ~200-fold, while the meiotic divisions and nuclear reformation only deliver a 20-fold volume decrease. Spermatid nuclei first elongate and thin somewhat, which includes a fivefold volume decrease, and take on a very asymmetric shape resembling (in transverse section) an oxbow at early ‘canoe’ stage; at late canoe stage they are kidney-shaped (Tokuyasu 1974). Drosophila sperm DNA is packaged with small nuclear basic proteins, related to histone H1, termed protamine-like proteins. These lysine-rich proteins are functionally similar to the arginine-rich vertebrate protamines. Three protamine-like proteins have been described for D. melanogaster, Mst35Ba (Protamine A), Mst35Bb (Protamine B) and Mst77F (Jayaramaiah Raja & Renkawitz-Pohl 2005). All three of these proteins are found in nuclei of mature sperm. As with vertebrates, the transition from nucleosomal packaging to protamines is facilitated by transition proteins, including Tpl94D (Rathke et al. 2007). The histone–transition protein–protamine switch occurs at late canoe stage, and the major loss of nuclear volume occurs from this stage onwards. Using the single cyst RT-PCR assay on testes expressing fluorescently-tagged histones, protamines and transition proteins, we found that the comet and cup transcription occurs just before the switch, when the majority of nuclear DNA is still not highly condensed (Barreau et al. 2008). Active Pol II has also been detected specifically at the late canoe stage of spermiogenesis (Rathke et al. 2007). Although not directly tested, we assume that transcription only occurs on nucleosomally packaged chromatin. The relative timing of cessation of transcription in vertebrate spermatids and the vertebrate histone–transition protein–protamine switch has not been determined. It is not clear if the shut-off of transcription depends on the initiation of chromatin compaction, whether chromatin compaction depends on stopping transcription, or whether these events are mechanistically independent. There is a clear difference between the post-meiotic transcription in Drosophila and that in mammals. Transcription is relatively high through early stages of spermatid differentiation in mammals, shutting-off during chromatin compaction. In Drosophila, we found a general shutdown of transcription at the end of the primary spermatocyte.

Figure 3 Post meiotic expression revealed by in situ hybridisation. sunz mRNA is detected at low levels in primary spermatocytes, not detected in early elongation spermatids, and is detected at high levels at the distal ends of mid-elongation spermatids.
stage; we did not detect transcript accumulation in early spermatids, rather, we found an abrupt re-activation of transcriptional capacity at mid-elongation stages.

At least one of the comet genes, soti, is required for male fertility, as we found that spermatid individualisation fails in soti homozygotes (Barreau et al. 2008). It is interesting to note that soti heterozygotes are fertile, and transmit the soti mutant chromosome. Therefore, spermatid cysts in Drosophila, as in mammals, are functionally diploid, i.e. the 32 haploid soti mutant spermatids can be rescued by the normal soti allele carried by the other 32 spermatids in the cyst.

We identified 24 comet and cup genes in an in situ hybridisation screen of ~1200 genes. Given our lack of systematic search strategy we have probably not identified all post-meiotically transcribed genes, and a genome scale approach is likely to yield more such genes. Analysis of the local chromatin environments of the set of 24, for example looking at gene density, clustering with co-expressed genes etc. yielded no clues as to why these genes are expressed in spermatids. The genes are found in unremarkable chromatin contexts. We have generated many transgenic lines containing genomic regions of several of the comet and cup genes. All the independent insertions mirror the endogenous expression patterns, so there appears not to be a specialised chromatin region permissive for post-meiotic transcription. As noted earlier, testis expressed genes tend to have short promoter regions, and the comet and cup genes also follow this pattern: ~1 kb of genomic flanking DNA is sufficient to recapitulate the normal expression pattern (we have not evaluated shorter fragments). Further experiments are needed to identify the DNA regions driving post-meiotic expression, and of course the transcriptional activators that promote this.

Why is there post-meiotic transcription in Drosophila testes? Given that the vast majority of genes involved in spermiogenesis are transcribed in primary spermatocytes, it is interesting to ask why there is a small exceptional class. All the comet and cup transcripts are detected in primary spermatocytes, albeit at low levels, so their transcription in these cells is clearly not detrimental, although these early-produced transcripts do not persist into elongation stages. It is attractive to speculate that the post-meiotic transcription and RNA localisations are linked. In our many in situ hybridisations the only genes whose transcript levels were significantly elevated in spermatids compared to spermatocytes also showed dramatic localisation of the mRNA to the spermatid elongating ends. It is possible that other, non-localised, transcripts are also made in spermatids, but that the qualitative nature of in situ hybridisation means this is not the most appropriate method for identifying other post-meiotically transcribed genes. For the comet and cup genes, perhaps the machinery required to localise these transcripts only becomes active during spermatid elongation. Further investigation is clearly warranted to determine how and why the comet and cup transcripts localise to the growing tips of spermatids.

Concluding remarks

As spermatogenesis proceeds gene expression patterns change considerably as a developmentally regulated transcriptional control machinery becomes engaged. The expression of a very large number of testis-specific genes in Drosophila primary spermatocytes is critically dependent on the activities of two protein complexes. The relative simplicity of this system compared to cascades of transcription factor activation in other developmentally regulated systems is attractive for studying how genes are specifically activated at the appropriate time in development.

Declaration of interest

The author declares that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

Work in my laboratory is funded by the Royal Society, the Wellcome Trust (meiotic arrest genes), and the BBSRC (in situ hybridisations and post-meiotic expression).

Acknowledgements

Thanks are due to my current and former lab colleagues for numerous stimulating discussions, most notably Jianqiao Jiang, Elizabeth Benson, Karen Doggett, Carine Barreau, and Nina Baushek. Thanks too to Ceri Morris for proof-reading the manuscript.

References


Schultz N, Hamra F & Garbers D 2003 A multitude of genes expressed solely in meiotic or postmeiotic spermatogenic cells offers a myriad of contraceptive targets. PNAS 100 12201–12206.


Received 2 March 2009
First decision 15 April 2009
Revised manuscript received 20 August 2009
Accepted 14 September 2009