Sperm chromatin proteomics identifies evolutionarily conserved fertility factors

Diana S. Chu[1], **Hongbin Liu**[2,3], **Paola Nix**[4,5], **Tammy F. Wu**[1], **Edward J. Ralston**[4,5], **John R. Yates III**[2], and **Barbara J. Meyer**[4,5]

[1]Department of Biology, 1600 Holloway Avenue, San Francisco State University, San Francisco, California 94132, USA.

[2]The Scripps Research Institute, Department of Cell Biology, 10550 North Torrey Pines Road, La Jolla, California 92037, USA.

[3]Agilent Technologies, Inc., 2850 Centerville Road, Wilmington, Delaware 19808, USA.

[4]Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, California 94720, USA.

[5]Howard Hughes Medical Institute, Berkeley, California 94720, USA.

Abstract

Male infertility is a long-standing enigma of significant medical concern. The integrity of sperm chromatin is a clinical indicator of male fertility and *in vitro* fertilization potential[1]: chromosome aneuploidy and DNA decondensation or damage are correlated with reproductive failure. Identifying conserved proteins important for sperm chromatin structure and packaging can reveal universal causes of infertility. Here we combine proteomics, cytology and functional analysis in *Caenorhabditis elegans* to identify spermatogenic chromatin-associated proteins that are important for fertility. Our strategy employed multiple steps: purification of chromatin from comparable meiotic cell types, namely those undergoing spermatogenesis or oogenesis; proteomic analysis by multidimensional protein identification technology (MudPIT) of factors that co-purify with chromatin; prioritization of sperm proteins based on abundance; and subtraction of common proteins to eliminate general chromatin and meiotic factors. Our approach reduced 1,099 proteins co-purified with spermatogenic chromatin, currently the most extensive catalogue, to 132 proteins for functional analysis. Reduction of gene function through RNA interference coupled with protein localization studies revealed conserved spermatogenesis-specific proteins vital for DNA compaction, chromosome segregation, and fertility. Unexpected roles in spermatogenesis were also detected for factors involved in other processes. Our strategy to find fertility factors conserved from *C. elegans* to mammals achieved its goal: of mouse gene knockouts corresponding to nematode proteins, 37% (7/19) cause male sterility. Our list therefore provides significant opportunity to identify causes of male infertility and targets for male contraceptives.

Basic features of DNA compaction and partitioning differ during spermatogenesis and oogenesis[2], but sperm-specific processes are conserved, making *C. elegans* appropriate for identifying and functionally validating male fertility factors. Whereas spermatocytes complete meiotic divisions rapidly to produce four haploid spermatids, oocytes complete meiosis after

fertilization, creating one haploid cell and two polar bodies (Fig. 1a). In most species sperm DNA is packaged uniquely from that of oocytes using small basic structural proteins called sperm nuclear basic proteins (SNBPs). SNBPs in *C. elegans* are yet to be identified, as are factors in any organism that mediate somatic histone displacement, incorporate SNBPs or implement other sperm-specific processes such as transcriptional silencing.

We developed a strategy to identify abundant proteins that co-purify with spermatogenic chromatin and analysed their functions in fertility (Fig. 1b). Chromatin was isolated from populations of spermatogenic or oogenic germ nuclei (Methods). Proteins co-purified with chromatin were identified by using a combination of MudPIT (mass spectrometric analysis to acquire tandem mass spectra[3]) coupled with the SEQUEST algorithm to match spectra to predicted peptides[4] and DTA Select to filter and assemble peptides into corresponding *C. elegans* proteins[5]. In all, 1,099 spermatogenic and 812 oogenic proteins were identified (Supplementary Tables 1–4).

Spermatogenic proteins were prioritized on the basis of abundance measurements derived from MudPIT, which was an advance over previous studies[6,7]. By choosing only abundant spermatogenic proteins that occurred in three or more preparations, we reduced the number of candidate proteins from 1,099 to 502 (Methods). The 597 eliminated proteins constitute only 12% of the relative mass of total spermatogenic proteins (Supplementary Table 5a).

A further reduction in the number of abundant proteins under consideration was achieved by subtracting factors common to spermatogenic and oogenic chromatin samples, eliminating 74% (370 of 502) of spermatogenic proteins (Supplementary Methods and Supplementary Table 5b, c).We expected the remaining proteins to include spermatogenesis-specific proteins and spermatogenesis-enriched proteins expressed in low abundance in oocytes. Thus, from 1,099 initial candidates that co-purify with chromatin, our strategy focused functional analysis on 132 abundant spermatogenic proteins (Supplementary Table 1).

The subcellular localization of 11 candidate proteins demonstrates that spermatogenesis-specific and highly spermatogenesis-enriched chromatin factors were correctly identified. By immunolocalization, eight proteins, representing four different families (GSP-3, GSP-4 (Glc-seven phosphatase), SMZ-1, SMZ-2 (sperm meiosis PDZ), HTAS-1 (histone two A sperm) and SPCH-1, SPCH-2, SPCH-3 (sperm chromatin enriched)), associate specifically with spermatogenic meiotic chromosomes and mature sperm chromatin (Fig. 2b–e and Table 1). These proteins resemble the single known *C. elegans* sperm-specific chromatin-associated protein SPE-11 (spermatogenesis defective), also identified by our proteomic analysis. SPE-11 is a paternal factor supplied to eggs for subsequent use in embryonic development[8] (Fig. 2a and Table 1). Three other proteins (TOP-1 (topoisomerase), GLH-2 (germ line helicase) and HCP-4[CENP-C] (holocentric chromosome-binding protein) known to function in somatic cells[9–11] also associate with sperm chromatin, indicating possible roles in spermatogenesis (Fig. 2f, h, Table 1 and Supplementary Fig. 3). In sum, all 11 tested candidates revealed previously undetected associations with spermatogenic chromatin.

Functional analysis of the 132 proteins further validated our approach. The importance of each protein for fertility was evaluated with the use of RNA interference (RNAi) against cognate genes to reduce gene products. The sterility of treated animals was assayed by brood counts, and chromosome or gonad abnormalities were assessed by cytology (Methods). Although genes required for spermatogenesis are notoriously insensitive to RNAi (S. Ward and S. L'Hernault, personal communication), sterility or embryonic lethality was found for 50 of 132 genes (Supplementary Table 6) and germline cytological defects for 20 of these 50 genes (Table 1 and Supplementary Table 7). RNAi treatment of 18 of 20 genes disrupted aspects of male fertility, including sperm meiotic chromosome segregation and male germline morphology

(Table 1, Supplementary Table 7 and Supplementary Figs 4 and 5). Thus, RNAi analysis showed spermatogenesis-enriched proteins to be important for fertility and embryonic development.

Proteomic identification of factors co-purified with sperm chromatin coupled with injection RNAi and quantitative screening resulted in a higher rate for identifying reproductive factors than that for previous studies. Our rate (38%; 50 of 132) is tenfold that (less than 3%) for RNAi-treated genes detected by microarray analysis for enriched transcript levels during spermatogenesis versus oogenesis[12]. Proteomic analysis identified factors overlapping with (42) and differing from (90) genes transcriptionally enhanced during spermatogenesis. Sterility and embryonic lethality caused by RNAi were observed for 41% of non-overlapping and 31% of overlapping factors. Our rate is also threefold that (10%) for identifying fertility genes through *C. elegans* genome-wide RNAi screens[13,14].

Many *C. elegans* homologues of mammalian fertility factors were identified: 29 of 132 *C. elegans* proteins (representing 24 protein families) correspond to 19 genetically disrupted mouse homologues (Table 1 and Supplementary Table 8); 37% (7 of 19) of mouse knockouts (corresponding to 14 worm homologues in 7 families) caused male infertility. Half (7 of 14) of these worm homologues function in spermatogenesis based on RNAi phenotype (GSP-3 and GSP-4) or have specific association with spermatogenic chromatin (HTAS-1, SPCH-1, SPCH-2, SPCH-3 and GLH-2). Our subtractive proteomics approach is therefore valid for finding evolutionarily conserved, spermatogenesis-enriched proteins essential for fertility.

Analysis of these diverse proteins in *C. elegans* revealed essential functions in spermatogenesis, including roles in meiotic chromosome segregation and chromosome condensation and architecture. Analysed proteins were divided into three categories on the basis of protein localization and RNAi phenotypes: category I proteins have spermatogenesis-specific localization, category II proteins have newly discovered roles in spermatogenesis and previously described roles in other cellular processes (Table 1), and category III proteins have other roles in fertility (Supplementary Table 7).

Category I proteins GSP-3 and GSP-4 are nearly identical homologues of PP1-γ, a glc7/PP1 phosphatase required for mouse male meiosis and spermiogenesis[15]. GSP-3 and GSP-4 encase chromosomes during meiotic divisions and in mature sperm (Fig. 2b). RNAi against either caused variably penetrant male sterility as a result of chromosome segregation defects, including chromosome bridges and incompletely separated chromosomes (Supplementary Fig. 4a–d). Previous *gsp-3* and *gsp-4* RNAi experiments showed hermaphrodite sterility[16,17]. Fertility in *gsp-4*(RNAi) (Supplementary Table 6) and *gsp-3*(RNAi)[17] hermaphrodites was restored by mating with wild-type males, indicating defective spermatogenesis. GSP-3 and GSP-4 may act as the sperm-specific counterparts to the other *C. elegans* Glc7/PP1 phosphatases, GSP-1 and GSP-2, which antagonize the Aurora B kinase-mediated release of chromosome cohesion during mitosis and oocyte meiosis[18]. Hence, studies in worms and mice suggest that GSP family members act directly on chromatin to achieve proper chromosome segregation and fertility in various species.

Also identified were spermatogenesis-specific proteins that lack homology to known fertility factors, yet are essential for sperm meiotic chromosome segregation and male fertility. SMZ-1 and SMZ-2 are 89% identical, contain PDZ (protein–protein interaction) domains, and represent a novel protein family vital for spermatogenic chromosome segregation. These proteins localize to male germline nuclei during and after meiosis and concentrate around sperm chromatin in 30–50% of spermatids (Fig. 2c). *smz-1*(RNAi) or *smz-2*(RNAi) spermatocytes failed to progress through meiotic divisions. Defective meiotic chromosomes did not congress to the metaphase plate or segregate (Supplementary Fig. 4e). Thus, our

approach identified highly conserved and *C. elegans* sperm-specific proteins that are important for male meiosis and fertility.

Many organisms use SNBPs such as histone variants and protamines to establish unique sperm chromosome structure. In mammals, histone H2A variants such as H2AX substitute for H2A during spermatogenesis and are required for male fertility[19]. We identified the first *C. elegans* SNBPs, including the histone H2A variant HTAS-1, which has 51% identity to canonical histones. It localizes with meiotic chromosomes and persists on mature sperm chromatin (Fig. 2e). Histone variants are therefore part of the unique constitution of sperm chromatin in worms and mammals.

Proteomic analysis further identified SPCH-1, SPCH-2 and SPCH-3, three small, abundant, highly basic proteins, which resemble invertebrate protamines such as the surf clam SNBP, PL-I (BLAST expectation score (E value) 10^{-23})[20]. As expected for SNBPs, SPCH proteins are spermatogenesis-specific, localize to meiotic DNA and encase mature sperm chromatin (Fig. 2d). They are also homologous to mouse HANP/H1T2 (ref. [21]) (Table 1), which localizes to mouse sperm chromatin, is required for the nuclear localization of protamines and is important for male fertility[22]. Although the rapid evolution of male reproductive genes[23] has made *C. elegans* SNBPs difficult to identify, proteomic analysis pinpointed SNBPs with similar features in vertebrates and invertebrates.

Unknown roles in DNA compaction and chromosome segregation during spermatogenesis were revealed for proteins with functions in other cell types (category II; Table 1). The topoisomerase I homologue TOP-1 is a nucleolar protein in somatic and germ cells[9], but unlike the subtracted nucleolar protein FIB-1 (Fig. 2g), TOP-1 surrounds mature sperm chromatin (Fig. 2f) and functions during spermatogenesis. Abnormally large sperm nuclei and aberrant progression through male meiosis were caused by *top-1*(RNAi), possibly indicating either defective DNA condensation or abnormal DNA content (Supplementary Fig. 5a, b).

Because topoisomerases alter DNA twist and writhe, TOP-1 may resolve *C. elegans* chromosomes during meiotic divisions, condense DNA during spermatogenesis or decondense DNA after fertilization, which are functions related to those of yeast topoisomerase I in mitosis[24]. Alternatively, because human topoisomerase I is a kinase that phosphorylates SR proteins (key regulators of RNA processing events such as splicing)[25], TOP-1 could control RNA regulation during sperm formation. In fact, *top-1*(RNAi) defects resembled those caused by RNAi of *rsp-6*, which encodes an SR protein (Table 1 and Supplementary Fig. 5b, c). Other sperm-enriched SR proteins, RSP-1 (three occurrences) and RSP-2 (one occurrence), function together during spermatogenesis: double-mutant males are sterile[26].

Of identified factors, 23% (30 of 132) show homology to translation or splicing factors, many belonging to large families. Our subtractive proteomic analysis differentiated between family members to identify those with spermatogenesis-specific chromosome association or function. For example, unlike the subtracted P-granule RNA helicase GLH-1 (Fig. 2i), the GLH-2 family member surrounds sperm meiotic chromosomes (Fig. 2h). Although disruption of *glh-2* had no effect, disruption of *vbh-1* (vasa and belle-like helicase; two occurrences) caused male sterility (R. Navarro, personal communication). Similarly, disruption of the homologous RNA helicase Mvh (murine VASA) causes male infertility[27].

Our study provides clues to the underlying causes of mammalian infertility by identifying and analysing spermatogenesis-enriched chromatin proteins in the nematode *C. elegans*. Cross-species comparison of sperm proteomes, characterized by approaches similar to ours, will elucidate the molecular basis for sperm evolution and male fertility.

Our list of proteins represents a resource for identifying causes of human infertility, thereby providing candidates for any disease locus residing in a multigenic region correlated with a fertility defect. One illustration is azoospermia, the complete lack of sperm[28]. *DBY*, which encodes a VASA RNA helicase, falls within a Y-chromosome region correlated with this defect, and is further implicated in spermatogenesis by our analysis.

Of the 132 proteins detected by our strategy, 70 (53%) have human homologues not yet tested for roles in fertility (Table 1 and Supplementary Table 9). Analysis of these proteins in mice and humans has the potential to define new mammalian fertility factors. These factors present opportunities for the development of diagnostic tests to assess sperm competence and human reproductive potential, and represent potential targets for the development of safe male contraceptives.

METHODS

Purification of spermatogenic and oogenic chromatin

Germ-cell isolation and chromatin purification are described in Supplementary Methods; 5–20% of spermatogenic nuclei were in various stages of meiosis as judged by a cytological examination of 4′,6-diamidino-2-phenylindole (DAPI)-stained chromosomes. Similarly, 10–30% of *fer-1(hc1)* oocytes were in diakinesis, 50–70% were oocytes that had undergone endomitotic reduplication (Emo), and 20–30% were fertilized embryos, allowing the subtraction of chromatin and meiotic factors. Subsequent SDS–PAGE and staining with colloidal blue (Novex) of purified chromatin revealed that core histones were the most abundant proteins in both cell types, showing similar enrichment of chromatin proteins (Supplementary Fig. 2). Major sperm proteins (MSPs), the most abundant proteins in sperm cytosol and pseudopods[29], were undetectable in our chromatin preparations by staining with colloidal blue, indicating that sperm cellular components had been effectively removed.

Proteomic identification and subtractive analysis

A total of eleven 12-step LC–LC–MS–MS experiments (six of spermatogenic chromatin and five of oogenic chromatin) were performed to enhance the detection of peptides from small proteins and low-abundance proteins. For each experiment, 50–100µg of protein was digested and analysed (Supplementary Methods). To ensure the subsequent subtraction of low-abundance oogenic factors, more stringent criteria were adopted for identifying proteins from spermatogenic germ cells (a minimum of two different peptides per protein from each individual preparation) than from oogenic cells (a minimum of two different peptides per protein from all data sets of five preparations combined).

Prioritization of enriched and reliably detected proteins by using two abundance measurements from MudPIT reduced the number of proteins to a manageable subset for functional studies. The first, total spectrum count (TSC), represents the total number of tandem mass spectra collected for each protein from all preparations and was used to assess relative protein abundance[30]. The second measure, 'occurrence', was the number of preparations in which each protein was identified and indicates the reliability of protein detection across different preparations. Occurrence is especially useful for proteins that are not sampled frequently in any one mass spectrometric analysis. By choosing only spermatogenic proteins represented by three or more occurrences, we reduced the number of candidate proteins from 1,099 to 502. These well-represented spermatogenic proteins comprise 46% (502 of 1,099) of individual proteins identified by MudPIT analysis and constitute 88% of the mass of total proteins in spermatogenic preparations, as estimated by TSC (Supplementary Table 5a). Enrichment of proteins such as core histones was confirmed by high TSC and high occurrence (6 of 6

preparations; Supplementary Fig. 2 and Supplementary Table 3). Correspondingly, MSPs were detected in only very low abundance (Supplementary Tables 2 and 3).

The entire oogenic protein list (812) was subtracted from the prioritized spermatogenic protein list (502). Subtractive analysis removed appropriate factors and reduced the number of proteins to 132 (Supplementary Methods).

Functional analysis

RNAi analysis was conducted for genes corresponding to each of the 132 abundant sperm chromatin proteins by injecting 1–4 mg ml^{-1} double-stranded RNA into *him-8(e1849)* animals. Quantitative screening and statistical analysis of F$_2$ animals from RNAi-treated grandparents were used to assess the functions of each gene in fertility (Supplementary Methods). F$_1$ parents producing statistically lower progeny numbers were further analysed by observing the morphology of germ-cell chromosomes (Supplementary Methods). To assess male fertility, FS male progeny were tested for their ability to rescue the fertilization defect of *spe-8 dpy-4* hermaphrodites.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Agarwal A, Said TM. Role of sperm chromatin abnormalities and DNA damage in male infertility. Hum. Reprod. Update 2003;9:331–345. [PubMed: 12926527]

2. Kimmins S, Sassone-Corsi P. Chromatin remodelling and epigenetic features of germ cells. Nature 2005;434:583–589. [PubMed: 15800613]

3. Washburn MP, Wolters D, Yates JR. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nature Biotechnol 2001;19:242–247. [PubMed: 11231557]

4. Eng JK, McCormack A, Yates JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J. Am. Soc. Mass Spectrom 1994;5:976–989.

5. Tabb DL, McDonald WH, Yates JR. DTA Select and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. J. Proteome Res 2002;1:21–26. [PubMed: 12643522]

6. Schirmer EC, Florens L, Guan T, Yates JR, Gerace L. Nuclear membrane proteins with potential disease links found by subtractive proteomics. Science 2003;301:1380–1382. [PubMed: 12958361]

7. Skop AR, Liu H, Yates J, Meyer BJ, Heald R. Dissection of the mammalian midbody proteome reveals conserved cytokinesis mechanisms. Science 2004;305:61–66. [PubMed: 15166316]

8. Browning H, Strome S. A sperm-supplied factor required for embryogenesis in *C. elegans*. Development 1996;122:391–404. [PubMed: 8565851]

9. Lee MH, Park H, Shim G, Lee J, Koo HS. Regulation of gene expression, cellular localization, and in vivo function of *Caenorhabditis elegans* DNA topoisomerase I. Genes Cells 2001;6:303–312. [PubMed: 11318873]

10. Gruidl ME, et al. Multiple potential germ-line helicases are components of the germ-line-specific P granules of *Caenorhabditis elegans*. Proc. Natl Acad. Sci. USA 1996;93:13837–13842. [PubMed: 8943022]

11. Moore LL, Roth MB. HCP-4, a CENP-C like protein in *Caenorhabditis elegans*, is required for resolution of sister centromeres. J. Cell Biol 2001;153:1199–1208. [PubMed: 11402064]

12. Reinke V, Gil I, Ward S, Kazmer K. Genome-wide germline-enriched and sex-biased expression profiles in *Caenorhabditis elegans*. Development 2004;131:311–323. [PubMed: 14668411]

13. Kamath RS, et al. Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. Nature 2003;421:231–237. [PubMed: 12529635]

14. Sonnichsen B, et al. Full genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. Nature 2005;434:462–469. [PubMed: 15791247]

15. Varmuza S, et al. Spermiogenesis is impaired in mice bearing a targeted mutation in the protein phosphatase 1c gamma gene. Dev. Biol 1999;205:98–110. [PubMed: 9882500]

16. Boag PR, Ren P, Newton SE, Gasser RB. Molecular characterisation of a male specific serine/ threonine phosphatase from Oesophagostomum dentatum (*Nematoda: Strongylida*), and functional analysis of homologues in *Caenorhabditis elegans*. Int. J. Parasitol 2003;33:313–325. [PubMed: 12670516]

17. Hanazawa M, Mochii M, Ueno N, Kohara Y, Iino Y. Use of cDNA subtraction and RNA interference screens in combination. Proc. Natl Acad. Sci. USA 2001;98:8686–8691. [PubMed: 11447255]

18. Rogers E, Bishop JD, Waddle JA, Schumacher JM, Lin R. The aurora kinase AIR-2 functions in the release of chromosome cohesion in *Caenorhabditis elegans* meiosis. J. Cell Biol 2002;157:219–229. [PubMed: 11940606]

19. Celeste A, et al. Genomic instability in mice lacking histone H2AX. Science 2002;296:922–927. [PubMed: 11934988]

20. Lewis JD, McParland R, Ausio J. PL I of *Spisula solidissima*, a highly elongated sperm specific histone H1. Biochemistry 2004;43:7766–7775. [PubMed: 15196019]

21. Lewis JD, Song Y, de Jong ME, Bagha SM, Ausio J. A walk though vertebrate and invertebrate protamines. Chromosoma 2003;111:473–482. [PubMed: 12743711]

22. Martianov I, et al. Polar nuclear localization of H1T2, a histone H1 variant, required for spermatid elongation and DNA condensation during spermiogenesis. Proc. Natl Acad. Sci. USA 2005;102:2808–2813. [PubMed: 15710904]

23. Swanson WJ, Vacquier VD. The rapid evolution of reproductive proteins. Nature Rev. Genet 2002;3:137–144. [PubMed: 11836507]

24. Wang JC. Cellular roles of DNA topoisomerases: a molecular perspective. Nature Rev. Mol. Cell Biol 2002;3:430–440. [PubMed: 12042765]

25. Rossi F, et al. Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I. Nature 1996;381:80–82. [PubMed: 8609994]

26. Kawano T, Fujita M, Sakamoto H. Unique and redundant functions of SR proteins, a conserved family of splicing factors, in *Caenorhabditis elegans* development. Mech. Dev 2000;95:67–76. [PubMed: 10906451]

27. Tanaka SS, et al. The mouse homolog of *Drosophila* Vasa is required for the development of male germ cells. Genes Dev 2000;14:841–853. [PubMed: 10766740]

28. Vogt PH. Azoospermia factor (AZF) in Yq11: towards a molecular understanding of its function for human male fertility and spermatogenesis. Reprod. Biomed. Online 2005;10:81–93. [PubMed: 15705299]

29. Ward S, Klass M. The location of the major protein in *Caenorhabditis elegans* sperm and spermatocytes. Dev. Biol 1982;92:203–208. [PubMed: 7049792]

30. Liu H, Sadygov RG, Yates JR. A model for random sampling and estimation of relative protein abundance. Anal. Chem 2004;76:4193–4201. [PubMed: 15253663]
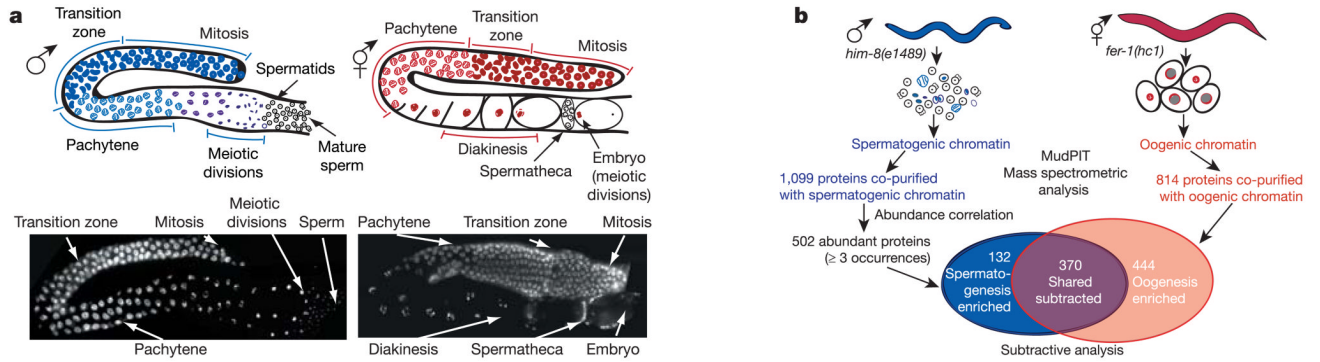
Acknowledgements

Figure 1. Abundance-correlated subtractive proteomic strategy to identify spermatogenesis-enriched proteins

a, Gametogenesis in *C. elegans*. Upper panels illustrate the progression of germ-cell nuclei in a wild-type male gonad (left) and in one arm of a wild-type bilaterally symmetric hermaphrodite gonad (right). Lower panels show corresponding DAPI-stained nuclei from dissected and fixed gonads. Visually distinct stages of gametogenesis are labelled. **b**, Proteomic strategy. Spermatogenic chromatin was isolated from germ nuclei of adult XO males derived from the *him-8(e1489)* X chromosome non-disjunction mutant. Oogenic chromatin was isolated from germ cells of *fer-1(hc1)* hermaphrodites. *fer-1(hc1)* mutants produce defective sperm, causing XX animals to be functional females in which some of the oocytes mature, are ovulated unfertilized, and undergo endoreduplication. Spermatogenic and oogenic chromatin was subjected to MudPIT for identification of associated proteins. Subtractive analysis of chromatin proteins identified 132 abundant spermatogenesis-enriched proteins.
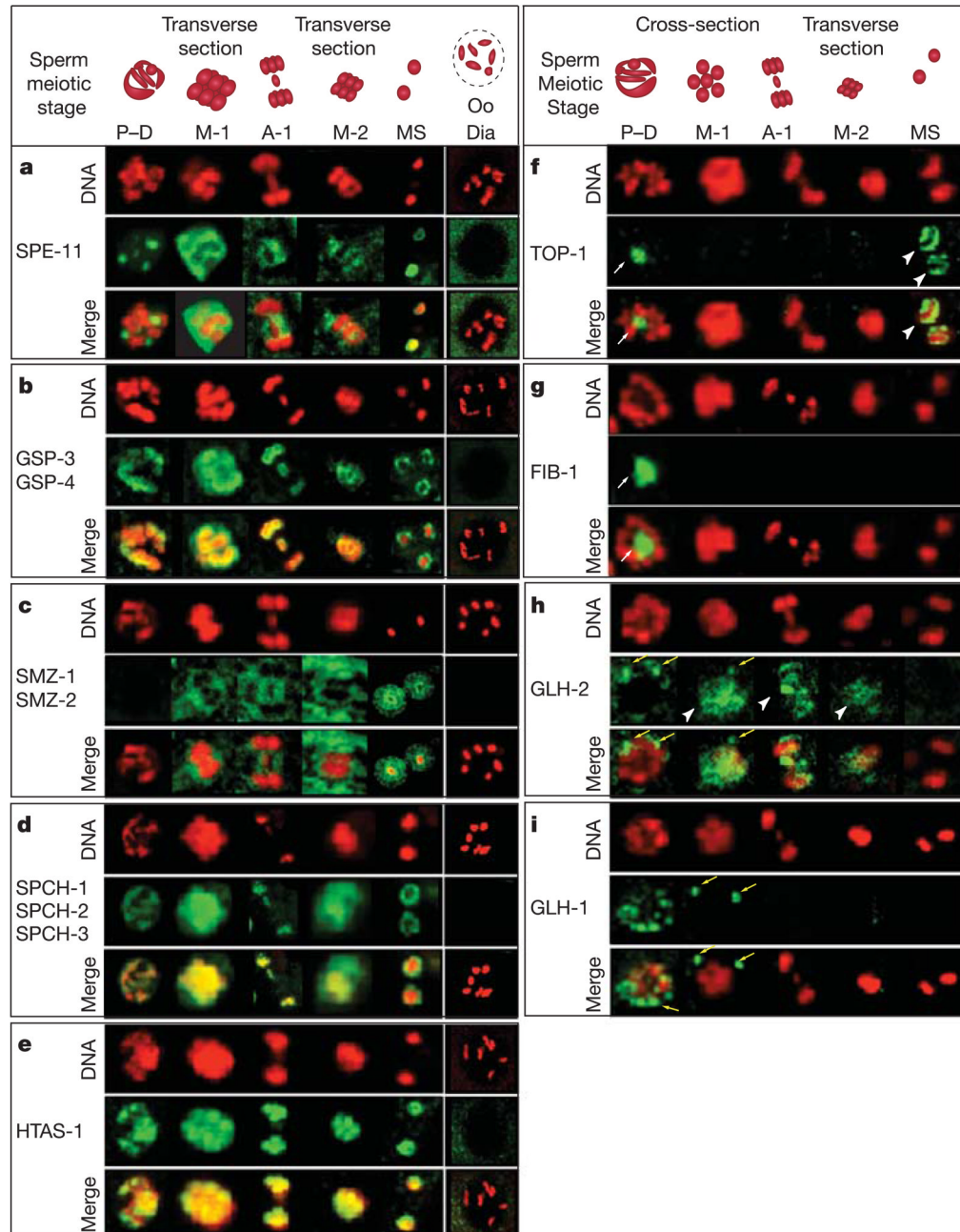
Figure 2. Identification of spermatogenic chromatin-associated proteins by proteomic analysis
Immunolocalization of spermatogenic proteins show chromatin association in mature sperm
and during meiotic divisions. Sperm meiotic stage: P–D, pachytene to diakinesis transition;
M-1, metaphase 1; A-1, anaphase 1; M-2, metaphase 2; MS, mature sperm. Oogenesis: Oo
Dia, oocyte diakinesis. DNA is shown in red, antibody staining in green. **a–e**, Spermatogenesis-
specific chromatin-associated proteins (proteins were not detected during any stage of
oogenesis): **a**, SPE-11; **b**, GSP-3 and GSP-4; **c**, SMZ-1 and SMZ-2; **d**, SPCH-1, SPCH-2 and
SPCH-3; **e**, HTAS-1. **f, g**, Spermatogenic localization of nucleolar proteins: **f**, TOP-1; **g**, FIB-1
(*C. elegans* fibrillarin). TOP-1 surrounds mature sperm chromatin (white arrowheads) but
FIB-1 does not (white arrows). Nucleolar localization (white arrows) is ubiquitous in

spermatogenic and oogenic germ lines (data not shown, and ref. [9]). **h, i**, Spermatogenic localization of P-granule components: **h**, GLH-1; **i**, GLH-2. GLH-2 associates with sperm meiotic chromatin (white arrowheads) but GLH-1 does not. Both localize to perinuclear P-granules (yellow arrows) during spermatogenesis and oogenesis (data not shown, and ref. [10]).

Table 1

Abundant *C. elegans* spermatogenesis-enriched chromatin proteins with roles in fertility

Gene ID	Gene name	Descriptor	*Caenorhabditis elegans*					Mammalian	
			Localization	RNAi cytological defects		Phenotype	Sum: new evidence for fertility function	Known homologue	Fertility link
				Male	Hermaphrodite				
Category I									
F48C1.7	spe-11	Paternal factor required for development[8]	High levels on sperm DNA during and after meiotic divisions	None detected	Embryos do not form hard eggshells, Emb (mutant)	Male Ste (mutant)/(RNAi resistant)			
W09C3.6*	gsp-3	Glc7/PP1 phosphatase	High levels in sperm, strong around meiotic and mature sperm DNA	Defective sperm chromosome segregation	Embryos do not form hard eggshells, Emb	High male Ste	Spermatogenesis-specific localization/ meiotic sperm chromosome segregation defects	M. Ppp1cc (1×10^{-102}) H. PPP1C (1×10^{-95})	M. Ppp1cc KO male infertile
T03F1.5*	gsp-4	Glc7/PP1 phosphatase	High levels in sperm, strong around meiotic and mature sperm DNA	Defective sperm chromosome segregation	Embryos do not form hard eggshells, Emb	High male Ste	Spermatogenesis-specific localization/ meiotic sperm chromosome segregation defects	M. Ppp1cc (1×10^{-102}) H. PPP1C (1×10^{-95})	M. Ppp1cc KO male infertile
C25G4.6†	smz-1	PDZ protein-protein interaction domain	High levels in sperm, strong around 50% of mature sperm DNA	Defective sperm chromosome segregation	Embryos do not form hard eggshells, Emb	High male Ste	Spermatogenesis-specific localization/ meiotic sperm chromosome segregation defects		
T21G5.4†	smz-2	PDZ protein-protein interaction domain	High levels in sperm, strong around 50% of mature sperm DNA	Defective sperm chromosome segregation	Embryos do not form hard eggshells, Emb	High male Ste	Spermatogenesis-specific localization/ meiotic sperm chromosome segregation defects		
ZK1251.1	htas-1	Histone H2A variant	High levels on sperm DNA during and after meiotic divisions	None detected (mutant)	None detected	Partial Ste (mutant)/(RNAi resistant)	Spermatogenesis-specific localization, mutant hermaphrodite has decreased fertility	M. H2AX (2×10^{-21}), macroH2A(3×10^{-27}) H. H2AX (3×10^{-21}), macroH2A (3×10^{-27})	M. H2AX KO male infertile
C04G2.8‡	spch-1	SNBP-like	High levels on sperm DNA during and after meiotic divisions	Low penetrance male meiotic problems	None detected	Low Ste/(RNAi resistant)	Spermatogenesis-specific localization	M. histone H1T2 (2.3×10^{-14}) H. SON DNA-binding protein (1.3×10^{-19})	M. Histone H1T2 KO male infertile
C10G11.9†	spch-2	SNBP-like	High levels on sperm DNA during and after meiotic divisions	None detected	None detected	(RNAi resistant)	Spermatogenesis-specific localization	M. Histone H1T2 (2.3×10^{-14}) H. SON DNA-binding protein (1.3×10^{-19})	M. Histone H1T2 KO male infertile
T27A3.4‡	spch-3	SNBP-like	High levels on sperm DNA during and after meiotic divisions	None detected	None detected	(RNAi resistant)	Spermatogenesis-specific localization	M. histone H1T2 (2.3×10^{-14}) H. SON DNA-binding protein (6.5×10^{-20})	M. Histone H1T2 KO male infertile

Category II

Gene ID	Gene name	Descriptor	*Caenorhabditis elegans* Localization	RNAi cytological defects Male	RNAi cytological defects Hermaphrodite	Phenotype	Sum: new evidence for fertility function	Mammalian Known homologue	Mammalian Fertility link
M01E5.5	top-1	Topoisomerase I[9]	Sperm: high levels around mature sperm DNA / Other: nucleolar, absent during meiosis	Abnormal gonad morphology, meiotic problems with large nuclei in later germ line	Abnormal gonad morphology, oogenesis arrest or Emo	Complete F_1 sterility, Gro	Localization around mature sperm chromatin	M. TOP-1 (3×10^{-154}); H, TOP-1 (1×10^{-156})	M, TOP-1 KO Embryonic lethal; H, decreased Top-1 activity in testes of infertile men with varicocele
C55B7.1	glh-2	RNA helicase[10]	Sperm: surrounding sperm meiotic chromosomes / Other: P-granule localization in germ line	None detected	None detected	WT	Localization around sperm meiotic chromatin	M. Ddx4 (VASA) (1×10^{-105}); H, DBY (1×10^{-92}), DDX4 (1×10^{-105})	M, Ddx4 KO male infertile; H, *DBY* is frequently deleted in infertile patients
T03F1.9	hcp-4	CENP-C centromere component[11]	Sperm: surrounding sperm chromosomes during and after meiotic divisions / Other: uniformly distributed on oocyte meiotic chromosomes, polewards on mitotic chromosomes	No F_1 progeny	No F_1 progeny	Complete F_1 lethality, Emb	Localization around mature sperm chromatin	M, CENP-C[11]; H, CENP-C[11]	

The categories are described in the text. Degrees of sterility are defined in Supplementary Methods. Symbols (*, †, ‡) denote highly identical genes whose products may be depleted simultaneously by RNAi. Mammalian fertility link references are listed in Supplementary Table 8. RNAi resistant means that the protein product was detectable by immunolocalization in animals subjected to RNAi of the corresponding gene. Extent of homology is indicated in parenthesis (BLAST e-values) in 'Known homologue' column. Emb, embryonic lethality; Emo, endomitotic reduplication; Gro, slow growing; H, human homologues; KO, mouse knockout of corresponding gene; M, mouse homologues; Ste, sterility; WT, wild type.