Regulation of transcription of meiotic cell cycle and terminal differentiation genes by the testis-specific Zn-finger protein *matotopetli*

Lucia Perezgasga[1,*,‡], JianQiao Jiang[2,‡], Benjamin Bolival, Jr[1,‡], Mark Hiller[1,†], Elizabeth Benson[2], Margaret T. Fuller[1] and Helen White-Cooper[2,§]

[1]Departments of Developmental Biology and Genetics, Stanford University School of Medicine, Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA
[2]Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK
*Present address: Departamento de Genética y Fisiología Molecular, Instituto de Biotecnología, UNAM, Avenida Universidad #2001 Col. Chamilpa Apdo. Postal 510-3, Cuernavaca, Mor. 62250 México
†Present address: Department of Biological Sciences, Goucher College, 1021 Dulaney Valley Road, Baltimore, MD 21204-2794, USA
‡These authors contributed equally to the work
§Author for correspondence (e-mail: helen.white-cooper@zoo.ox.ac.uk)

Accepted 11 December 2003

Development 131, 1691-1702
Published by The Company of Biologists 2004
doi:10.1242/dev.01032

Summary

A robust developmentally regulated and cell type specific transcriptional programme is activated in primary spermatocytes in preparation for differentiation of the male gametes during spermatogenesis. Work in *Drosophila* is beginning to reveal the genetic networks that regulate this gene expression. The *Drosophila aly*-class meiotic arrest loci are essential for activation of transcription of many differentiation-specific genes, as well as several genes important for meiotic cell cycle progression, thus linking meiotic cell cycle progression to cellular differentiation during spermatogenesis. The three previously described *aly*-class proteins (*aly*, *comr* and *achi*/*vis*) form a complex and are associated with chromatin in primary spermatocytes. We identify, clone and characterize a new *aly*-class meiotic arrest gene, *matotopetli* (*topi*), which encodes a testis-specific Zn-finger protein that physically interacts with Comr. The *topi* mutant phenotype is most like *achi/vis* in that *topi* function is not required for the nuclear localization of Aly or Comr, but is required for their accumulation on chromatin. Most target genes in the transcriptional programme depend on both *topi* and *achi/vis*; however, a small subset of target genes are differentially sensitive to loss of *topi* or *achi/vis*, suggesting that these *aly*-class predicted DNA binding proteins can act independently in some contexts.

Supplemental data available

Key words: Spermatogenesis, Transcription, Chromatin, Meiosis, Differentiation.

Introduction

Tissue and cell type specific transcriptional activation is a fundamental feature of cellular differentiation that mediates many changes in cell behaviour and morphology during development. One of the most astonishing developmentally regulated changes in cell morphology occurs in the production of sperm. The cellular differentiation events that constitute spermiogenesis require many gene products used at no other time in development. The unique features of male gamete differentiation are reflected in a developmentally specific transcription programme that initiates in primary spermatocytes in preparation for spermiogenesis. In *Drosophila* spermatogenesis, transcription is essentially shut down as mature primary spermatocytes enter the meiotic divisions. Therefore, gene products required during differentiation are largely produced from transcripts expressed pre-meiotically (Olivieri and Olivieri, 1965; Schafer et al., 1995). A two-part genetic module, the meiotic arrest genes, is responsible for most of this transcriptional activation, and ensures that a large number of genes required for normal cellular morphogenesis are co-expressed in a cell type and developmental stage-specific manner (Lin et al., 1996; White-Cooper et al., 1998).

Males mutant for any of the meiotic arrest genes [including *always early* (*aly*), *cannonball* (*can*), *meiosis I arrest* (*mia*), *spermatocyte arrest* (*sa*), *cookie monster* (*comr*), *achintya/vismay* (*achi/vis*) and *no-hitter* (*nht*)] are viable but sterile. Testes from flies mutant for any one of these genes contain morphologically normal spermatogenic cells up to mature primary spermatocytes. However, the mutant cells then arrest and fail to initiate either the meiotic divisions or spermatid differentiation (Ayyar et al., 2003; Jiang and White-Cooper, 2003; Lin et al., 1996). Although transcription of a number of broadly expressed genes (such as *cyclin A*) is normal in the mutant spermatocytes, transcription of many spermiogenesis genes (e.g. the mitochondrial fusion protein *fzo*) is very low or undetectable. The *aly*-class meiotic arrest genes (*aly*, *comr*, *achi/vis*) are likely to act in a pathway distinct from the *can*-class genes, because they are required for transcription of a wider range of target genes. In addition to a

role in transcription of spermiogenesis genes, *aly*-class genes are also essential for the transcriptional activation in primary spermatocytes of several genes required for progression into the meiotic divisions (namely *twine*, *cycB*, *boule*). By contrast *can*, *mia*, *sa* and *nht* (can-class genes) are not required for transcription of these cell cycle genes; however, they are required for translation of twine. The *can*-class meiotic arrest genes analysed to date encode testis specific homologues of more generally transcribed TATA-binding protein associated factors (TAFs) (Aoyagi and Wassarman, 2000; Hiller et al., 2001) (M.H. and M.T.F., unpublished). TAFs are subunits of the basal transcription factor TFIID, which is involved in recruitment of the RNA polymerase II holoenzyme to the proximal promoter region. Thus, the *can*-class genes are likely to activate full levels of transcription of spermatid differentiation genes through intimate association with target promoters (Hochheimer and Tjian, 2003).

Not all genes transcribed in primary spermatocytes are controlled by the meiotic arrest pathway, so it is likely the targeting of particular promoters is mediated through sequence-specific DNA-binding activity. Thus, we would expect that one (or more) of the meiotic arrest genes encode a DNA-binding protein(s), whereas others are likely to encode regulatory proteins. *aly* is one of two *Drosophila* homologues of the *C. elegans* vulval differentiation regulator, *lin-9* (Beitel et al., 2000; White-Cooper et al., 2000). The SynMuvB pathway, in which *lin-9* acts, is thought to negatively regulate adoption of vulval fate and promote hypodermal fate, through activation of a histone deacetylase chromatin remodelling complex (Ferguson and Horvitz, 1989; Lu and Horvitz, 1998; Solari and Ahringer, 2000). *comr* encodes a novel acidic protein with no significant similarities to other proteins in current sequence databases (Jiang and White-Cooper, 2003). *achi*/*vis* was the first meiotic arrest locus to be shown to encode products with sequence-specific DNA-binding activity (Ayyar et al., 2003; Wang and Mann, 2003). The TALE class homeodomain proteins encoded by the recent gene duplication pair *achi/vis* are virtually identical to each other and are homologues of the human TGIF sequence-specific DNA-binding factor (Ayyar et al., 2003; Wang and Mann, 2003). Both *achi* and *vis* are expressed in many cells throughout development. However, flies that are null mutant for both genes are viable but male sterile, with a meiotic arrest phenotype. Achi/Vis proteins are localized to chromatin in wild-type primary spermatocytes and in *aly* mutant testes (Wang and Mann, 2003).

Despite lacking any predicted DNA-binding motifs, both Aly and Comr proteins are concentrated on chromatin in primary spermatocytes. However, the nuclear localizations of Aly and Comr are mutually dependent: Aly remains cytoplasmic in *comr* mutant testes and vice versa. By contrast, the localization of Aly and Comr to the nucleus is independent of *achi/vis* (Ayyar et al., 2003; Jiang and White-Cooper, 2003; Wang and Mann, 2003; White-Cooper et al., 2000). Aly, Comr and Achi/Vis co-immunoprecipitate from testis protein extracts, suggesting that they interact in a common complex in vivo (Wang and Mann, 2003).

We describe the identification, cloning and characterization of the *Drosophila aly*-class meiotic arrest gene, *matotopetli* (*topi*). The similarity in phenotype between *topi* and *aly*, *comr* and *achi/vis* suggests that these genes act together in a common pathway. *topi* encodes a testis-specific predicted Zn-finger protein, making Topi the second putative sequence-specific transcription factor to be placed into the *aly* class. At least two separate Zn-finger containing regions of Topi are sufficient to bind directly to a 100 amino acid region in the middle of the Comr protein. Like *achi*/*vis*, *topi* activity is not required for the nuclear localization of Aly and Comr proteins. We suggest that the DNA-binding activities of Topi and Achi/Vis are required to recruit Aly and Comr to the promoters of target genes in primary spermatocytes. Consistent with this idea, the majority of target genes require both the *topi* and *achi/vis* DNA-binding factors for full transcriptional activation. However, rare exceptions to this rule have been revealed via microarray analysis. A small number of target promoters rely to a much greater extent on one or other of these DNA-binding activities, indicating that in at least some contexts the activities of *achi*/*vis* and *topi* are independent.

Materials and methods

Fly strains and husbandry

Flies were raised on standard cornmeal molasses (or sucrose) agar at 25°C. Visible markers and balancer chromosomes are described in FlyBase except where otherwise noted (FlyBase, 1999). *red e* or *bw; st/TM6B* was used as the wild-type strain. *aly*^{5} *red e/ TM6C, w; can*^{3} *red e/ TM6C, cn achi*^{Z3922}*vis*^{Z3922} *bw/ CyO, cn comr*^{Z1340} *bw/ CyO, mia st/ TM6B* and *sa*^{1} *red/ TM3* stocks were used to generate homozygotes for comparison to other meiotic arrest loci (Ayyar et al., 2003; Jiang and White-Cooper, 2003; Lin et al., 1996). *topi*$^{Z3-2139}$, *topi*$^{Z3-0707}$ and *topi*$^{Z3-3767}$ were provided by C. Zuker from his collection of EMS-generated, viable lines, screened for male sterility by B. Wakimoto and D. Lindsley.

Microscopy and immunofluorescence

Live testes were dissected, squashed in 4 μg/ml Hoechst 33342 in testis buffer (183 mM KCl, 47 mM NaCl, 10 mM Tris pH 6.8) and examined by phase contrast and fluorescence microscopy. *topi* mutant testes contained many balls of cells, and no elongating cysts, hence the name *matotopetli*, which means 'balls' in the Aztec language, nahuatl. Aly and Comr proteins were visualized using a BioRad Radiance Plus confocal microscope, using rabbit anti-Aly or rabbit anti-Comr antibodies detected with FITC-conjugated secondary antibodies (Sigma), DNA was co-stained with propidium iodide (Jiang and White-Cooper, 2003; White-Cooper et al., 2000). Phase contrast images were captured using a Photometrics cooled CCD camera connected to a Zeiss Axiophot microscope, or a Q-imaging Retiga 1300 CCD camera linked to an Olympus BX50 microscope using IP Lab Spectrum or Openlab software (Improvision) and then imported into Photoshop. Images of RNA in situ hybridization to testes were captured on colour slide film, this was scanned and imported into Photoshop.

Northern blotting and in situ hybridization

Northern blotting of poly-A+ RNA from whole males, whole female, agametic males and embryos using *topi* and *rp49* probes was carried out as described previously (Hiller et al., 2001). RNA in situ hybridization was carried out as previously described (White-Cooper et al., 1998). Dig-labelled RNA probes were generated using dig-RNA labelling mix (Roche). Probes for *topi*, *cyclinA*, *cyclinB*, *twine*, *fzo* and *Mst87F* were generated by transcription off a linearized cDNA-clone plasmid. For *TrxT* and *CG8349* a T3 RNA polymerase promoter was included in the 3′ RT-PCR primer, and purified PCR product was used as the template for transcription of the labelled RNA probe.

Microarray analysis and RT-PCR

Full details of the microarray analysis will be presented elsewhere. Testes and seminal vesicles were dissected from 0- to 1-day-old males

homozygous for *red e* (control), *aly*^{5}, *can*^{3}, *comr*^{z1340}, *achi*^{Z3922}*vis*^{Z3922} and *topi*$^{Z3\text{-}2139}$ in testis buffer, and placed at –80°C within 30 minutes of dissection. For each genotype, three independent RNA samples were made from 200 testes. Testes were homogenized in Trizol (Roche), then shipped on dry ice to the BBSRC IGF facility in Glasgow for RNA extraction, labelling and Affymetrix array hybridization (http://www.mblab.gla.ac.uk/igf/index.html). Total RNA (6 μg) per sample was used for probe synthesis. Normalized signal intensities were averaged over the three replicates and compared between genotypes. Genes were selected for analysis from a list of genes expressed in wild-type testes where the mean signal for *topi* was eight times higher or lower than that for *achi/vis*. *P* values were calculated with a one tailed *t*-test on this pair of samples.

For RT-PCR, total RNA was extracted from dissected testes with Trizol reagent. The RNA was resuspended in RNAse free water at a concentration of three testes worth/μl. First-strand cDNA from 4 μl of this sample was generated using oligo-dT primers with the SuperScript II reverse transcriptase system (Invitrogen) following the manufacturer's instructions. As a negative control, wild-type testis RNA was used as the template and reverse transcriptase was omitted from the reaction. cDNA derived from 0.18 testes was used for each RT-PCR reaction and amplified with Taq DNA polymerase (Qiagen) with 24 amplification cycles. Genomic DNA from wild-type flies was used as a positive PCR control.

PCR primers (synthesized by MWG) were designed to amplify 400-1000 bp fragments from the transcript of the genes of interest. The T3 RNA polymerase promoter site, preceded by 6 bp of random sequence, was incorporated into the 3′ primer to facilitate production of RNA probes directly from the RT-PCR product. Primer sequences were: CG8349-5′ GCTCCTTCAGCGCTACATGC; CG8349-3′T3 GCAACGAATTAACCCTCACTAAAGGGCGCATAGGCACATCG; TrxT-5′ TCGGCGAGGGCAGAGCTC; TrxT-3′T3 GCAACGAATTAACCCTCACTAAAGGGCATTCTCGTCGTGGGC.

Cloning of *topi*

topi was mapped by recombination between *th* and *cu*, and further localized by deficiency complementation to polytene interval 85E9-13. The *topi* region was defined by the right breakpoints of *Df*(*3R*)*GB104*, which complemented *topi*, and *Df*(*3R*)*by10*, which failed to complement *topi*. This identified a 60 kb region of genomic DNA containing 23 predicted genes. We used representation of ESTs from different tissues as a crude guide to the expression pattern of these genes. Seventeen were not represented in the EST set derived from testis. Four had ESTs from both testis and other tissues, and two (*CG8484* and *CG8526*) had ESTs from only testis libraries. The premature stop codons in *topi*$^{Z3\text{-}2139}$ and *topi*$^{Z3\text{-}0707}$ and the mis-sense mutation in *topi*$^{Z3\text{-}3767}$ were identified by PCR amplifying the genomic region containing *CG8484* from homozygous males, and sequencing the bulk PCR product. Predicted Zn-finger motifs within Topi were identified by eye and using the InterPro analysis tool (http://www.ebi.ac.uk/InterProScan/). *topi* homologues from *Anopheles gambiae* (on AAAB01008944.1) and *Drosophila pseudoobscura* (on Contig815_Contig5737) were identified using Blast searches of the genome sequence databases.

Yeast two hybrid interaction screen

A Comr-Gal4-DNA-binding-domain fusion construct was made by subcloning the ORF from a full-length *comr* cDNA clone into the vector pGBKT7 using *Nde*I (at the start codon) and *Not*I. As no testis cDNA libraries suitable for two-hybrid screening were available, we generated and screened a testis cDNA-Gal4-Activation Domain (AD) fusion protein library by in vivo recombination using the Matchmaker library construction and screening kit (Clontech). Total RNA (1 μg), isolated from wild-type testes with Trizol, was used to synthesize first-strand cDNA using an oligo d(T) primer. Double-strand cDNA was synthesized with SMART III and CDS III anchors. The AD fusion library construction and two-hybrid screen were carried out in one step by co-transforming the yeast strain AH109 with ds cDNA, pGADT7-Rec and pGBKT7/comr. Colonies were picked from SD/-Ade/-His/-Leu/-Trp/X-α-Gal selection plates after 7 days. A screen of 10^6 independent co-transformants yielded 39 colonies that grew under selective conditions and were blue in the presence of X-α-Gal. AD/library plasmids were isolated from each positive yeast colony, transformed into *E. coli* and sequenced; one of the positive clones contained the full ORF of CG8484 – *topi*.

Deletion analysis plasmid construction

Determination of the protein interaction domains necessitated construction of *comr* and *topi* deletion series. PCR products for the deletion derivatives were subcloned into pGBKT7 using *Nde*I and *Not*I sites engineered into the primers. Co-transformation of HA109 cells was with pGADT7-RecTopi (full length Topi) or pGADT7-RecComr (full length Comr) as appropriate. A 622 bp fragment covering five zinc fingers in Kruppel was generated by PCR from wild-type genomic DNA.

Transfection, expression and immunoprecipitation in mammalian tissue culture cells

The Topi ORF was subcloned into HA-tagged pCDEF3 and the Comr ORF into FLAG-tagged pCDEF3 respectively at *Nde*I and *Not*I sites. 293T cells were cultured in DMEM with 10% FCS to 90% confluence then co-transfected with 10 μg HA-tagged Topi and Flag-tagged Comr using Lipofectamine 2000 reagent (Invitrogen). As a negative control, 293T cells were co-transfected with HA-tagged Topi and Flag-tagged human Smad1. Forty-eight hours after transfection, cells were collected and processed for immunoprecipitation (Bennett and Alphey, 2002), except that protein G sepharose incubation was carried out at 4°C overnight.

Results

Wild type function of *matotopetli* is required for male meiotic division and spermatid differentiation

Mutations in *matotopetli* (*topi*) block both progression of the meiotic cell cycle in males and onset of spermatid differentiation. Wild-type adult *Drosophila* testes contain a spatially organized temporal array of differentiating male germ cells (Fig. 1A) (reviewed by Fuller, 1993). A small population of germline stem cells resides at the testis apical tip. Division of a germline stem cell and associated somatic cyst progenitor cells gives one spermatogonial daughter, which is committed to differentiation and encapsulated by two somatically derived cyst cells. Four mitotic amplification divisions of the spermatogonium results in a cyst of 16 interconnected primary spermatocytes, which then enter an extended period of growth and gene expression followed by meiotic division to produce a cyst of 64 round spermatids. By a programme including extensive elongation and other morphological changes, the spermatids differentiate into mature sperm. Testes lacking *topi* function contained germ cell cysts up to and including mature primary spermatocytes, but lacked cells in meiotic division or spermatid differentiation stages, placing *topi* into the meiotic arrest class of genes (Fig. 1B). *topi* null flies were viable and female fertile. Three alleles of *topi* were identified by screening a collection of EMS induced, viable, male sterile mutations (generated by E. Koundakjian and C. Zuker, screened for male fertility by B. Wakimoto and D. Lindsley). *topi*$^{Z3\text{-}2139}$ was identical in phenotype to *topi*$^{Z3\text{-}0707}$. *topi*$^{Z3\text{-}3767}$ appeared to be a hypomorphic allele, as occasional progression of mature primary spermatocytes into the meiotic divisions was observed in homozygotes and hemizygotes (data not shown).

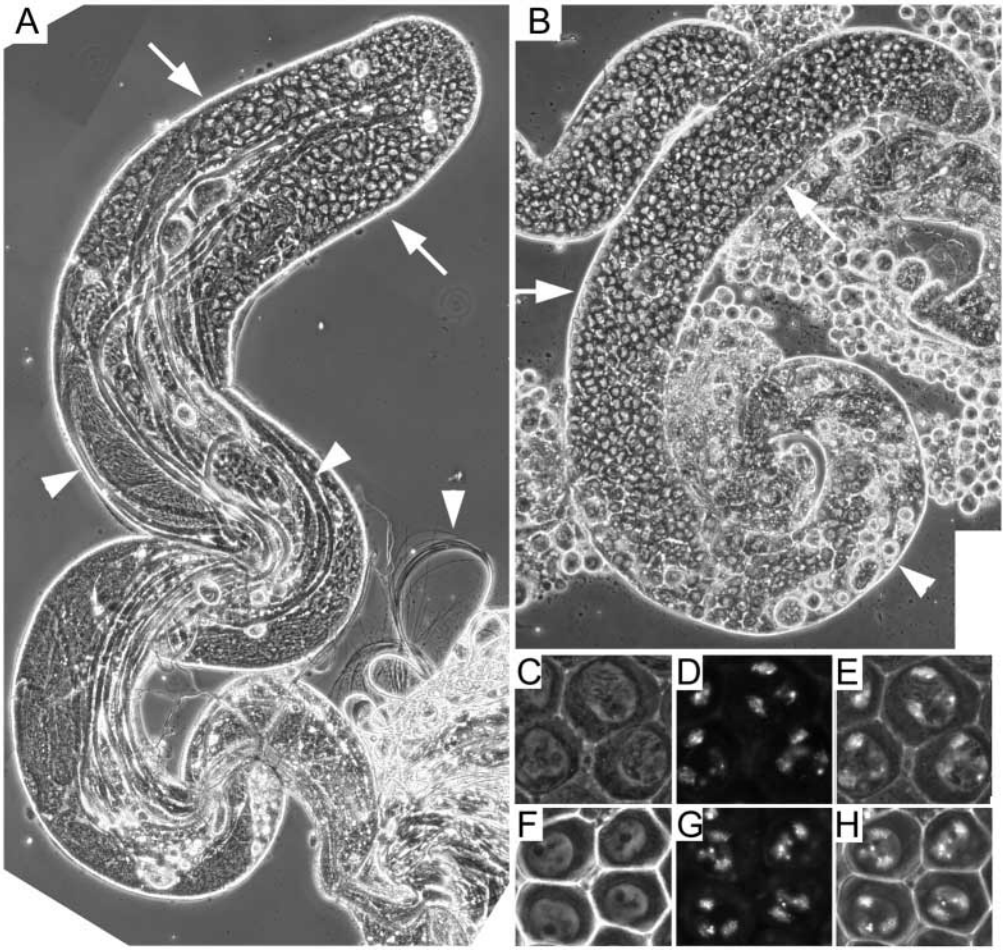

Fig. 1. The *topi* meiotic arrest phenotype. (A) Phase-contrast image of wild-type whole testis, apical tip in top right-hand corner. Primary spermatocytes occupy most of the apical end (arrows) Elongating spermatid bundles (arrowheads) are seen inside the testis, and spilling out from the distal end. (B) Phase-contrast image of $topi^{Z3\text{-}2139}$ whole testis. Only stages up to mature primary spermatocytes (arrows) are seen. Degenerating primary spermatocytes are present in the distal region of the testis (arrowheads). (C-H) High-magnification views of mature spermatocytes from squashed preparations of wild-type (C-E) and $topi^{Z3\text{-}2139}$ (F-H) testes viewed by phase-contrast (C,F) and Hoechst DNA staining (D,G) or both visualization methods simultaneously (E,H).

The mature spermatocytes present in $topi^{Z3\text{-}3767}$/*Df(3R)by10* or $topi^{Z3\text{-}3767}$ homozygous mutant testes resembled wild-type mature primary spermatocytes by phase contrast microscopy (Fig. 1C,F). However the chromatin morphology of mutant spermatocytes was subtly different from wild type when visualized by inclusion of Hoechst 33342 in the squash buffer. Mature wild-type primary spermatocytes have decondensed chromosomes, which are typically found as three chromatin domains close to the nuclear membrane (Fig. 1D,E). As primary spermatocytes pass through the G2-Meiosis I transition, the chromosomes condense and move away from the nuclear envelope. Chromosomes in *topi* arrested cells were somewhat more compact in appearance than in mature wild-type cells in late G2 (Fig. 1G,H), suggesting a block in the G2-M transition of meiosis I in *topi* mutant males. *topi* mutant chromatin most closely resembled that found in *achi/vis* mutant testes, and did not show the fuzzy, decondensed morphology typical of *aly* and *comr* (Ayyar et al., 2003; Jiang and White-Cooper, 2003).

matotopetli belongs to the *aly*-class of meiotic arrest genes

RNA in situ hybridization analysis of the effects of *topi* mutants on transcription of meiotic cell cycle and spermatid differentiation genes placed *topi* into the *aly* class of meiotic arrest mutants. The two strong alleles of *topi* showed greatly reduced levels of expression in primary spermatocytes of the meiotic cell cycle transcripts *twine*, *cyclin B* (Fig. 2C-F) and *boule* (not shown). The transcript levels of several spermatid differentiation genes, including *fzo* and *Mst87F*, were also dramatically reduced compared with wild-type levels (Fig. 2G-J). Some expression of *fzo* was detected in $topi^{Z3\text{-}3767}$ mutant testes, consistent with the weaker phenotype seen in this allele by phase-contrast microscopy (data not shown).

matotopetli function is not required for normal association of *Aly* and Comr proteins with chromatin in primary spermatocytes

Our phenotypic analysis indicated that *topi* is a new member of the *aly*-class of meiotic arrest genes, which includes *aly*, *comr* and *achi/vis*. All three of these proteins localize to nuclei of wild-type primary spermatocytes. The behaviour of Aly and Comr proteins in *achi/vis* mutant testes and *topi* mutant testes was extremely similar. Immunofluorescence analysis with anti-Aly and anti-Comr antibodies revealed that both Aly and Comr proteins localized to nuclei in spermatocytes from males carrying strong loss of function mutations in *topi* (Fig. 3). In wild-type primary spermatocytes, Aly and Comr proteins were concentrated on chromatin as well as being distributed throughout the nucleoplasm. By contrast, in *topi* mutant testes, Aly and Comr proteins did not appear to concentrate on chromatin, but were uniformly distributed within the nucleus.

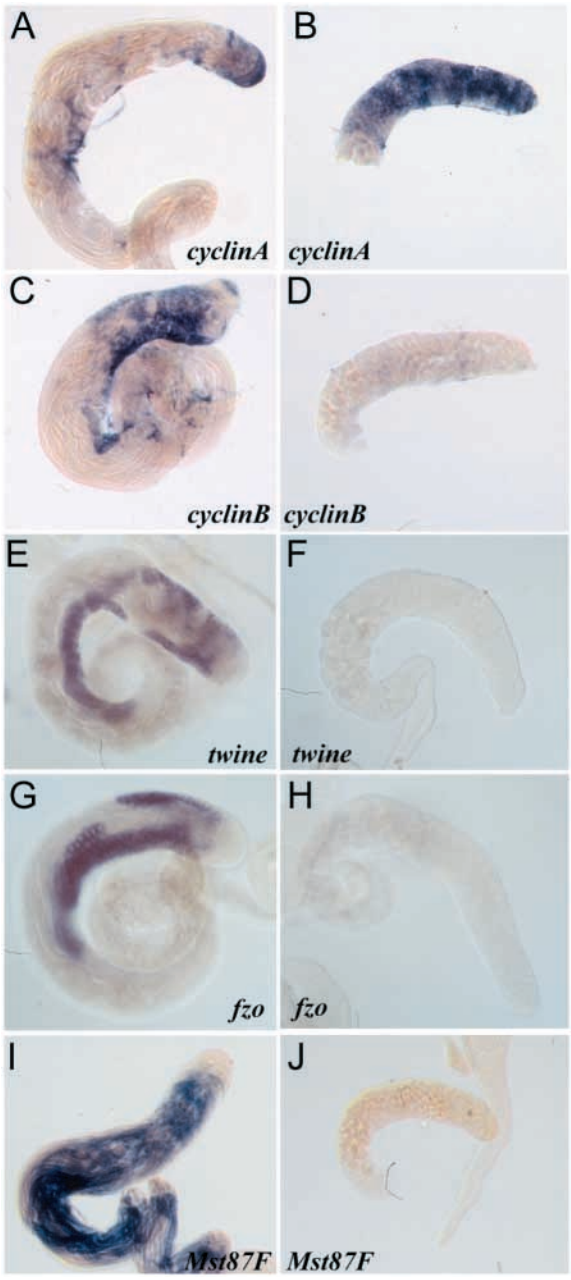

Fig. 2. In situ hybridization to whole testis. In situ hybridization to wild-type (A,C,E,G,I), $topi^{Z3\text{-}2139}$ (B,D,J) or $topi^{Z3\text{-}0707}/Df(3R)by10$ (F,H) mutant testes revealed that *topi* is an *aly*-class meiotic arrest gene. The distal ends of some mutant testes were broken off during the staining. Probes were anti-sense transcripts from *cyclinA* (A,B), *cyclinB* (C,D), *twine* (E,F), *fzo* (G,H) or *Mst87F* (I,J).

matotopetli encodes a Zn finger protein expressed specifically in primary spermatocytes

The *topi* gene was localized by a combination of recombination mapping and deficiency complementation (see Materials and methods). Sequence analysis of genomic DNA fragments amplified from $topi^{Z3\text{-}3767}/Df$ (*3R*)*by10*, $topi^{Z3\text{-}2139}$ and $topi^{Z3\text{-}0707}$ mutant animals by PCR revealed that each of the three *topi* alleles caused lesions in *CG8484*, identifying this predicted gene as *matotopetli*. $topi^{Z3\text{-}0707}$ and $topi^{Z3\text{-}2139}$ introduced stop codons, in the first third and in the last third of the predicted protein, respectively. The weaker allele ($topi^{Z3\text{-}3767}$) carried a mis-sense mutation.

topi encodes a 814 amino acid multiple Zn-finger predicted protein with a predicted molecular weight of 92 kDa (Fig. 4). A near full length cDNA isolated from testes (AT25463) was identified in the BDGP EST project and was fully sequenced. The conceptually translated Topi protein contains 10 C2H2 and one C2HC class Zn fingers predicted by the PFAM and prosite protein domain analysis tools. The predicted Zn fingers are clustered in the central region of the protein (amino acids 230-650), with the first two fingers being separated from the remaining nine by a 64 amino acid spacer. The mis-sense mutation in $topi^{Z3\text{-}3767}$ alters a conserved residue, G(516)D, just after the second Zn-binding cystine in the seventh predicted Zn finger. The non-sense mutation in $topi^{Z3\text{-}0707}$ would truncate the protein after the first predicted Zn finger (at amino acid 261). The non-sense mutation in $topi^{Z3\text{-}2139}$ would truncate the predicted Topi protein within the seventh predicted Zn finger (amino acid 521), potentially encoding a protein with six Zn fingers. Blast searches were used to identify putative *topi* orthologues, with sequence conservation extending beyond the Zn finger domains, in the mosquito *Anopheles gambiae* and in *Drosophila pseudoobscura*. Unlike both fly genes, the mosquito protein had only 10 predicted Zn fingers owing to lack of three crucial Zn-binding residues from the second Zn finger.

The *matotpetli* gene encodes a 2.7 kb transcript that was detected in northern blots of polyA+ RNA from whole adult males but not from adult males lacking a germ line, females or embryos (Fig. 5A). RNA in situ hybridization to whole testes revealed that *topi* transcript was expressed specifically in primary spermatocytes. *topi* transcript was not detected in spermatogonia or elongating spermatids (Fig. 5B). *topi* transcript was low or not detectable in spermatocytes from $topi^{Z3\text{-}0707}/Df$ males (Fig. 5C).

matotopetli protein physically interacts with Comr

A full-length clone of the *topi* predicted Zn-finger protein was identified independently in a yeast-two hybrid screen for testis cDNAs encoding proteins that interact with Comr (see Materials and methods). Full-length Comr did not interact in the two-hybrid system with a domain containing five C2H2 Zn fingers from the transcriptional repressor protein Kruppel, indicating that the Comr-Topi interaction is specific. The interaction between Topi and Comr proteins was confirmed using co-expression and immunoprecipitation from mammalian tissue culture cells (Fig. 6C). FLAG-tagged Comr and HA-tagged Topi were co-expressed in 293T tissue culture cells. HA-tagged Topi protein had an apparent M_r of 92 kDa, correlating well with the predicted protein size. Immunoprecipitation with anti-FLAG antibodies, followed by western blotting with anti-HA antibodies showed that Topi co-immunoprecipitated with Comr. The reciprocal experiment, immunoprecipitation with anti-HA antibodies and blotting with anti-FLAG, also showed co-immunoprecipitation of Comr and Topi. To control for non-specific interactions in the cultured cells we co-expressed HA-Topi with FLAG-Smad1. We were not able to co-immunoprecipitate this pair of proteins.

The domains of Comr and Topi responsible for the

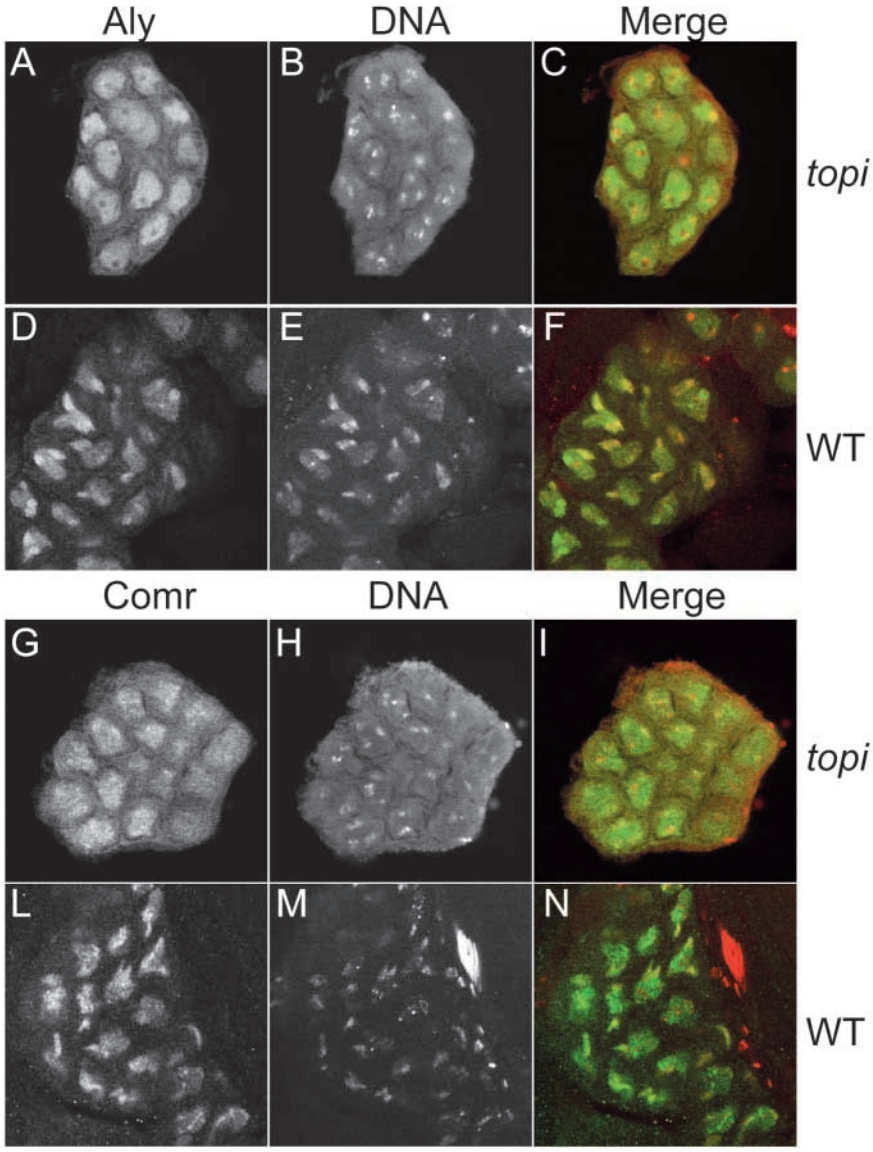

Fig. 3. Antibody stain of Aly and Comr in wild type versus topi. Immunofluorescence staining of Aly (A,D, green in merge) and DNA (B,E, red in merge) on wild-type (D-F) and $topi^{Z3\text{-}2139}$ (A-C) testes revealed that Aly protein is nuclear in both wild type and mutant, but that Aly fails to accumulate on the chromatin in $topi^{Z3\text{-}2139}$ mutant spermatocytes. Immunofluorescence staining of Comr (G,L, green in merge) and DNA (H,M, red in merge) on wild-type (L-N) and $topi^{Z3\text{-}2139}$ (G-I) testes similarly revealed that Comr protein is nuclear in both wild type and mutant, but that Comr fails to accumulate on the chromatin in $topi^{Z3\text{-}2139}$ mutant spermatocytes.

interaction were mapped by constructing N-terminal and C-terminal deletion series for each of the proteins and assaying interactions in yeast (Fig. 6A,B). Fragments of Comr containing residues 200-300 were capable of interacting with full-length Topi, whereas deletions lacking this region did not bind Topi. The Topi deletion series indicated that the Comr binding activity lies within the Zn-finger region. Specifically, fragments containing either the first two Zn fingers, fingers 7-10, or both, were capable of interacting with full-length Comr.

Differential requirements for *topi* and *achi/vis* for expression of some target genes

To explore whether *topi* and *achi*/*vis* might differ in their target gene specificities, we carried out a set of DNA microarray experiments to examine gene expression in wild-type and meiotic arrest mutant testes (see Materials and methods) (H.W.-C., unpublished). About 1000 genes were expressed at reduced levels (at least fourfold) in *aly*-class meiotic arrest mutants than in wild type, most of these transcriptional targets of the meiotic arrest genes depended equally on the functions of *aly*, *comr*, *topi* and *achi*/*vis*. However, transcription of a subset (about 42) of genes was much more dependent on *achi*/*vis* than on any of the other meiotic arrest loci. Transcription of a different subset (about 44) of genes seemed to be much more dependent on *topi*. The raw and log ratio array data for these genes are presented in Table S1 at http://dev.biologists.org/supplemental/, along with data for genes whose expression in wild-type and meiotic arrest mutant testes has already been examined by RNA in situ hybridization or northern blotting (White-Cooper et al., 1998). Further analysis of representative genes by RT-PCR from wild-type and mutant testes confirmed the *achi*/*vis* versus *topi* differential dependence.

CG8349 encodes a predicted cytidine deaminase. 11 testis ESTs for this gene have been sequenced, no *CG8349* ESTs have been sequenced from other tissue libraries, indicating that *CG8349* transcription is highly testis enriched. Microarray analysis indicated that expression of this gene was approximately five times lower in *aly*, *comr* and *topi* than in wild type, while expression of *CG8349* was reduced by two-hundred times in *achi/vis* compared with wild type. By RT-PCR, we found that *CG8349* transcript was detectable, but levels were dramatically reduced in *aly*, *comr*, *topi*, *mia* and *sa* mutant testes. However, no transcript was detected in *achi/vis* mutant testes (Fig. 7A). To further investigate these changes in gene expression, we used in situ hybridization against wild-type and mutant testes. The relative level of transcript can be compared between wild-type and mutant testes using this method as the tissues were mixed before hybridization. In situ hybridization revealed that *CG8349* was expressed in primary spermatocytes, and the transcript persisted into mid-late stages of spermatid elongation (Fig. 7B). As predicted by the RT-PCR and array analysis, the in situ hybridization signal intensity for *CG8349* was lower than wild type in *aly*, *comr* and *topi* testes (Fig. 7C,E), and the transcript was undetectable by in situ hybridization in *achi/vis* mutant testes (Fig. 7D).

The microarray analysis indicated that transcripts of *TrxT*, a testis specific isoform of thioredoxin (Svensson et al., 2003),

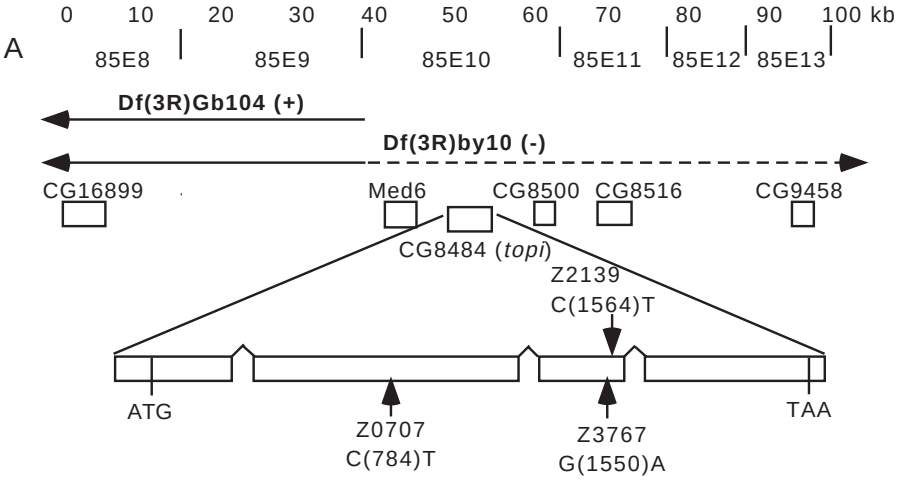

Fig. 4. Molecular cloning of *topi*. (A) *topi* genetic region was defined by the right breakpoints of Df(3R)Gb104 and Df(3R)by10, an interval of about 60 kb, which contains 23 predicted genes of which five are shown. *topi* gene structure and the location of lesions in mutant alleles are shown. (B) *topi* predicted amino acid sequence (mel) aligned with the putative *topi* orthologues from *D. pseudoobsura* (pseud) and *A. gambiae* (gamb). Sequence identity between all three proteins is indicated by black boxes, sequence similarity is indicated by grey boxes. Predicted Zn-finger motifs are indicated with lines above the sequences. The amino acids affected in the mutant alleles are indicated with an asterisk above the sequence.

B

```
mel     MKVKVSGEYTLAEIEVELDQQLTPDDLQPGATVLATNESTGGLERIVSHEELSRFFAVGP  60
pseud   ----VPGEYTLAEIEVGDKSQLNFSHL-PDATVLAPSEGNSQVKRVLGQEELVQLFNRPP  55
gamb    ------------------------------------------------------------

mel     AG--ALPMPTDVVVERTLADPAFKQILQEADGKKGFDPQAEQMKIRDFLAGVTSSKMTTE 118
pseud   PPPPIIQMSTEVVLQRTLEHPALKQILQEADGQPDFDPAAEQMKIREFLAGVTNSKMTTE 115
gamb    ------------------------------------------------------M-SG 10

mel     QSVFHGSRSNS----------SASTVNRI-KCPTCLVQFDAVAFQNHSCE----AKFVEV 163
pseud   ESVFQGRQLQS-LFPPAQRMSTGSTGSRIYKCRHCGGKYDAVYFSSHICTSEDVAKFGHE 174
gamb    NTLFSYWKFRR---------------------------------------------------- 21

mel     AVPQQEKFHLVPT---VSAPFAPLSKPASERVIRENQVRLRRYIKDEMKYDLATGIESSR 220
pseud   AFGYQDKFRALKTSMEMGTVFAPRYVRSTERIILENQVRLRRYMKDELKYDVTTGVDSSV 234
gamb    ------------------QAEETFRSSSRRLRQNIESLQCTLLQDFGTGIGAGG---S  58

mel     -KNAAKGPNECTMCDRKFVHASGLVRHMEKHALDLIP---SQTSEQPHTIP-AAGLHVVV 275
pseud   PRKAGKQINECSICDRQFVHFSGLVRHMEKHALNLIA---PQANAPSNPAPPAPGLAVVI 291
gamb    VRLASYKKYTCDQCLRKFVHESGLARHVAKFHPDVSTNVEPPTGQQQQHPELPATVCSDS 118

mel     KCNSCGRIFYDPQVAFRHGLIHDSEHSTMRQSPMTQVPSNRADFNELLLDGEML---IDN 332
pseud   KCNDCGRIFYDSKQAMEHARVHYPDQPKLYEPLTSENESEKKDFCKRMLKGDWLFLCVQS 351
gamb    KPTPASEAYIDAVVKCRHSGAIFP------SAEQYLCHAMRYKFWEQLFAAEPE------ 166

mel     DPAFATSNQNTNPPEKEMFSSLILGSVLQCEFCEYIFADIAELLVHSASHAAERRFECTA 392
pseud   KKRSVDQKQEKRAFEAGLSSTLILGSVLQCEFCEYVFADIEKLLVHSAAHVAERRFECTA 411
gamb    ------ERERPAVLLRRQFSIVAVHTVLRCEYCDACFSDCAGLFAHESQHSPTGGFWCTF 220

mel     CDIQMNTAKEASIHFQTDCIFMREAIRSLNVTLSRYFVCNVCELKFANTDLLQEHR---C 449
pseud   CGITMNTTKEASVHFQTECIFMREELKRLDVSLFRYFVCNVCEQKFSTADLLQEHR-SNC 470
gamb    CRLHFPTLTDVLAHR-AECLEGSSFIGCHFQNVQLQYCCNVCMALFATLPELYAHR---- 275

mel     TS--------FHYFPRLNENGK-------KLLLPCDFCDVNFEFAHDFLAHSEEKHLN-- 492
pseud   PSLQTPRYIYYHYFPRLTADGS-------TLVLPCEYCDANFTHAHEVEEHNQEKHLN-- 521
gamb    -------YDEIHIFPRRMARAKGQTGGYGGLWVSCELCVFVCDHIELLLLVHRAECHPGGA 328

mel     ----------------------KKKREKETRNTG----AGRIRQYLCDICGKSYTQSSH 525
pseud   ----------------------RKRREKETRTNGPSYGNGRLRQYLCDVCGKSYTQSSH 558
gamb    VTNKPPTMAAPSPSAGNPPKQRQQQQQQLAAPKRPEGEESSNERPYLCEKCGKTYTQSSH 388

mel     LWQHLRFHQG--------------------VKPFVCQEENCDRKFTIRPDLNDHIRKCH 564
pseud   LWQHLRFHQGKCQLCQRSPWPITSSPVGLTGVKPFACQEANCDRKFTIRPDLNDHIRKCH 618
gamb    LWQHLRFHNG--------------------VRPFACDRDGCSRRFTIRPDLSDHIRKCH 427

mel     TGERPYLCLVCGKRFLTGSVFYQHRLIHRGERRYECEECGKRFYRADALKNHQRIHTGEK 624
pseud   TGERPYHCLVCGKRFLTGSVFYQHRLIHRGERRYECDECGKRFYRADALKNHQRIHTGEK 678
gamb    TGERPYECNVCHKRFLTGSVYYQHRLIHRGERRYGCPDCDRRFYRADALKNHQRIHTGEK 487

mel     PYSCLFCTKTFRQRGDRDKHIRARHSHLDAN-SRLMMQMQKFQLETAAAQKAQSHNPEQQ 683
pseud   PYSCLFCTKTFRQRGDRDKHIRARHSHLDAN-SRLMMQMQKFQLESAQAAKARIRQGQMG 737
gamb    PYECGHCERKFRQRGDREKHVRVKHKNGGSPESDQPDSDSLFAIHSGADSDAEIDSDS-- 545

mel     DNDVAGGASTSDVPSGSGFMSTEPSVAEMQYSITPEQQEEMVCVPIDEVN-NSFFMSHYM 742
pseud   D-----GDSYSAEASGSHVGPEE----DGPNNFATEEQDGMYCDPMDGFTPNQFEIQQCV 788
gamb    ------------------------------------------------------------

mel     QAVP---MEEDGSGQHIIVFEQPGQNMDMMSIYDQQQVGEPMHESGVPKRFAEENARVVV 799
pseud   EEAEPTPMNED-VDQEIIVYESPEPTVTIS--YEPKPEPQQMKQE-IPVKFERPKRKIVV 844
gamb    ------------------------------------------------------------

mel     VKNNPTKPIFSDTYL 814
pseud   VKNNPSLPFYPNSYM 859
gamb    ---------------
```

were 2.5 times less abundant in *aly* and *comr* than in wild-type or *achi*/*vis* mutant testes. However, *TrxT* transcript was at least fifty times less abundant in *topi* mutant testes than in wild type. Consistent with these data, we were unable to RT-PCR *TrxT* from *topi* mutant testes, while the RT-PCR of *TrxT* revealed only a mild reduction in transcript levels compared with wild type in *aly*, *comr* and *achi*/*vis* mutant testes (Fig. 7F). No reduction in transcript level was evident by RT-PCR in *mia* or *sa* mutant testes. *TrxT* expression was first detected by in situ hybridization in primary spermatocytes in wild-type testes, the transcript persisted after the meiotic divisions and was detected until mid-elongation stages of spermiogenesis (Fig. 7G). *aly* and *comr* mutant testes showed expression in the primary spermatocytes; however, the signal intensity in these cells was lower than in the wild-type control (Fig. 7H). Thus, although significant expression of *TrxT* was detected in *aly* and *comr* testes, *aly* and *comr* are both required for full expression of *TrxT*. Signal intensity was similar to control levels in *achi/vis* mutant testes (Fig. 7I). *TrxT* transcript was undetectable in *topi* mutant testes by in situ hybridization (Fig. 7J), confirming that *topi* is absolutely required for expression of *TrxT*.

Discussion

A pathway for recruitment of the *aly*-class meiotic arrest proteins to chromatin

The *aly* class meiotic arrest genes of *Drosophila* form a functional module required for transcription in primary spermatocytes of a number of meiotic cell

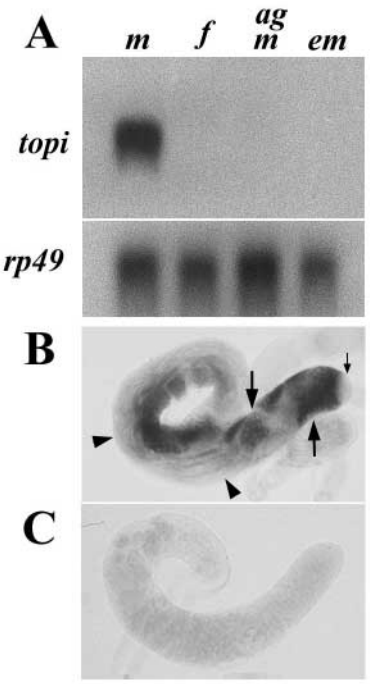

Fig. 5. *topi* is expressed specifically in primary spermatocytes. (A) Northern blot with *topi* (top) and *rp49* (bottom) probes. A 2.7 kb *topi* transcript was detected in RNA from whole males (m), but not from females (f), agametic males (ag m) or embryos (em). *rp49* was detected at similar levels in all the samples. (B) RNA in situ hybridization with *topi* probe to wild-type testis reveals expression in primary spermatocytes (arrows) and absent from spermatogonia (small arrow) and spermatids (arrowheads). (C) RNA in situ hybridization with *topi* probe on $topi^{Z3\text{-}0707}/Df(3R)by10$ mutant testis indicates lack of *topi* transcript.

cycle regulatory components and spermatid differentiation genes. *aly* encodes a homologue of the *C. elegans* SynMuvB gene *lin-9*, which is thought to be involved in recruiting the NURD nucleosome remodelling and histone deacetylase complex to repress target genes involved in vulval development (Beitel et al., 2000; White-Cooper et al., 2000). The *Drosophila* Aly protein is found in a complex with Comr, a novel protein that also associates with chromatin in *Drosophila* primary spermatocytes (Jiang and White-Cooper, 2003; Wang and Mann, 2003) (J.J. and H.W.-C., unpublished). A pathway by which Aly and Comr are assembled onto chromatin in primary spermatocytes is now beginning to emerge (Fig. 8). Interaction between Aly and Comr is required for each of the two proteins to localize to the nucleus, because absence of Aly leads to cytoplasmic accumulation of Comr and vice versa (Jiang and White-Cooper, 2003). Once the Aly/Comr complex has entered the nucleus, wild-type function of the predicted DNA-binding proteins Topi and Achi/Vis appears to be required for concentration of the Aly/Comr complex on chromatin. The Aly/Comr complex interacts structurally with Achi/Vis in primary spermatocytes (Wang and Mann, 2003) and Topi directly binds Comr. We propose that the Zn-finger protein Topi, via its physical interaction with Comr, and the TALE-class homeodomain proteins Achi/Vis, via their physical interaction with Aly/Comr, recruit an Aly/Comr complex to chromatin. We propose that a typical target gene, e.g. *dj*, has both *topi* and *achi/vis* target sequences in its promoter. Topi and Achi/Vis would bind to these regions, either independently or co-operatively, and recruit a pre-formed Aly/Comr complex from the nucleoplasm onto the promoter. Both Topi and Achi/Vis would need to be present to efficiently recruit Aly/Comr. Alternatively, Topi and Achi/Vis could bind to an Aly/Comr complex in the nucleus, with the target sequence for either DNA-binding protein sufficient to localize the entire assemblage to the target promoter. The *aly*-class meiotic arrest gene *topi* plays a crucial role in function of this regulatory module required for activation of a novel transcriptional programme in primary spermatocytes. Loss of function of *topi* results in the arrest of mutant cells as mature primary spermatocytes, and wild-type function of *topi* is required to generate a store of transcripts to support meiosis and the post-meiotic morphological changes that occur during spermatid differentiation.

Structure of target promoters

Although the meiotic arrest genes are crucial for transcription of a large suite of primary spermatocyte specific genes, they are not global regulators, as many transcripts accumulate to normal levels in the primary spermatocytes of mutant testes. Cis-acting sequences at target promoters would make them dependent on the function of *aly* and *comr*. Targeting of the Aly/Comr complex activity to particular sites on the genome may be crucial for understanding the molecular mechanism of this genetic module. The identification of two potential DNA-binding activities within the *aly*-class of meiotic arrest genes raises questions about the requirements for each of these proteins at different target gene promoters. In our microarray analysis, all genes we found to depend on *aly* or *comr* were also dependent on *topi* and/or *achi/vis*. The overwhelming majority of *topi*-dependent genes are also dependent on *aly*, *comr* and *achi/vis* for their full activation; genes in this class include *Mst87F*, *dj* and *fzo*. This indicates that both Topi and Achi/Vis are usually needed at the promoters of meiotic arrest target genes for full activity. Many transcription factors function as homo- or heterodimers, and many developmentally regulated promoters have binding sites for more than one transcription factor. These characteristics allow a combinatorial mechanism to ensure spatiotemporal accuracy of the transcription of any particular gene. The simplest explanation for the requirement for both Topi and Achi/Vis is that the promoters of the majority of terminal differentiation genes have binding sites for both factors. *achi* and *vis* are expressed at many stages of development, yet their testis target genes remain silent in these other tissues. This silence is likely to be due, at least in part, to the lack of *topi* expression in other tissues. Thus, the cooperation between these two transcription factors is essential for the developmental accuracy of expression of genes required for spermatogenesis.

It is intriguing that transcription of some genes (including *TrxT* and *CG8349*) showed a strong requirement for *topi* but not *achi*/*vis* (or vice versa), and that expression of these genes was less strongly dependent on the activity of the other meiotic arrest genes. This indicates that, at least in these contexts, the binding of the Achi/Vis and Topi transcription factors to DNA is independent. Another feature these genes have in common is that they are members of clusters of related genes, with different expression patterns. *TrxT* is found adjacent in the

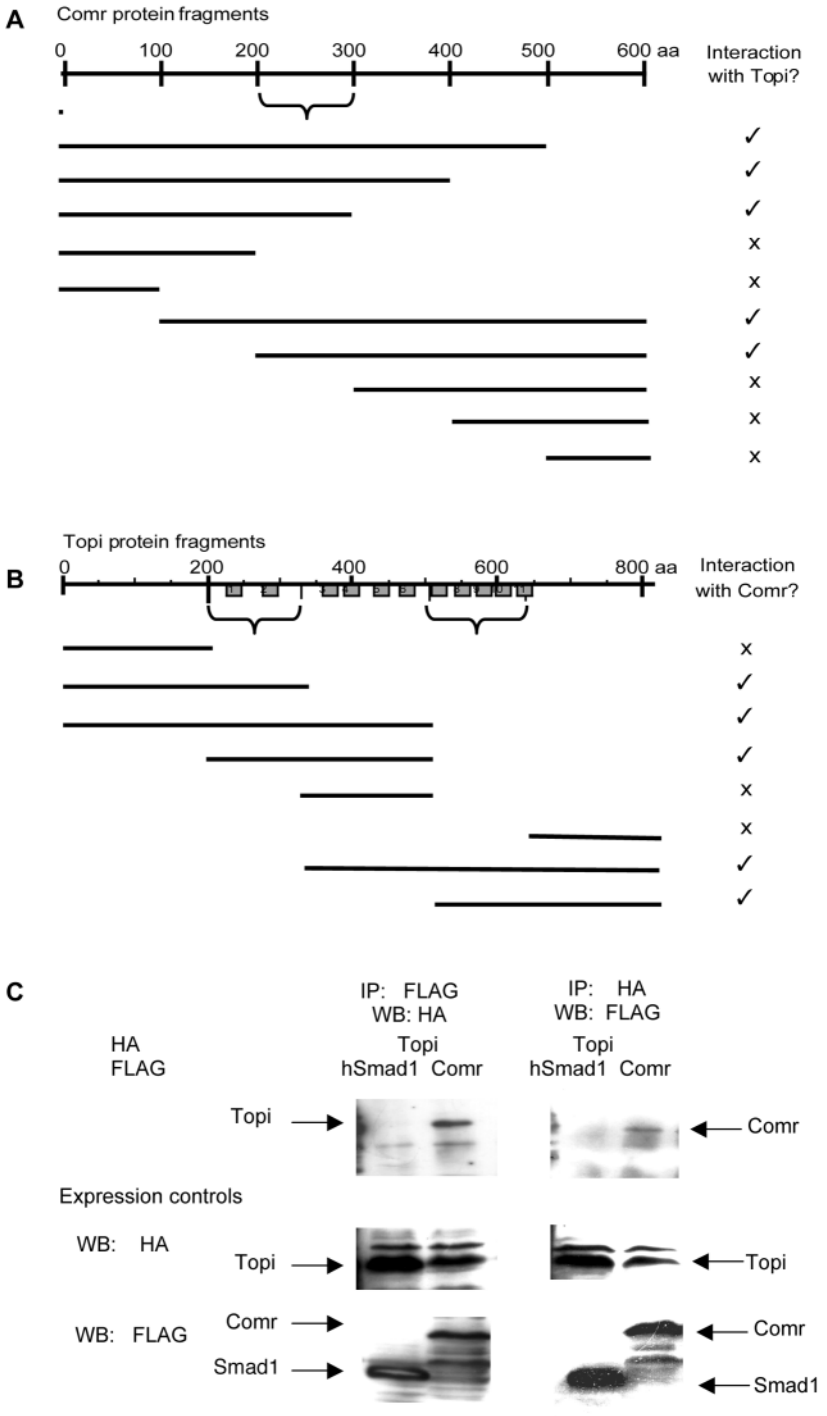

Fig. 6. Structural interaction between Topi and Comr proteins. (A) Mapping domains of Comr responsible for the interaction with full-length Topi. Solid bars represent protein fragments tested in the yeast two-hybrid system against Topi. The region responsible for the interaction lies in the 200-300 amno acid fragment (bracket). (B) Mapping domains of Topi responsible for the interaction with full-length Comr. Zn-finger motifs within Topi are indicated with boxes. Fragments containing either the first two, or the final five, Zn fingers can bind Comr (brackets). (C) Co-immunoprecipitation of Comr and Topi from tissue culture. Cells expressed HA-tagged Topi and FLAG-tagged Comr or Smad1. Binding was assessed by IP-ing with anti-FLAG, and blotting with anti-HA (first pair of lanes) or vice versa (second pair of lanes). Topi and Comr co-immunoprecipitate in both cases (top). Total protein expression was assessed by western blotting of the cell lysate (bottom).

genome to the female specific thioredoxin, *deadhead* (Svensson et al., 2003). *CG8349* is the 3′-most of a cluster of three predicted cytidine deaminase genes (the others being *CG8360* and *CG8353*). The structure of these promoters may differ from the canonical structure for meiotic arrest target genes, perhaps by having several binding sites for only one of the *aly*-class transcription factors. In this situation partial activation of gene expression would be achieved on binding of one transcription factor. Further recruitment of Aly and Comr (and usually the other DNA-binding factor) would be needed for the robust expression levels seen in wild type.

Functional domains within the Topi protein

The meiotic arrest gene *topi* encodes a protein with multiple Zn fingers. The final five Zn-fingers of Topi form an evolutionarily constrained unit, which we suggest may function via sequence specific DNA binding activity. *topi* has a large number of transcriptional targets (H.W.-C., unpublished). If, as is likely, Topi binds directly to *cis* acting regulatory elements associated with target genes in a sequence-specific manner, we would expect the protein sequence in the DNA-binding domain to be evolutionarily constrained. Changes in the affinity of Topi for any particular DNA sequence would have to be compensated for by changes in the sequence of a great many promoters, and are therefore likely to be selected against. Conservation of *topi* protein sequences within Diptera is extremely high across the final five Zn-finger motifs, and much lower across the first six Zn-fingers and at the termini. This conservation is particularly apparent in the loop regions of the final five Zn fingers, which are likely to be exposed, and important for interaction with DNA. Additionally these final five Zn finger motifs are more similar to those found in many transcription factors from other taxa. $topi^{Z3\text{-}2139}$ has a non-sense mutation that would truncate the protein before this predicted DNA binding region, and would presumably therefore be unable to bind DNA. The mis-sense mutation in the hypomorphic allele $topi^{Z3\text{-}3767}$ may reduce the affinity of the protein for its target site, but not totally abolish DNA-binding activity, thus allowing some residual function.

We propose that the first Zn finger of Topi is most important for the binding to Comr. Zn-finger motifs were first identified as nucleic acid binding domains. However Zn fingers are also known to be important for some protein-protein interactions (Evans and Hollenberg, 1988; Tsai and Reed, 1998). There is precedent among known Zn-finger transcription factors, as we

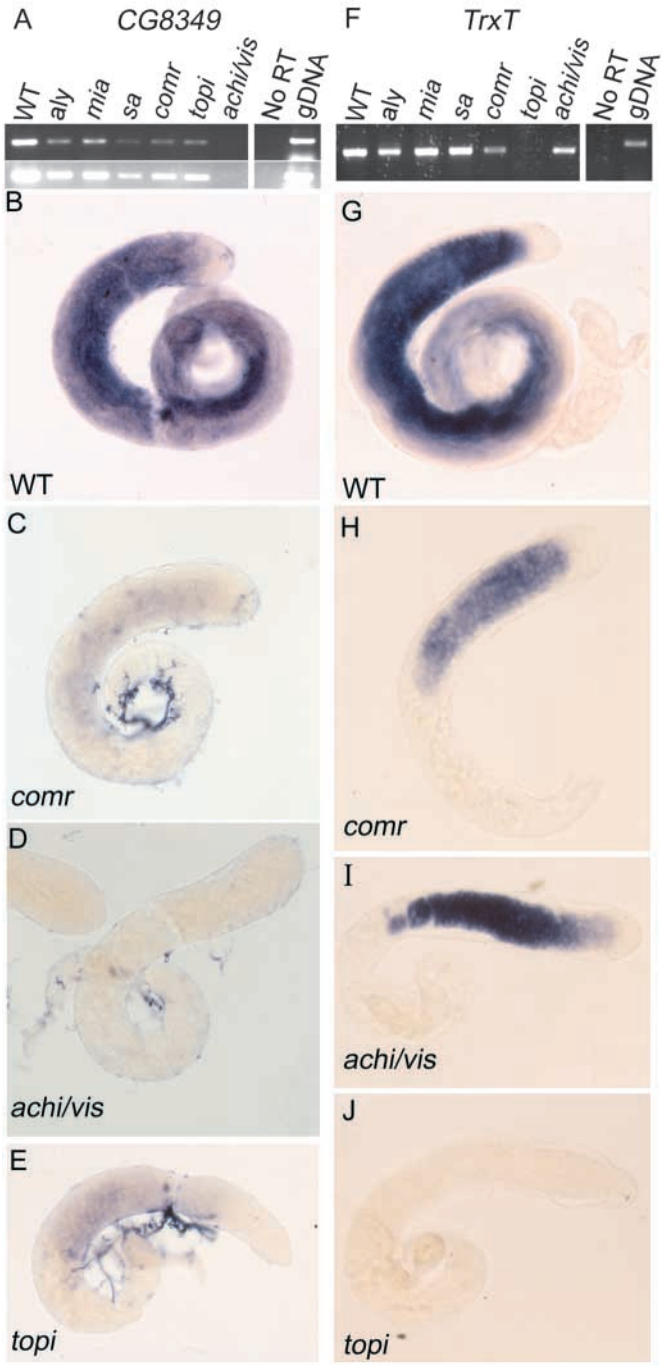

Fig. 7. Differential requirement for *topi* and *achi*/*vis* in expression of some target genes. (A) RT-PCR of *CG8349* from wild-type and meiotic arrest mutant testes. No-RT is a negative control, gDNA is PCR on genomic DNA from wild-type flies. Two exposures of the same gel are shown to emphasize the differences between the mutants and wild type (top), and the lack of product in *achi/vis* (bottom). (B-E) In situ hybridization of *CG8349* probe to testes revealed strong expression in wild-type (B) primary spermatocytes, persisting until mid-late elongation stages, weaker expression in *comr* (C) and *topi* (E), and no detectable signal in *achi/vis* (D) testes. (F) RT-PCR of *TrxT* from wild-type and meiotic arrest mutant testes. No-RT is a negative control, gDNA is PCR on genomic DNA from wild-type flies. (G-J) In situ hybridization of *TrxT* probe testes revealed strong expression in wild-type (G) primary spermatocytes and early-mid elongation stages. The signal in *achi/vis* testes was of a similar intensity to wild type (I). Weaker expression was seen in *comr* (H) and there was no detectable signal in *topi* (J) testes.

propose for Topi, for certain of the fingers to bind proteins, while others contact the DNA, allowing the transcription factor to target a multi-subunit complex to specific promoters (Kalenik et al., 1997). The Topi protein interacts structurally with Comr, and the region of Topi containing the first two Zn fingers was sufficient for interaction with Comr in yeast two hybrid assays. We propose that the first Zn finger is more important for Comr binding because the second of the two Zn fingers in mosquito Topi lacks crucial Zn-binding residues, so is unlikely to form a conventional finger structure in the folded protein. The final five Zn fingers, which we propose function primarily as DNA-binding motifs, also showed some Comr-binding activity in the yeast two-hybrid system. This interaction may stabilize, or increase the binding affinity of, the Topi-Comr complex. Strikingly, the N-terminal Comr binding region of *topi* is much less well conserved than the last five Zn finger containing domain, even among Diptera (*D. melanogaster*, *D. pseudoobscura*, *A. gambiae*). Also apparent is the relatively low conservation within the loop regions of the Zn fingers. If *comr* is rapidly evolving, the Comr interaction domains within Comr-binding proteins would also be expected to show low levels of conservation. It is possible that the N-terminal Zn-finger Comr-binding domain of Topi has evolved rapidly in concert with the rapid evolution of Comr. Searches of sequence databases had previously indicated failed to identify a homologue of *D. melanogaster comr* in the genome sequence of *A. gambiae*, which diverged from Drosophila 250 Mya (Zdobnov et al., 2002). Partial sequence of a *comr* homologue (H.W.-C., unpublished) was found in the set of unassembled genome sequence contigs from another Drosophilid, *D. pseudoobscura*, which diverged from *D. melanogaster* 46 Mya (Bergman et al., 2002). The conservation with *D. melanogaster comr* (31% identical, 52% similar) was remarkably low considering the relatedness of the species. This relative lack of conservation was apparent even in the Topi-binding region.

Transcriptional activation by the *aly*-class genes

The exact mechanism by which the *aly* homologue *lin-9* functions is not well understood. However, some insight has been provided by the molecular analysis of other genes in the *C. elegans* Syn-MuvB pathway, including *lin-35* (Rb, Retinoblastoma), *lin-53* (RbAp48) and *hda-1* (histone deacetylase) (Lu and Horvitz, 1998). Other components of the NURD histone deacetylase and chromatin remodelling complex also have SynMuvB activity (Solari and Ahringer, 2000). The biochemical function of NURD is to de-acetylate histones and re-position nucleosomes (Xue et al., 1998; Zhang et al., 1998). Typically, histone de-acetylation is associated with transcriptional repression, although exceptions to this rule have been found (De Rubertis et al., 1996; Struhl, 1998). Another gene in the SynMuvB pathway, *lin-13*, has been shown to encode a putative Rb-interacting protein with multiple C2H2 Zn-fingers, which may be important for targeting Rb and the NURD complex to specific sites

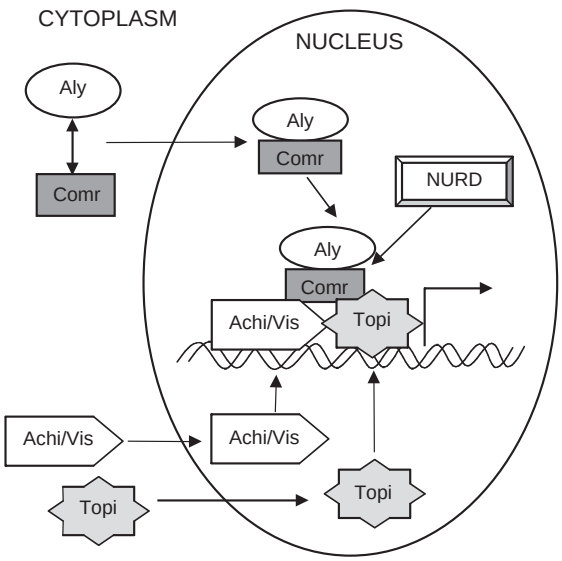

Fig. 8. Model for assembly and function of an *aly*-class protein complex at target promoters. Aly and Comr may be imported into the nucleus together or may only form a stable complex in the nucleus. Either protein alone remains in the cytoplasm. Topi and Achi/Vis proteins would enter the nucleus independently. These putative DNA-binding proteins would recognize specific sequences at target promoters, and once bound they would together recruit the Aly/Comr complex to these promoters. Alternatively, Topi and Achi/Vis could assemble with Aly/Comr in the nucleus, and this entire complex would then bind target promoters. This complex would then recruit the NURD histone deacetylase chromatin remodelling complex and promote transcription.

(Melendez and Greenwald, 2000). The loop regions of the Zn fingers in LIN-13 showed no significant similarity with those in Topi, however the two proteins could be playing similar roles in linking *lin-9* and NURD to target promoters. Analogous to the *C. elegans* system, we have proposed that *aly*, and by extension other genes in the same pathway as *aly*, functions through a NURD-complex chromatin-remodelling activity. Many genes require the *aly*-class meiotic arrest genes for activation of expression in primary spermatocytes, while a few may require these genes for transcriptional repression (H.W.-C., unpublished). Several SynMuvB genes have also been shown to be important for expression of transgenes in repetitive extrachromosomal arrays in *C. elegans*, indicating that, at least in certain specific contexts, the SynMuvB pathway activates gene expression (Hsieh et al., 1999). We favour a model in which the putative sequence-specific DNA-binding proteins, Topi and Achi/Vis target Aly and Comr to specific promoters, where Aly/Comr may then recruit a NURD complex to alter local chromatin structure. In the second part of the model, transcription of specific terminal differentiation genes in primary spermatocytes in some way requires the altered chromatin structure set up by the function of the Aly/Comr complex. The *can*-class meiotic arrest genes appear to form a different functional module that is also required for normal levels of transcription of many of the same target genes that require *aly*/comr. So far, all of the *can* class meiotic arrest genes molecularly identified encode homologues of subunits of TFIID. A TFIID complex containing the testis specific TAFs encoded by the *can*-class meiotic arrest genes may have a higher affinity for the altered chromatin conformation than conventional TFIID, and thus be required for full transcriptional activity.

We thank James Wakefield and Daimark Bennett for critical reading of the manuscript; and Luke Alphey, Myles Axton and members of the Fuller and White-Cooper laboratories for helpful discussions throughout this work. We thank Cricket Wood for help with the recombination mapping of *topi*. We are indebted to Barbara Wakimoto, Dan Lindsley and Charles Zuker for the *topi* alleles. L.P. thanks Alberto Darszon and Mario Zurita for allowing some of the experiments to be carried out in their respective laboratories. The microarray analysis was conducted by the BBSRC funded *Drosophila* IGF facility in Glasgow – thanks to Steve Russell, Julian Dow and Jing Wang. This work was supported by a Stanford Medical School Dean's postdoctoral fellowship and a CONACyT Fellowship (990544) to L.P., an NIH grant (1R01 1HD32936) to M.T.F., and Royal Society, MRC and Wellcome Trust grants to H.W.-C.

References

Aoyagi, N. and Wassarman, D. A. (2000). Genes encoding *Drosophila melanogaster* RNA polymerase II general transcription factors: diversity in TFIIA and TFIID components contributes to gene-specific transcriptional regulation. *J. Cell Biol.* **150**, F45–F49.

Ayyar, S., Jiang, J., Collu, A., White-Cooper, H. and White, R. (2003). *Drosophila* TGIF is essential for developmentally regulated transcription in spermatogenesis. *Development* **130**, 2841-2852.

Beitel, G. J., Lambie, E. J. and Horvitz, H. R. (2000). The *C. elegans* gene *lin-9*, which acts in an Rb-related pathway, is required for gonadal sheath cell development and encodes a novel protein. *Gene* **254**, 253-263.

Bennett, D. and Alphey, L. (2002). PP1 binds Sara and negatively regulates Dpp signaling in *Drosophila melanogaster*. *Nat. Genet.* **31**, 419-423.

Bergman, C., Pfeiffer, B., Rincón-Limas, D., Hoskins, R., Gnirke, A., Mungall, C., Wang, A., Kronmiller, B., Pacleb, J., Park, S. et al. (2002). Assessing the impact of comparative genomic sequence data on the functional annotation of the *Drosophila* genome. *Genome Res.* **3**, research0086.1-0086.20.

De Rubertis, F., Kadosh, D., Henchoz, S., Pauli, D., Reuter, G., Struhl, K. and Spierer, P. (1996). The histone deacetylase RPD3 counteracts genomic silencing in *Drosophila* and yeast. *Nature* **384**, 589-591.

Evans, R. and Hollenberg, S. (1988). Zinc fingers: gilt by association. *Cell* **52**, 1-3.

Ferguson, E. L. and Horvitz, H. R. (1989). The multivulva phenotype of certain *Caenorabditis elegans* mutants results from defects in two functionally redundant pathways. *Genetics* **123**, 109-121.

FlyBase (1999). The FlyBase database of the *Drosophila* genome projects and community literature. *Nucleic Acids Res.* **27**, 85-88.

Fuller, M. T. (1993). Spermatogenesis. In *The Development of Drosophila*, Vol. 1 (ed. M. Bate and A. Martinez-Arias), pp. 71-147. Cold Spring Harbor, New York: Cold Spring Harbor Press.

Hiller, M. A., Lin, T.-Y., Wood, C. and Fuller, M. T. (2001). Developmental regulation of transcription by a tissue-specific TAF homolog. *Genes Dev.* **15**, 1021-1030.

Hochheimer, A. and Tjian, R. (2003). Diversified transcription initiation complexes expand promoter selectivity and tissue-specific gene expression. *Genes Dev.* **17**, 1309-1320.

Hsieh, J., Liu, J., Kostas, S. A., Chang, C., Sternberg, P. W. and Fire, A. (1999). The RING finger / B-Box factor TAM-1 and a retinoblastoma-like protein LIN-35 modulate context-dependent gene silencing in *Caenorhabditis elegans*. *Genes Dev.* **13**, 2958-2970.

Jiang, J. and White-Cooper, H. (2003). Transcriptional activation in *Drosophila* spermatogenesis involves the mutually dependent function of aly and a novel meiotic arrest gene cookie monster. *Development* **130**, 563-573.

Kalenik, J. L., Chen, D., Bradley, M. E., Chen, S. J. and Lee, T. C. (1997). Yeast two-hybrid cloning of a novel zinc finger protein that interacts with the multifunctional transcription factor YY1. *Nucleic Acids Res.* **25**, 843-849.

Lin, T.-Y., Viswanathan, S., Wood, C., Wilson, P. G., Wolf, N. and Fuller,

M. T. (1996). Coordinate developmental control of the meiotic cell cycle and spermatid differentiation in *Drosophila* males. *Development* **122**, 1331-1341.

Lu, X. and Horvitz, R. H. (1998). *lin-35* and *lin-53*, two gens that antagonize a *C. elegans* Ras pathway, encode proteins similar to Rb and its binding protein RbAp48. *Cell* **95**, 981-991.

Melendez, A. and Greenwald, I. (2000). *Caenorhabditis elegans lin-13*, a member of the LIN-35 Rb class of genes involved in vulval development, encodes a protein with Zinc fingers and an LXCXE motif. *Genetics* **155**, 1127-1137.

Olivieri, G. and Olivieri, A. (1965). Autoradiographic study of nucleic acid synthesis during spermatogenesis in *Drosophila melanogaster*. *Mutat. Res.* **2**, 366-380.

Schafer, M., Nayernia, K., Engel, W. and Schafer, U. (1995). Translational control in spermatogenesis. *Dev. Biol.* **172**, 344-352.

Solari, F. and Ahringer, J. (2000). NURD-complex genes antagonise Ras-inducted vulval development in *Caenorhabditis elegans*. *Curr. Biol.* **10**, 223-226.

Struhl, K. (1998). Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev.* **12**, 599-606.

Svensson, M., Chen, J., Pirrotta, V. and Larsson, J. (2003). The ThioredoxinT and deadhead gene pair encode testis- and ovary-specific thioredoxins in *Drosophila melanogaster*. *Chromosoma* **112**, 133-143.

Tsai, R. and Reed, R. (1998). Identification of DNA recognition sequences and protein interaction domains of the multiple-Zn-finger protein Roaz. *Mol. Cell. Biol.* **18**, 6447-6456.

Wang, Z. and Mann, R. S. (2003). Requirement for two nearly identical TGIF-related homeobox genes in *Drosophila* spermatogensis. *Development* **130**, 2853-2865.

White-Cooper, H., Schafer, M. A., Alphey, L. S. and Fuller, M. T. (1998). Transcriptional and post-transcriptional control mechanisms coordinate the onset of spermatid differentiation with meiosis I in *Drosophila*. *Development* **125**, 125-134.

White-Cooper, H., Leroy, D., MacQueen, A. and Fuller, M. T. (2000). Transcription of meiotic cell cycle and terminal differentiation genes depends on a conserved chromatin associated protein, whose nuclear localisation is regulated. *Development* **127**, 5463-5473.

Xue, Y., Wong, J., Moreno, G. T., Young, M. K., Cote, J. and Wang, W. (1998). NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. *Mol. Cell* **2**, 851-861.

Zdobnov, E., von Mering, C., Letunic, I., Torrents, D., Suyama, M., Copley, R., Christophides, G., Thomasova, D., Holt, R., Subramanian, G. et al. (2002). Comparative genome and proteome analysis of *Anopheles gambiae* and *Drosophila melanogaster*. *Science* **298**, 149-159.

Zhang, Y., Leroy, G., Seelig, H. P., Lane, W. S. and Reinberg, D. (1998). The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. *Cell* **95**, 279-289.

Table S1A. Genes more dependent on *topi* than *achi/vis*

Gene	Function	topi-v-achi/vis	aly-v-wild type	comr-v-wild type	topi-v-wild type	achi/vis-v-wild type	can-v-wild type	Wild-type rank	Map
CG11591	Unknown	-4.61	-3.80	-2.98	-7.70	-3.09	-1.96	8	64A
CG4439	Leucyl aminopeptidase	-3.46	-1.73	-1.97	-7.17	-3.70	-2.33	46	53C
CG7046	DNA binding	-3.48	-0.67	-0.88	-5.15	-1.66	-1.07	50	94B
CG3213	Cytoskeleton	-6.55	-10.02	-10.15	-10.21	-3.66	-2.39	74	23F
CG5017	Nucleosome assembly	-3.18	-4.24	-5.27	-6.55	-3.37	-1.74	96	98B
CG7164	Cytoskeleton	-3.77	-8.07	-7.95	-7.48	-3.72	-2.40	170	28C
TrxT	**Thioredoxin**	**-6.06**	**-1.37**	**-1.33**	**-6.05**	**0.02**	**0.61**	**174**	**4F**
Menl-2	Malate dehydrogenase	-3.01	-5.31	-5.28	-5.80	-2.79	-2.03	272	53C
CG6790	Unknown	-3.28	-2.80	-1.92	-5.73	-2.45	-0.38	320	86E
CG15177	Unknown	-3.45	-6.08	-5.84	-6.41	-2.96	-1.10	342	84A
CG11601	Unknown	-3.13	-0.93	-0.96	-3.94	-0.81	-0.65	399	21B
CG9875	5-Formyltetrahydrofolate cyclo-ligase	-4.21	-5.78	-6.78	-7.00	-2.78	-1.76	424	59D
CG31473	Pyridoxamine-phosphate oxidase	-3.52	-1.76	-3.37	-6.69	-3.17	-1.47	510	84E
CG15742	Cell adhesion	-4.26	-0.42	-1.33	-6.19	-1.93	-0.31	521	11E
CG4198	Cytoskeleton	-4.10	-2.15	-2.46	-7.23	-3.14	-1.24	537	4F
CG13186	Unknown	-4.62	-6.70	-6.15	-7.14	-2.51	-3.24	706	48E
CG3994	Metal ion transporter	-3.21	-3.52	-2.45	-5.66	-2.44	-0.37	731	35C
CG3121	Cytoskeleton	-4.87	-1.86	-4.91	-7.33	-2.46	-0.92	809	60B
CG11068@	Unknown	-4.37	-2.81	-2.38	-7.20	-2.83	-1.12	1090	12C
CG11700$	DNA metabolism	-7.24	0.70	0.92	-4.79	2.45	2.94	1342	5E
CG16781	Unknown	-4.73	-1.32	-1.71	-5.68	-0.95	-0.28	1563	3D
CG10881	Translational initiation	-4.44	-6.79	-7.11	-6.57	-2.13	-0.53	1639	92E
I-t@	Protein phosphatase inhibitor	-4.27	-5.19	-3.63	-6.48	-2.21	-0.92	1826	86E
CG14708	Defense response	-5.37	0.43	0.70	-6.06	-0.69	0.99	1829	86E
CG11700$		-8.96	0.79	1.02	-6.35	2.61	3.23	1866	5E
CG13829	Oxidative phosphorylation	-3.65	-3.33	-2.98	-5.06	-1.41	0.24	1964	94E
CG6301	Cytoskeleton	-3.68	0.24	-1.46	-4.84	-1.17	-1.44	2122	53D
CG14811	Unknown	-3.48	2.07	1.44	-1.80	1.68	0.40	2124	2B
CG18472	Chaperone	-3.89	-2.89	-0.79	-5.49	-1.60	0.08	2364	97D
CG15296	Unknown	-3.57	-5.78	-5.24	-5.31	-1.74	-0.55	2416	9D
CG10459	Receptor kinase	-4.69	0.23	-0.36	-5.93	-1.24	0.06	2708	46B
CG10513	Unknown	-3.74	0.66	0.86	-2.11	1.63	1.29	3001	96C
CG5065	Unknown	-4.97	1.39	1.43	-4.43	0.54	1.72	3090	53C
CG10748	Tricarboxylic acid cycle	-4.61	2.03	2.56	-4.46	0.15	0.58	3362	69F
GstD5	Glutathione S-transferase	-3.34	0.77	3.40	0.30	3.64	-0.03	3598	87B
CG1924	Chaperone	-4.14	-1.74	-2.28	-4.30	-0.16	0.29	4079	11A
CG5342	Unknown	-5.64	0.76	1.54	-3.83	1.81	1.73	4273	86E
CG1494	ABC transporter	-3.74	-3.65	-3.87	-3.76	-0.02	0.77	4307	19F
CG18518	RNA polymerase	-4.05	2.00	2.47	-3.49	0.56	0.17	4717	35D
CG12725	Ubiquitin-like	-5.14	-0.19	0.68	-4.43	0.71	1.03	5179	11F
Cyp4p3	Cytochrome P450	-3.07	0.35	2.89	0.35	3.42	0.70	5200	45B
CG1268	Hydrogen-exporting ATPase	-3.03	-0.58	1.88	-0.56	2.47	0.25	5781	64A
Arc-p34	Cytoskeleton	-3.50	1.75	2.10	-1.13	2.37	1.83	5850	38C
CG15730	Unknown	-3.60	2.17	0.14	-3.18	0.42	0.39	6076	11A
CG6704	Aryl sulfotransferase/ retinol dehydratase	-3.71	0.67	1.55	-0.93	2.78	1.47	6629	50C

Table S1B. Genes more dependent on *achi*/*vis* than *topi*

Gene	Function	topi-v-achi/vis	aly-v-wild type	comr-v-wild type	topi-v-wild type	achi/vis-v-wild type	can-v-wild type	Wild-type rank	Map
Mst84Dc	Unknown	3.12	-0.65	-0.99	-3.16	-6.27	-1.64	1	84D
CR32658	Non-coding RNA	3.33	-0.82	-1.01	-3.14	-6.47	-1.61	30	11A
CG15219	Unknown	3.76	-1.12	-1.92	-1.79	-5.54	-2.37	33	40E
CG12816	Unknown	5.68	-0.32	-0.55	-2.30	-7.97	-0.87	85	86A
CG17736	Unknown	4.48	-2.86	-2.73	-3.77	-8.24	-3.26	87	76D
CG10396	Cytochrome-c oxidase	3.55	-1.98	-2.14	-4.16	-7.72	-2.39	132	41F
CG7211	Hydrogen-exporting ATPase	4.19	0.59	0.11	0.44	-3.75	0.08	193	28C
Mst57Db @ψ	Oviposition	3.23	-1.10	-1.46	0.72	-2.51	0.22	289	96F
CG9222	Protein kinase	4.22	-0.01	-1.68	-2.87	-7.09	-2.07	386	26B
CR32661	Non-coding RNA	3.12	-0.83	-1.57	-3.65	-6.77	-2.83	427	10F
CG16825	Amino acid transport	3.70	-0.23	-0.96	-0.51	-4.21	-1.39	462	34B
Arp53D	Cytoskeleton	3.53	-0.16	0.22	-3.62	-7.15	0.11	497	53D
CG8349	**Cytidine deaminase**	**5.60**	**-2.63**	**-2.62**	**-2.20**	**-7.80**	**-3.04**	**575**	**28F**
CG7570	Unknown	3.96	-1.31	-1.91	-1.47	-5.44	-1.86	588	76D
CG32316-ORFB	Tricarboxylic acid cycle	3.21	-1.45	-2.88	-2.67	-5.88	-1.93	798	62A
Mst57Da @ψ	Oviposition	3.27	-0.90	-1.13	1.26	-2.02	0.47	811	96F
CG14757	Unknown	4.02	0.20	0.08	-0.54	-4.56	-0.47	882	44B
CG5103	Transketolase	3.42	-3.12	-2.78	-2.83	-6.25	-2.80	909	75C
CG13032	Cytoskeleton	3.09	-0.15	-0.52	-0.63	-3.71	-1.38	967	73B
CG12376	Unknown	4.18	-0.79	-1.93	-2.90	-7.08	-0.58	1196	44E
GstE1	Glutathione S-transferase	3.87	0.18	-3.89	0.70	-3.17	0.11	1232	55C
CG1950	Unknown	5.52	-2.72	-0.91	-0.70	-6.23	-1.29	1297	11A
CG7257	Proteasome regulation	3.58	0.95	0.88	1.49	-2.08	0.57	1391	68E
Acp70A @	Hormone, female mating / oviposition	3.07	-1.44	-2.26	0.04	-3.02	-0.79	1454	70A
CG31872 @#*	Triacylglycerol lipase	3.25	-0.90	-1.94	0.73	-2.52	0.62	1875	32A
CG13501	Unknown	5.39	0.99	0.00	-0.86	-6.24	-1.82	2638	58B
CG12817	Unknown	5.77	3.04	2.17	2.16	-3.60	0.78	2681	86A
CG9227	Unknown	3.43	2.46	2.36	1.20	-2.23	1.28	2737	26B
CG10694	Base-excision repair	3.53	-0.03	0.04	-0.19	-3.71	-0.92	2858	95E
vis	Transcription factor	3.47	3.19	2.71	2.51	-0.96	2.06	2951	49A
msopa @	Unknown	3.52	-0.68	-1.09	1.48	-2.04	0.48	3077	79B
CG17097 @*	Triacylglycerol lipase	3.53	-1.22	-2.33	-0.05	-3.58	0.27	3113	32A
CG14838	Cytoskeleton	4.89	0.49	0.85	0.78	-4.11	0.09	3190	66A
CG14839	Cytoskeleton	5.18	-0.51	-0.31	0.14	-5.05	-1.32	3756	88B
CG11052	Unknown	5.08	1.70	1.62	1.69	-3.39	1.10	3818	84E
Cyp4p1	Cytochrome P450	5.87	-1.61	-6.67	0.72	-5.15	-2.38	4000	45B
CG2667	Diacylglycerol kinase	3.14	-2.16	-1.78	-0.52	-3.67	-1.92	4177	84E
CG9777	mRNA processing	4.27	2.75	2.46	1.16	-3.11	0.34	4557	14F
CG31872 @#*	Triacylglycerol lipase	3.12	-0.54	-1.97	0.95	-2.17	0.41	4967	32A
CG14087	Unknown	6.59	2.71	2.58	3.05	-3.54	1.59	5011	76B
CG1442	Translation initiation	3.49	3.73	3.53	3.41	-0.08	2.99	5139	98F
CG6244	Unknown	5.79	2.66	2.08	1.83	-3.97	2.26	5229	72A
CG31633	Unknown	3.22	2.84	2.98	3.06	-0.16	2.51	5409	26F
Acp33A @	Hormone, regulation of female mating	3.19	-1.43	-3.16	0.86	-2.33	-0.03	5534	33A
CG18284 @*	Lipase	3.20	-1.10	-0.73	0.97	-2.23	0.19	6463	32A
CG18240	Unknown	4.47	-1.39	-1.69	3.06	-1.41	3.49	6981	47D
CG2467	Cell communication	3.35	1.32	0.87	1.50	-1.86	0.01	7027	10F
Obp56d	Pheromone response	3.96	-0.39	-1.70	1.70	-2.26	-0.03	7386	56E
CG13617	Unknown	3.75	3.58	3.03	3.61	-0.14	1.64	7961	96A

Table S1C. Genes whose relative expression in wild-type and mutant testes has previously been determined by northern blotting or RNA in situ hybridisation (White-Cooper et al., 1998)

Gene	Function	aly-v-wild type	comr-v-wild type	topi-v-wild type	achi/vis-v-wild type	can-v-wild type	Wild-type rank	Map
Reduced in aly and can class on northern or by RNA in situ								
Mst84Dc	Unknown	-0.65	-0.99	-3.16	-6.27	-1.64	1	84D
Mst84Dd	Unknown	-1.35	-2.53	-4.42	-4.30	-2.09	3	84D
Mst84Da	Unknown	-0.72	-1.03	-2.88	-4.37	-1.40	4	84D
Mst84Db	Unknown	-1.03	-1.59	-4.04	-5.05	-1.86	9	84D
Mst98Ca	Unknown	-1.17	-2.14	-5.20	-6.34	-1.78	16	98C
Mst87F	Unknown	-2.12	-4.00	-4.97	-5.92	-2.06	26	87F
Mst98Cb	Unknown	-0.76	-1.07	-5.62	-6.38	-2.21	34	98C
dj	Sperm tail	-6.39	-6.82	-8.65	-9.45	-2.36	124	84C
janB	Unknown	-3.74	-4.50	-7.94	-7.89	-2.95	152	99D
gdl	Unknown	-3.84	-3.58	-3.51	-3.89	-2.44	407	71C
fzo	Mitochondrial fusion	-6.75	-6.98	-6.24	-6.56	-2.41	2647	94E
Reduced in aly-class specifically by RNA in situ								
bol	Cell cycle	-2.13	-2.71	-3.41	-3.57	-1.38	345	66F
twe	Cell cycle	-1.06	-1.40	-1.67	-2.65	0.18	2711	35F
CycB	Cell cycle	-0.88	-1.03	-0.84	-0.47	1.16	2944	59A
Expression detected in mutants at similar levels to wild-type by RNA in situ								
rux	Cell cycle	0.11	-0.38	-0.79	-1.19	0.17	972	5D
polo	Cell cycle	0.89	1.38	0.47	0.28	0.47	2174	77B
neb	Cytoskeleton	-1.04	-0.97	-0.76	-1.34	0.02	2267	38B
pelo	Cell cycle	1.20	1.34	0.68	0.93	1.09	2965	30C
CycA	Cell cycle	1.82	1.37	1.45	1.09	1.50	4695	68D

Fold-change expression differences between testis RNA samples analysed on the Affymetrix chips. Fold changes are expressed as the log2 of the expression ratio for each genotype pair. Bright green indicates at least an eightfold reduction (log2 ≤ -3); dark green indicates at least a twofold reduction (log2 ≤ -1); yellow indicates at least a twofold increase (log2 ≥ 1); red indicates at least an eightfold increase (log2 ≥ 3). Wild-type rank is the rank order of the signal intensity for the wild-type samples (so CG11591 was the eighth most intense signal on the wild-type arrays). Predicted function and cytogenetic map position were taken from FlyBase. Genes chosen for further analysis by RT-PCR and in situ are shown in bold.

@, genes for which the one-tailed *t*-test *P*-value for *achi/vis*>*topi* (A) or *topi*>*achi/vis* (B) was greater than 0.005 but less than 0.05. For all other genes $P<0.005$.

$ and # indicate genes with two probes on the array

* and ψ indicate pairs (or more) of homologous genes from the same chromosomal region (i.e. recent duplications).

Table S1D. Raw data for genes in Table S1A-C

Affy oligo	Gene	aly mean	aly s.d.	can mean	can s.d.	comr mean	comr s.d.	topi mean	topi s.d.	achi/vis mean	achi/vis s.d.	red e mean	red e s.d.	*P*<
l48143_at	CG11591	418.2	109.4	1495.7	114.9	736.2	55.4	28	16.9	683.9	114.4	5819.8	36.8	0.005
152130_at	CG4439	968.4	266.9	640.4	39.5	824.7	59.4	22.4	13.1	247.1	15.5	3221.1	94.9	0.005
150373_at	CG7046	1887.1	60.3	1424.6	82.5	1621.2	39.4	84.6	4.4	944.3	87	2993.7	113.3	0.005
l45717_at	CG3213	2.4	0.8	474.8	92.3	2.2	1.1	2.1	0.3	196.8	45	2495.8	23.2	0.005
150757_at	CG5017	113	17.9	641.2	43.3	55.6	12.2	22.9	0.1	207	49.9	2140.8	102.4	0.005
l45983_at	CG7164	5.6	1.8	285	25.7	6.1	1.7	8.4	3.7	114.2	11.8	1504.2	75.6	0.005
l44514_at	TrxT	572.4	83.3	2262.1	119.4	587.6	28.6	22.4	1.1	1499	153.5	1482	136.2	0.005
147306_at	Menl-2	30.2	10.2	292.8	47.3	30.8	8.2	21.4	3	172.4	35.8	1194.8	32.5	0.005
l51867_at	CG6790	158.3	14.1	851.4	36	290.9	14.8	20.8	4.5	201.7	29.5	1104.6	92.9	0.005
l49415_at	CG15177	15.7	3.3	496	38.3	18.6	1.9	12.5	3.8	136.8	6.2	1062.4	13.6	0.005
l42164_at	CG11601	503.6	52.4	612.2	22.8	492.5	9.4	62.4	30.5	545	26.4	958.7	14.2	0.005
147767_at	CG9875	16.8	4.1	271.8	15.9	8.4	2.3	7.2	3.3	133.7	64	921.3	110.3	0.025
l49460_at	CG2649	244.1	38.6	297.5	38.8	80	6.6	8	6.4	91.9	3.8	824.7	33.3	0.005
l44994_at	CG15742	606.4	95.4	655.2	26.3	322.1	23.7	11.1	5.7	213.2	26.1	811.6	4.5	0.005
152409_at	CG4198	179.4	30.4	337.1	71.4	144.2	9.6	5.3	1.5	90.6	14.3	796	9.3	0.005
l47052_at	CG13186	6.2	3.1	68.6	4.3	9.1	5.1	4.6	5	113.4	6.8	646.8	31.9	0.005
l44224_at	CG3994	54.5	19	484.8	33.5	114.2	9.6	12.4	12.3	114.9	5.1	625.3	17.3	0.005
152491_at	CG3121	155.3	31.7	297.3	13.1	18.7	12.3	3.5	0.8	102.3	20.6	561.9	5.7	0.005
l45042_at	CG11068	60.7	9.1	195.6	13.6	81.9	4.4	2.9	1.2	59.8	26.8	426.4	16.5	0.025
l43042_r_at	CG11700	564.9	80.6	2682.1	161.1	657.8	53.9	12.6	7.4	1903.9	285.1	348.8	20.2	0.005
l44441_at	CG16781	117.5	28.4	240.9	18.9	89.3	8.2	5.7	1.1	151.5	16.5	292.5	29.7	0.005
150262_at	CG10881	2.5	0.7	190.5	16.6	2	0.3	2.9	0.8	62.9	10.7	275.8	7.8	0.005
l44099_at	I-t	6.6	5.4	127.5	43.1	19.5	3.1	2.7	1.6	52.2	24.1	240.9	24.2	0.025
l49691_at	CG14708	324	58.1	479	74.7	391.6	15.9	3.6	2.6	148.7	11.6	240.5	6.9	0.005
l43041_i_at	CG11700	408.5	50.2	2223.6	152.7	479.2	41.2	2.9	2.7	1440.2	253.1	236.6	17.3	0.005
150429_at	CG13829	22.2	7	264	14.8	28.2	5.3	6.7	4	83.9	11.9	223.2	25.6	0.005
l47335_at	CG6301	238.3	0.6	74.3	23.8	73.1	29	7	4.6	89.5	30.1	201.1	14.3	0.005
l44389_at	CG14811	844	79.1	265.6	25.1	545.8	38.8	57.6	16.7	641.3	17.1	200.7	8.7	0.005
l50693_at	CG18472	23.6	4.2	185.6	11.1	101.5	21.4	3.9	3	57.9	1.2	175.2	17.8	0.005
l44811_at	CG15296	3.1	1.8	116.3	12.7	4.5	3.6	4.3	1.2	51	8.1	170.4	10.9	0.005
l46900_at	CG10459	171.4	23.3	152.8	9.2	113.5	2.8	2.4	0.6	61.9	8.7	146.1	14.9	0.005
152513_at	CG10513	197.5	78.9	306.7	25.5	227.8	30.9	29	3.5	386.5	3	125.3	21.4	0.005
l47321_at	CG5065	310.1	80.4	390.5	27.2	318.8	12.9	5.5	3.5	171.9	1.2	118.5	7.4	0.005
l48684_at	CG10748	430.1	58.1	157.9	12	623.7	31.4	4.8	2.2	117.1	21.4	105.6	11.7	0.005
l49759_at	GstD5	162.6	127.6	93.2	16.6	1005.3	74	117.2	44.5	1189.3	118.1	95.1	10.2	0.005
l44924_at	CG1924	23.6	3.7	96.1	10.8	16.2	1.4	4	0.5	70.7	7.4	78.8	10	0.005
l49693_at	CG5342	124.7	14.1	245.5	45.3	214.8	4	5.2	1.2	259.3	22	73.8	19.6	0.005
l45470_at	CG1494	5.8	1.6	124.2	10.6	5	0.9	5.4	1.5	72.2	7	73	11.7	0.005
l44298_at	CG18518	252.8	39.3	70.9	12.5	348.7	57.1	5.6	3.5	92.8	4.4	63	6.6	0.005
l45007_at	CG12725	47.5	14.2	110.1	13.1	86.7	11.2	2.5	0.8	88.1	9.6	54	6.8	0.005
l46873_at	Cyp4p3	68.4	20.6	87	10.2	397.4	21	68.1	5.7	572.9	167.3	53.6	19.5	0.005
l48161_at	CG1268	29.7	14	52.9	10.4	164.2	9.1	30.1	8	246.4	38.6	44.5	25.6	0.005
l54503_at	Arc-p34	146.5	16.1	155.1	16.1	187.1	34.3	19.9	6.6	225.8	27.9	43.6	9.9	0.005
l44942_at	CG15730	183.7	11.5	53.4	5.8	44.9	12.6	4.5	2.7	54.7	0.9	40.8	5	0.005
l42360_at	CG6704	54.7	7.6	95.4	18	100.5	3.5	18	2.5	235.8	64.6	34.4	3.1	0.005
l43435_at	Mst84Dc	4529.3	141.7	2273.9	229.4	3589.3	370.5	797.1	51.4	91.8	59.5	7107.6	370	0.005
l51334_at	CR32658	2223.4	97.4	1281.5	98.3	1941.7	110.9	445.6	24.7	44.2	17.3	3922.1	171	0.005
l51006_at	CG15219	1766.2	250.5	740	33	1014.9	49.9	1110.2	84	82.2	14.9	3834	169.7	0.005
l49622_at	CG12816	1809.7	207.8	1235.1	64.2	1541.3	90.6	460.3	71	9	10.9	2263.8	59.1	0.005
l49083_at	CG17736	310.1	58.8	234.4	26.9	339.1	24.3	164.6	7.4	7.4	1	2244.2	133	0.005
l46641_at	CG10396	447.5	70.4	335.8	18.3	401.7	19.8	98.5	9.6	8.4	7.5	1765.2	107.3	0.005
l45978_at	CG7211	2141.3	101.6	1505.4	151.9	1543	187.9	1932.6	199	106.2	14.6	1425.5	39.1	0.005
l43672_at	Mst57Db	544.3	903.4	1357.2	1122.2	422.6	647.8	1922	743.4	205.1	300.4	1165.7	1122.6	0.025
l45874_at	CG9222	970.9	93.2	232.8	24.8	306.7	30.9	134.2	27.6	7.2	5.5	980.4	31.3	0.005
l44907_at	CR32661	517.7	84.9	129.3	17.3	310.2	17.1	73.2	6.8	8.4	1.2	919.3	28.1	0.005
l46329_at	CG16825	746	17.7	335.1	9.8	451.1	23.8	615.7	62	47.4	5	875.3	58.8	0.005
l43686_at	Arp53D	751.2	112.9	906.1	51	975.4	45.6	68	7.3	5.9	2.7	837.6	60.2	0.005
146015_at	CG8349	122	13.4	91.8	9.7	123.2	6.6	164.4	20.4	3.4	2	755.2	55.1	0.005
l49082_at	CG7570	301.6	67.7	205.9	6.6	199	12.6	269.5	42.7	17.3	4.7	748.8	71.3	0.005
l47989_at	CG32316-ORFB	208.3	68.2	149.2	22.7	77.2	10.7	89.8	15.8	9.7	1.9	570.2	55.6	0.005
l43671_at	Mst57Da	300.9	513.1	774.4	686.8	255.6	418.1	1339.2	432.3	138.7	232	560.7	573.4	0.01
l46793_at	CG14757	603.2	80	376.9	7	554.1	61.6	360.4	24.2	22.2	12.2	523.7	23.3	0.005
l48992_at	CG5103	58.8	14.1	73.2	9.5	74.1	0.9	71.9	17.5	6.7	5	510.2	35.7	0.005
l48908_at	CG13032	428.8	18.5	182.7	14.2	332.4	14.3	308.1	37.8	36.2	8	475.2	22.3	0.005
152105_at	CG12376	226.9	24.9	263	39.1	102.6	7.2	52.7	2.1	2.9	1.5	392.1	29.9	0.005
l52598_at	GstE1	430.5	164.9	411.2	33.6	25.6	4.8	616	134.8	42.2	8.6	380.4	33.8	0.005
l44917_at	CG1950	54.7	9.7	146.5	27.2	191.4	9.2	221	15.5	4.8	1.2	359.2	15.6	0.005
l48618_at	CG7257	641.2	21.7	493.5	64	609.1	24.5	936	44.3	78.5	12.6	332.1	13	0.005
l43289_at	Acp70A	117.4	175.9	183.9	149.4	66.4	68.3	327.9	172.5	39.1	25.6	318.2	327	0.025
146174_s_at	CG31872	126.4	193	362.6	337	61.1	85.2	390.5	159.4	41	47.9	235.2	230.6	0.025
l47654_at	CG13501	301.8	57.2	42.8	3.4	151.1	15.7	83.6	9	2	0.4	151.6	27.1	0.005
l49621_at	CG12817	1219.4	46.2	254.9	16.5	668.7	42.6	664.7	57.4	12.2	6.6	148.4	18.1	0.005
l45873_at	CG9227	792.1	203.4	349.5	34.2	735.4	69.5	330.1	28.3	30.6	17.4	143.7	11.2	0.005
150480_at	CG10694	132	16.8	71.3	7.1	138.6	1	118.7	9.9	10.3	2.5	135	13.1	0.005
147099_s_at	vis	1169.1	100.3	535.9	41.5	836.6	91.3	729.7	46.9	66	10	128.2	17.2	0.005
l43463_at	msopa	74.5	104.1	166.8	133.2	56	60.9	333.3	193.3	29	23.9	119.3	110.3	0.05
146175_at	CG17097	50.4	84	141.4	144.9	23.3	34.4	113.3	54	9.8	10.9	117.2	127.9	0.025
l48347_at	CG14838	159.4	25.8	121.2	3.7	204.6	12.8	196	25.8	6.6	1.2	113.8	6.8	0.005
l49888_at	CG14839	62.6	2.8	35.7	10.9	72.2	12.6	98.1	11.8	2.7	1.7	89.2	6.9	0.005
l51011_at	CG11052	284.1	34.4	187	14	268	25.6	281.3	26.7	8.3	6	87.3	12.2	0.005
l43782_at	Cyp4p1	26.8	15.3	15.7	2.5	0.8	0.1	134.7	28.7	2.3	1.9	81.6	8.5	0.005
l49462_at	CG2667	17	12.9	20.2	13.8	22.2	5.1	53	10.8	6	2.8	76.2	3.7	0.005
l45196_at	CG9777	449.6	97.2	84.6	8.5	366	53.2	148.9	8.3	7.7	6.6	66.7	6.5	0.005
l46173_at	CG31872	39.6	60.2	76.7	69.8	14.7	13.4	111.2	49.5	12.8	10.2	57.7	54.2	0.025
l49044_at	CG14087	373.1	45.3	171.2	20.7	341.7	11.2	471.2	27.9	4.9	2.9	57	10.2	0.005
l50798_at	CG1442	723.4	83.8	432.6	58.6	631.3	2.5	580	15.7	51.6	1.2	54.6	2.1	0.005
l48823_at	CG6244	335.1	86.4	254	35.3	224.4	14.3	188.5	16.9	3.4	2.1	53.2	14.4	0.005
l54616_at	CG13770	360.9	49.7	286.9	14.1	398.7	31	420.6	17.9	45	4.5	50.4	4.1	0.005
l43983_r_at	Acp33A	17.9	28.9	47.1	34.3	5.4	1.5	87.4	31.1	9.6	3.8	48.2	35.6	0.01
152307_r_at	CG18284	16.9	26.9	41.3	37.3	21.8	36.1	71	30.9	7.7	11.1	36.2	29.2	0.025
l54692_at	CG18239	11.8	0.9	347.4	25.4	9.6	5.3	257.2	17.9	11.6	3.3	30.9	2.5	0.005
152280_at	CG2467	75.8	7.5	30.6	3.5	55.4	8.1	85.9	4.5	8.4	5	30.4	6.7	0.005
l47516_at	Obp56d	20.8	3.2	26.7	0.3	8.4	3.1	89	7.4	5.7	3.4	27.3	8.9	0.005
150512_at	CG13617	274.2	33.8	71.6	22.1	187	14.7	279.7	47.1	20.8	10.4	22.9	7.8	0.005
l43435_at	Mst84Dc	4529.3	141.7	2273.9	229.4	3589.3	370.5	797.1	51.4	91.8	59.5	7107.6	370	
l43436_at	Mst84Dd	2457.7	134.7	1469.2	81.8	1083.1	134.2	291.5	16.7	317.1	18.8	6249.7	384.2	
l43433_at	Mst84Da	3797.3	101.4	2368.9	24.9	3061.2	208.3	850.1	84.3	301.4	80.6	6244.7	429	
l43434_at	Mst84Db	2792.5	134.9	1566.7	150.6	1888.6	106.8	345.4	60.3	171.5	47.1	5698.3	400.2	
152226_at	Mst98Ca	2237.9	318.6	1466.3	90.4	1144.6	8.8	136.5	7.3	62.1	46.2	5030.5	102.6	
l43274_at	Mst87F	912.5	93	947.3	92.3	248.2	28.8	126	30.1	65.4	24	3962.6	90.1	
l43432_at	Mst98Cb	2239.5	117.7	822.5	9.1	1806.5	193.1	77.1	15.5	45.6	31	3797.9	282.5	
l43902_at	dj	21.7	4.9	354.2	1.2	16	4	4.5	3.3	2.6	1.6	1813.8	56.7	
l43206_at	janB	119.1	20.6	206.4	1.8	70.5	0.4	6.5	2.5	6.7	4.9	1594.1	134.7	
l43176_at	gdl	66.1	12.7	174.8	18.7	79.4	8.7	83	5.2	63.9	5.8	948	84.2	
l43666_at	fzo	1.4	0.3	28.4	6.3	1.2	0.3	2	0.6	1.6	0.9	151	28.3	
l43637_at	bol	241.3	63.2	407.1	48.6	161.9	3.4	99.6	10.4	89.1	1.3	1059.7	52.7	
153947_at	twe	69.8	19.3	165.3	30.8	55.3	5	45.9	13.7	23.3	5.5	146	24.7	
l54177_at	CycB	70	8.7	288	14	63.1	8.4	71.7	6.8	92.8	11.1	128.7	3.5	
l52858_at	rux	512.4	12	533.7	16.2	363.1	5.5	274.4	31.6	207.7	20.8	473.7	42.4	
153742_at	polo	360.1	26.3	269.3	18.3	506.8	75.2	269.5	27.5	236.1	53.5	194.6	4.7	
l41579_at	neb	90	17.8	187.3	29.8	94.4	6.7	109.3	3.2	72.7	8.9	184.6	16.5	
l54677_at	pelo	291.4	1.8	271.2	52.1	321.7	53.2	204.2	27.5	242	30	127.2	23.6	
155068_at	CycA	224.8	36.8	179.8	9.4	164.4	26.4	173.9	21.2	135.6	29.3	63.6	11.4	

Affy oligo is the name of the probe on the array. aly mean is the mean normalised signal intensity for the three replicates of the *aly* testis RNA sample. aly s.d is the standard deviation for these three replicates. Mean and s.d. are similarly shown for the other genotypes used. red e was our control wild-type strain for these experiments. *P* values are the result of a one tailed *t*-test comparing *topi* with *achi/vis*.