Molecular Cell, Vol. 6, 605–616, September, 2000, Copyright ©2000 by Cell Press

A Global Profile of Germline Gene Expression in *C. elegans*

Valerie Reinke,* Harold E. Smith,†# Jeremy Nance,†# John Wang,*# Carrie Van Doren,* Rebecca Begley,* Steven J. M. Jones,§ Elizabeth B. Davis,† Stewart Scherer,‡ Samuel Ward,† and Stuart K. Kim*‖
*Department of Developmental Biology
Stanford University School of Medicine
Stanford, California 94305
†Department of Molecular and Cellular Biology
University of Arizona
Tuscon, Arizona 85721
‡Rosetta Inpharmatics
Kirkland, Washington 98034
§Genome Sequence Center
British Columbia Cancer Research Center
Vancouver, British Columbia V5Z 1L3
Canada

Summary

We used DNA microarrays to profile gene expression patterns in the *C. elegans* germline and identified 1416 germline-enriched transcripts that define three groups. The sperm-enriched group contains an unusually large number of protein kinases and phosphatases. The oocyte-enriched group includes potentially new components of embryonic signaling pathways. The germline-intrinsic group, defined as genes expressed similarly in germlines making only sperm or only oocytes, contains a family of *piwi*-related genes that may be important for stem cell proliferation. Finally, examination of the chromosomal location of germline transcripts revealed that sperm-enriched and germline-intrinsic genes are nearly absent from the X chromosome, but oocyte-enriched genes are not.

Introduction

The completed sequence of the *C. elegans* genome has allowed the prediction of nearly all the genes in this animal that permit it to develop from a fertilized egg, to learn and behave, to mate and reproduce, and to age (*C. elegans* Sequencing Consortium, 1998). The next challenge is to combine this vast amount of sequence information with functional information to understand the coordinated activity of these genes. An important step in establishing function is to determine the location, timing, and level of expression of each gene. DNA microarrays can be used to rapidly assess the expression of many genes in parallel and thus determine gene expression patterns on a global scale (Schena et al., 1995; Shalon et al., 1996). The completed *C. elegans* sequence allows DNA microarrays to be constructed that contain not only known genes but also previously unstudied genes.

‖To whom correspondence should be addressed (email: kim@cmgm.stanford.edu).
#These authors contributed equally to this work.

As a first step in determining the gene expression patterns that underlie the development of *C. elegans*, we have used DNA microarrays to identify genes whose expression is increased in the germline. The germline differs from somatic tissue in several ways. The germline is totipotent because it generates every tissue in an individual of the next generation. It is also immortal, since it perpetuates itself indefinitely by producing subsequent generations. In contrast, every cell in the soma differentiates into a particular type and survives only as long as the individual. The complete molecular depiction of germline gene expression generates an overview of all the biological processes in that tissue. In addition, the list of germline-enriched genes can be used to identify particular genes with specific roles in stem cells, meiosis, or gamete formation.

The hermaphrodite germline of *C. elegans* is a complex tissue with diverse biological processes occurring in different regions of the syncytial gonad (reviewed in Schedl, 1997). The distal end of each tubular gonad arm contains a germline stem cell population that undergoes mitotic proliferation. As germ cell nuclei move more proximally, they stop mitotic proliferation and initiate meiosis, during which the chromosomes undergo pairing, synapsis, and recombination. Germ cells then travel into the proximal portion of the gonad and initiate gametogenesis. Germ cells that mature during the fourth larval stage (L4) develop into spermatocytes, which then differentiate into nonflagellated crawling sperm (reviewed in L'Hernault, 1997). Germ cells in the adult differentiate into oocytes, which make the eggshell and contain the maternal components to direct embryonic patterning events.

We have produced DNA microarrays containing genomic PCR products corresponding to 11,917 *C. elegans* genes. To identify genes expressed in the germline, we compared wild-type gene expression levels to that of *glp-4* mutants, in which the germline precursor cells do not proliferate. We also compared a mutant strain making only sperm, *fem-3(gf)*, to a mutant strain producing only oocytes, *fem-1(lf)*, to identify both sperm-enriched and oocyte-enriched genes. Using a statistical criterion for significance, these experiments define 1416 germline-expressed genes that fall into three categories: 650 sperm-enriched genes, 258 oocyte-enriched genes, and 508 germline-intrinsic genes. We have determined the temporal expression pattern of each of these genes during development. These germline genes comprise a molecular definition of germline components, and they also provide the framework to identify individual genes involved in specific germline functions. Surprisingly, we observed that few germline-intrinsic and sperm-enriched genes are found on the X chromosome, while oocyte-enriched genes are evenly distributed throughout the genome. We discuss possible causes for the strong bias in the chromosomal locations of these genes.

Results

Identification of Germline-Enriched Genes

We constructed DNA microarrays containing 11,917 genes, representing 63% of the 19,099 predicted genes

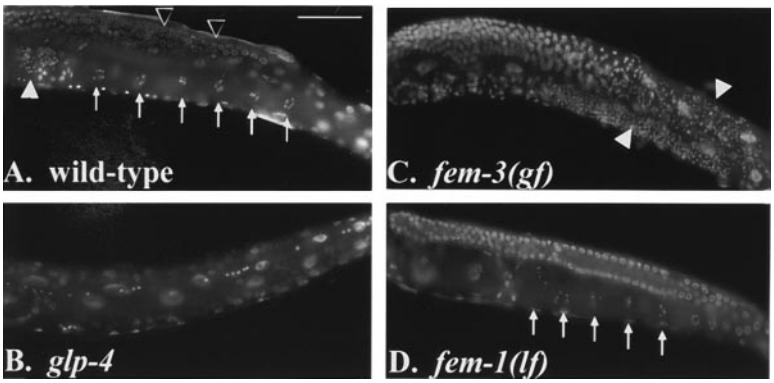

Figure 1. Germline Mutants

(A) Wild-type. Germline has bright compact sperm nuclei stored in the spermatheca (black triangles), oocytes with condensed chromosomes in the proximal arm (arrows), and meiotic and mitotic cells (white triangles).

(B) *glp-4(bn2)* animals lack germline nuclei present in the somatic gonad.

(C) *fem-3(gf)* animals have large numbers of developing sperm in the proximal gonad (white triangles), but no oocytes.

(D) *fem-1(lf)* nematodes lack sperm, but produce oocytes(arrows). Animals are stained with DAPI.

Bar, 50 μm.

in the genome, to identify genes preferentially expressed in the *C. elegans* germline. We first compared wild-type hermaphrodites (Figure 1A) to a strain that lacks virtually all germ cells, due to a defect in germline stem cell proliferation caused by a temperature-sensitive mutation in the *glp-4* gene (Beanan and Strome, 1992) (Figure 1B). To determine the temporal pattern of gene expression during the formation of the germline, we collected three or four independent poly(A)$^+$ RNA samples from wild-type and *glp-4* mutant hermaphrodites at each of four developmental stages: three larval stages (L2, L3, and L4) and young adults. We synthesized Cy3-labeled cDNA from each staged experimental RNA and compared its hybridization to the microarray with Cy5-labeled reference cDNA synthesized from mixed-stage wild-type hermaphrodite poly(A)$^+$ RNA. We calculated that about 59% of all the genes on the microarray are expressed significantly above background in mixed-stage wild-type hermaphrodite poly(A)$^+$ RNA ($p < 0.05$, Student's t test, see http://cmgm.stanford.edu/~kimlab/germline).

For each hybridization, we calculated the expression level in the experimental sample relative to the reference and then averaged the results for samples of the same genotype and developmental stage. Since all of the samples are measured against the same reference RNA, we can compare average gene expression levels from one genotype and developmental stage to that of any other. We used repeated DNA microarray hybridizations from a variety of experiments to estimate the standard error in measuring the expression ratio of each gene (61 repeats involving 15 different types of experiments, S. K. K., unpublished data) and then used the standard error of each gene to determine whether its expression in wild-type and *glp-4* mutants was significantly different (Student's t test, Experimental Procedures). Germline-expressed genes were defined as those with higher levels in wild-type than in *glp-4* mutant hermaphrodites in at least one of the four developmental stages at the 99.9% confidence level, which corresponds to expression ratios from 1.8- to 104-fold. We detected more germline-enriched genes as the germline increased in size and complexity: 1 gene was identified in the L2 stage, 16 genes in the L3 stage, 307 genes in the L4 stage, and 674 genes in the young adult. Upon combining the results from each developmental stage, we identified a total of 875 potential germline-enriched genes.

We next identified genes differentially expressed during spermatogenesis and oogenesis by comparing mRNA isolated from mutant hermaphrodites making only sperm to mRNA isolated from mutants making only oocytes. The temperature-sensitive, gain-of-function mutation *q23* in the gene *fem-3* causes the hermaphrodite germline to produce only sperm instead of oocytes, while the soma is apparently normal (Barton et al., 1987) (Figure 1C). The temperature-sensitive loss-of-function mutation *hc17* in the gene *fem-1* yields hermaphrodites incapable of initiating spermatogenesis so that they make only oocytes (Nelson et al., 1978) (Figure 1D). We prepared five independent RNA samples from young adult *fem-3(gf)* and *fem-1(lf)* animals. For each comparison, we directly hybridized Cy3-labeled *fem-1(lf)* cDNA and Cy5-labeled *fem-3(gf)* cDNA to a microarray and determined average gene expression differences. At the 99.9% confidence level, 908 genes showed differing expression levels between *fem-1* and *fem-3* mutants, which corresponds to expression ratios between 1.6- and 71-fold. Six hundred fifty genes had higher expression in the *fem-3* strain making sperm, and 258 genes had higher expression in the *fem-1* strain making oocytes. We refer to these classes as sperm enriched and oocyte enriched, respectively. Both classes likely include some genes expressed in the somatic tissue but induced by the presence of gametes.

We combined the results from the wild-type/*glp-4* microarray experiments with those from the *fem-1(lf)/fem-3(gf)* experiments to identify a total of 1416 germline genes. This number represents 11.8% of all of the genes on the microarray. In addition to the 650 sperm-enriched genes and 258 oocyte-enriched genes that are likely to function in gametogenesis and early embryonic development, 508 germline-enriched genes are expressed at levels that are not significantly different between mutants making only sperm or only oocytes. We refer to this class of genes as germline intrinsic, and these genes may have a role in common germline activities, such as mitotic proliferation, meiosis, and recombination. The *fem-1(lf)* and *fem-3(gf)* mutant strains produce more gametes than wild-type hermaphrodites, so many genes differentially expressed in the *fem-1(lf)/fem-3(gf)* experiment were not detected as differentially expressed in the wild-type/*glp-4* experiment. The complete results from these DNA microarray experiments can be accessed at http://cmgm.stanford.edu/~kimlab/germline.

Validation of the DNA Microarray Results

The 11,917 genes on the DNA microarrays include 71 genes known from prior work to be germline-enriched genes that can serve as positive controls. These genes

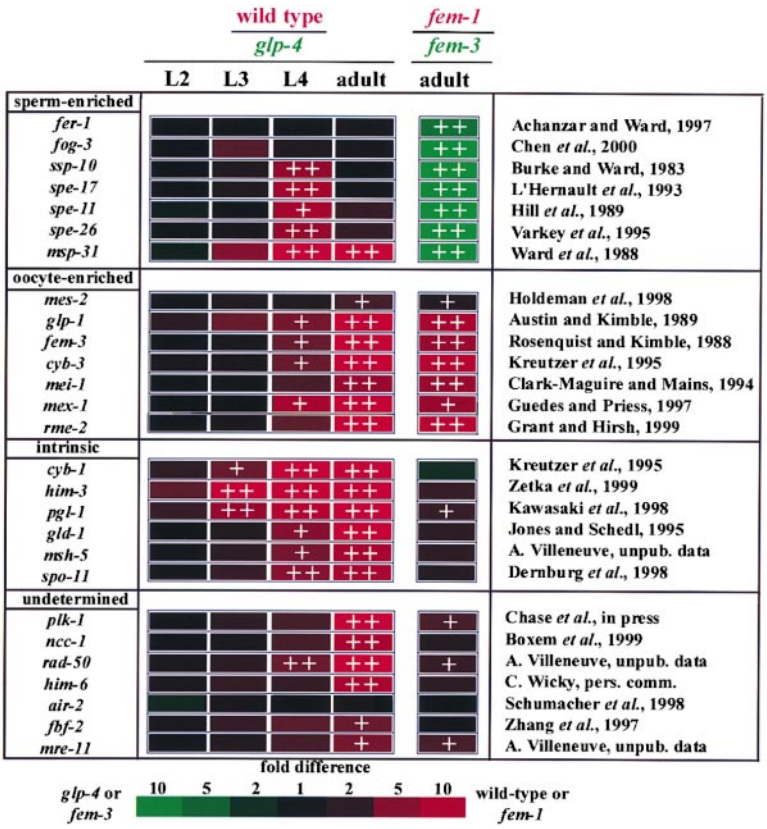

Figure 2. Expression Levels of 27 Genes Previously Known to Have Germline-Enriched Expression

Red boxes indicate increased expression in wild-type or *fem-1*, and green indicates increased expression in *glp-4* or *fem-3*. "+" signifies 99% confidence interval; "++" denotes 99.9% confidence interval. The microarray data show germline-enriched expression for 26 of the 27 previously identified germline genes, with the exception of *air-2*. Four genes were poorly spotted or missing (*him-14*, *spe-12*, *cya-1*, and *ssr-1*). In addition to the 27 genes shown here, 40 additional genes are listed at http://cmgm.stanford.edu/~kimlab/germline that are previously known to be germline enriched, and all 40 were identified in these microarray experiments.

are involved in a wide variety of processes, including gamete sex determination, gamete differentiation, embryonic patterning, germline specification, germline proliferation, and meiosis. The positive control genes included 43 genes known to be sperm enriched, 10 genes known to be oocyte enriched, and 9 genes known to be expressed in the germline but at similar levels during spermatogenesis and oogenesis (referred to as germline-intrinsic genes). A fourth group contains 11 genes known to be germline enriched but whose expression during spermatogenesis or oogenesis was not previously examined. Figure 2 shows 27 representative genes; the entire set can be viewed at the accompanying website. Of these positive controls, 55 are strongly germline enriched in our experiments (at a 99.9% confidence interval), and an additional 11 are moderately enriched in the germline (at a 99% confidence interval). Thus, the microarray experiments identified 66/71 (93%) of these positive control genes as germline enriched. Of the five genes that were not identified, four were not analyzed due to poor spotting on the array, and one, *air-2*, had prior data on protein levels but not RNA levels (Schumacher et al., 1998). In addition to these positive controls, we used PCR screening of cDNA libraries made from *fem-1* and *fem-3* mutants to test the expression pattern of 25 genes that were predicted to be sperm enriched by the microarray results. We found that 21 genes were sperm enriched and that the remaining four genes were not detected in either library (H. E. S., data not shown). Together, these results provide strong evidence that most of the other genes identified in our microarray experiments are germline enriched.

Germline Gene Expression Patterns during Development

To investigate the kinetics of germline gene expression during development, we used a hierarchical clustering algorithm (Eisen et al., 1998) to sort the germline-enriched genes according to their similarity in expression pattern during the wild-type/*glp-4* developmental time course (Figure 3A). The hierarchical clustering sorted the genes into two main nodes. Almost all of the genes in the first node (cluster A) have peak wild-type expression levels during the L4 stage, while almost all of the genes in the second node (cluster B) have peak wild-type expression in the adult. Almost all of the sperm-enriched genes (95%) are in cluster A, while almost all of the oocyte-enriched and intrinsic genes (96% and 98%, respectively) are in cluster B (Figure 3B). The hermaphrodite germline makes sperm for a short time in the L4 stage before switching to oogenesis as an adult (Schedl, 1997), and this switch is exemplified in the expression patterns of the germline genes during germline development. The majority (61%) of sperm transcript levels decrease sharply between the L4 and the adult stages, since spermatocyte transcripts are degraded once the haploid sperm are formed (Kimble and Ward, 1988).

Gene Expression during Spermatogenesis

To display the expression pattern of the entire complement of sperm-enriched genes, we used the hierarchical clustering method to group the 650 sperm-enriched genes based on their expression in the wild-type/*glp-4* time course and also the *fem-1/fem-3* experiment (see http://cmgm.stanford.edu/~kimlab/germline). We sorted

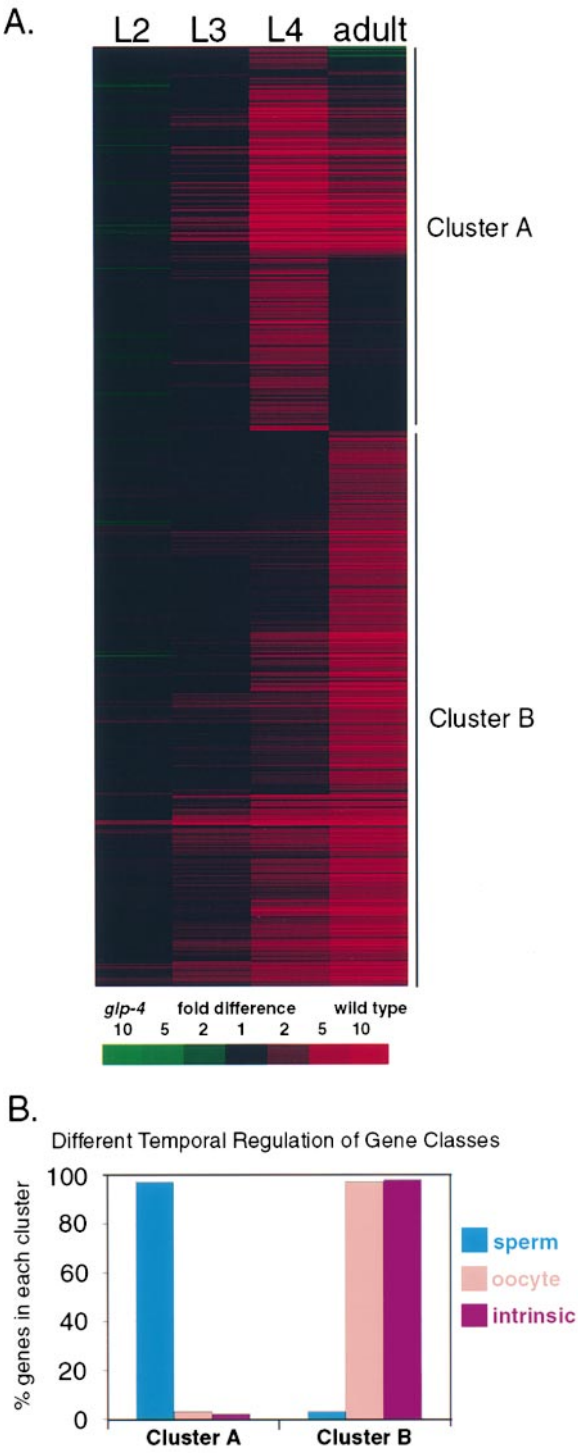

Figure 3. Expression Pattern of Germline Genes during Larval Development

(A) Hierarchical clustering was used to display the expression ratios of the germline-enriched genes in the wild-type/*glp-4* experiments at different stages during development. Included in the cluster were genes with a difference that was 99.9% significant in the *fem-3(gf)/fem-1(lf)* comparison and had a difference of at least 95% significance in one or more stages of the wild-type/*glp-4* comparison, as well as those genes with a 99.9% significant difference in the wild-type/*glp-4* comparison that did not have a 99.9% significant difference in the *fem-3(gf)/fem-1(lf)* comparison. Therefore, of the 1416 germline-enriched genes, 203 genes differentially expressed in the *fem-1/fem-3* experiment but not in the wild-type/*glp-4* experiment are not included in the cluster. Rows represent genes, and columns represent different stages of development. Levels of expression are shown by color and intensity, according to the key. A fully annotated version of this figure can be found at http://cmgm.stanford.edu/~kimlab/germline. In Cluster B, intrinsic genes with adult-only expression probably represent genes with low levels of expression that require a full-sized germline before they are detectable on the array.

(B) Percent of the sperm-enriched, oocyte-enriched, or germline-intrinsic genes that are in each of the clusters is shown.

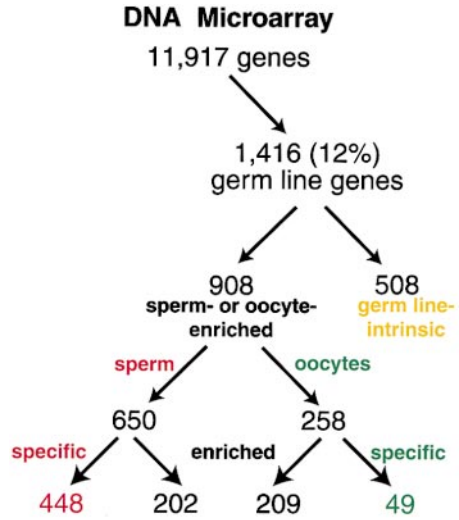

Figure 4. Summary of Gene Classes

The division between sperm-specific and sperm-enriched genes, as well as oocyte-specific and oocyte-enriched genes, is shown. The numbers represent the quantity of genes in each category, which are listed at http://cmgm.stanford.edu/~kimlab/germline.

the sperm-enriched genes by identifying those that appeared to be sperm specific, defined as those expressed in *fem-3(gf)* mutants, but not above background in *fem-1(lf)* mutants. Background levels in the *fem-1(lf)* hybridizations were determined by measuring the spot intensity of 43 genes that were previously known to be sperm specific (see Experimental Procedures). Four hundred forty-eight (69%) genes were sperm specific, while 202 (31%) of the sperm genes showed detectable expression in the soma or oocytes (Figure 4). In contrast, only 49 of the 258 oocyte-enriched genes (19%) appeared to be oocyte specific using these same criteria. These results support the idea that sperm are highly specialized cells that require many functions unique to sperm.

To identify cellular functions that might be critical for proper development of sperm, we determined which functional classes of genes were overexpressed among sperm-enriched genes (Table 1). Among the 650 sperm-enriched genes, genes encoding protein kinases were more than three times as abundant as expected from a random distribution, and protein phosphatases were nearly nine times as abundant, representing nearly half the protein phosphatases in the genome (Table 1). Both serine/threonine and tyrosine protein kinases are represented, and the vast majority are predicted to be cytoplasmic kinases rather than receptor kinases. Likewise,

Table 1. Overrepresented Functional Classes

Functional Class	Total Number on Array	Number of Sperm (% of Total)	Number of Oocyte (%)	Number of Intrinsic (%)
All	11917	650 (5.45%)	258 (2.16%)	508 (4.26%)
Protein kinase	395	71 (18.5%)	12 (3.04%)	23 (5.82%)
Protein phosphatase	106	53 (50%)	2 (1.89%)	7 (6.60%)
β-oxidation	63	10 (15.9%)	0 (0%)	1 (1.59%)
Chitin synthesis	5	0 (0%)	4 (80%)	1 (20%)
DNA replication	28	0 (0%)	10 (35.7%)	9 (32.1%)
RNA regulation	122	7 (5.73%)	13 (10.7%)	52 (42.6%)
Chromatin/chromosome	59	1 (1.7%)	4 (6.7%)	18 (30.5%)

both serine/threonine and tyrosine protein phosphatases are present. Because much of sperm development is regulated posttranslationally (see Discussion), these microarray results suggest that kinases and phosphatases are important components of that regulation. A second class of genes overrepresented in the sperm-enriched group encode nearly all the enzymes required for β-oxidation of fatty acids (acetyl-coA synthetase, ligase, lyase, dehydrogenase, and reductase). The expression of these enzymes in sperm likely reflects the high energy requirement for sperm motility.

Gene Expression in Oocytes

The expression patterns of the entire complement of 258 oocyte-enriched genes in our experiments are displayed at http://cmgm.stanford.edu/~kimlab/germline. The oocyte-enriched genes delineated in these experiments provide a molecular framework for gene functions required for both oogenesis and early embryogenesis (Table 1 and Figure 5). At least five oocyte-enriched genes are likely to contribute to eggshell formation, since they encode proteins similar to components of a chitin synthesis pathway, which functions in cell wall synthesis in *S. cerevisiae*.

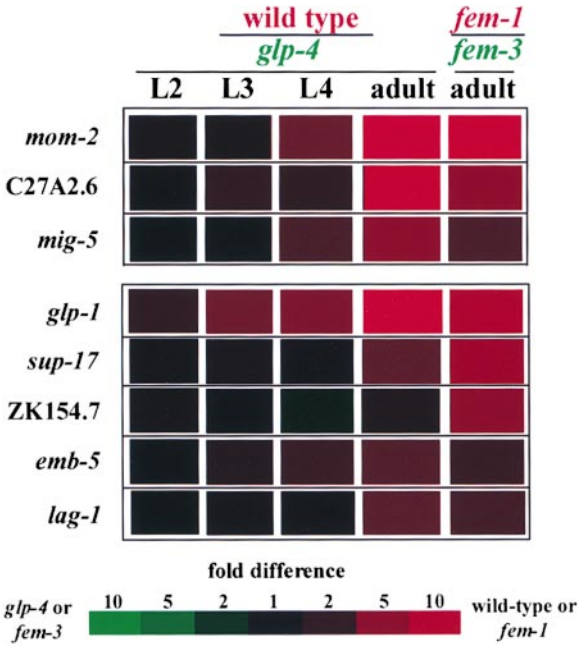

Figure 5. Oocyte-Enriched Genes

The expression ratios of the oocyte-enriched genes in the *glp-1* Notch and *mom-2* Wnt pathways are shown. Levels of expression are shown by color and intensity, according to the key. Three components of *mom-2* Wnt signaling, *mom-1*, *mom-5*, and *wrm-1*, were not significantly oocyte enriched, most likely because they are expressed in somatic tissues.

The oocyte-enriched genes include the maternal RNAs that drive the first few rounds of cell division before the onset of transcription in the zygote. We find that maternal transcripts encoding proteins predicted to function in DNA replication (ten genes) are overrepresented in the oocyte-enriched class of genes. Maternal transcripts are also required for proper patterning of the embryonic body axis prior to the onset of zygotic transcription. Embryonic patterning involves polarization of the EMS blastomere by the *mom-2* Wnt signaling pathway (Rocheleau et al., 1997; Thorpe et al., 1997). The *dishevelled* gene is a conserved member of the Wnt signaling pathway in flies, frogs, and mammals (reviewed in Dierick and Bejsovec, 1999). However, genetic experiments in *C. elegans* have not yet identified a *dishevelled*-like gene acting in the *mom-2* Wnt signaling pathway. Two genes (*mig-5* and C27A2.6) encode proteins with sequence similarity to Dishevelled and are oocyte enriched in our microarray experiments (Figure 5), suggesting that these genes may act redundantly in the *mom-2* Wnt pathway in early embryos.

The *glp-1* Notch signaling pathway specifies the proper fates of the ABp blastomere at the 4-cell stage (Priess et al., 1987) and the descendants of ABa at the 12-cell stage (Mello et al., 1994). The list of oocyte-enriched genes includes two genes previously known to act in early embryonic *glp-1* signaling: *glp-1* Notch and *lag-1* Su(H) (Christensen et al., 1996). In addition, two genes previously known to act in *glp-1* or *lin-12* Notch signaling at other times in development are oocyte enriched: *emb-5* and *sup-17* Kuzbanian (Hubbard et al., 1996; Tax et al, 1997; Wen et al., 1997). Kuzbanian is a *Drosophila* protease that is required for proper interactions between the Notch receptor and its ligand (Pan and Rubin, 1997). ZK154.7 is another gene similar to Kuzbanian that is oocyte enriched, suggesting that ZK154.7 might act redundantly with SUP-17 in the early embryo (Figure 5).

Germline-Intrinsic Gene Expression

The 508 germline-intrinsic genes may be involved in germline functions common to both sperm and oocytes, such as meiosis and recombination, stem cell proliferation, and germline determination (see http://cmgm.stanford.edu/~kimlab/germline). We can identify genes involved in meiosis and recombination by combining protein sequence information with gene expression data. For instance, some proteins in the meiosis-specific synaptonemal complex structure contain a coiled-coil

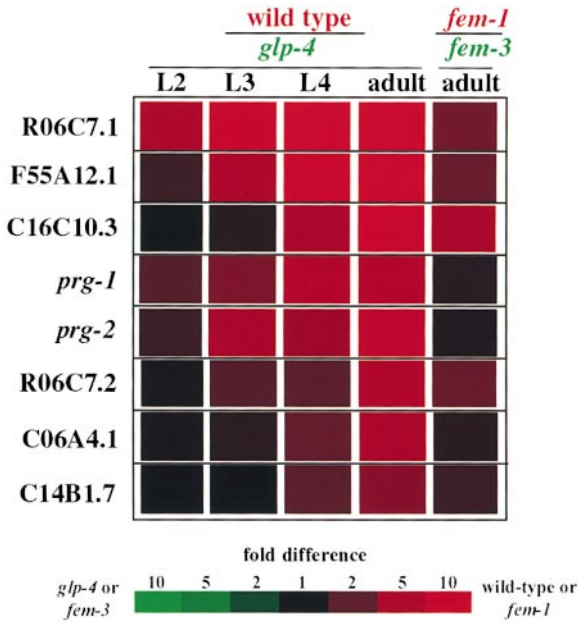

Figure 6. Germline-Intrinsic Genes

The expression ratios of the germline-intrinsic *piwi*-related genes are presented. Levels of expression are shown by color and intensity, according to the key.

domain, which mediates protein–protein interactions in proteins of diverse functions (e.g., Meuwissen et al., 1992; Sym et al., 1993). Using the multicoil algorithm (Wolf et al., 1997), we identified 446 genes out of the 11,917 on the DNA microarray with a probability greater than 50% of encoding a protein containing a coiled-coil domain. Only 24 of the 446 are germline-intrinsic genes, and these 24 are candidates for genes that encode components of the synaptonemal complex (Figure 6). Preliminary studies using RNA interference suggest that at least two of these genes function in chromosome synapsis (Reddy et al., unpublished data).

Included in the germline-intrinsic class are genes that potentially function in the stem cells, which divide mitotically and populate the germline with germ cells. Germline stem cells are present in the L2 and L3 stages, whereas mature gametes do not appear until the L4 and adult stages (Schedl, 1997). Among the 508 germline intrinsic genes, 72 begin to be differentially expressed in the L2 and L3 stages (at the 99% confidence interval) and thus may be important for germline stem cell function. Among these early-expressed germline-intrinsic genes, we noticed four genes related to the *Drosophila piwi* gene. *piwi* is required for germline stem cell maintenance in *Drosophila* (Cox et al., 1998), and a *piwi*-related gene (ZWILLE) is required for the maintenance of stem cells in *Arabidopsis* (Moussian et al., 1998). Of the 18 *piwi*-related genes present on the DNA microarrays, five are germline enriched at greater than 99.9% confidence (ranging from 3.6- to 28.9-fold at the adult stage), and three are germline enriched at greater than 99% confidence (2.4- to 3.7-fold) (Figure 6B). The two *C. elegans* genes most closely related to *piwi*, *prg-1* and *prg-2* (Cox et al., 1998), are only mildly enriched in the germline (3- to 4-fold), whereas other *piwi*-related genes such as R06C7.1 are highly germline-enriched (28.9-fold).

Control of gene activity at the posttranscriptional level mediates both germline specification and gamete sex determination. RNA-binding proteins bind target transcripts and control their localization, stability, or availability to the translational apparatus. The importance of RNA regulation in the germline is underscored by the large number of germline-intrinsic genes (52) that are involved in controlling RNA metabolism (Table 1 and Figure 6). Among these, 24 encode predicted RNA-binding proteins, including *pgl-1* and *gld-1,* two known germline RNA-binding proteins. FBF is an RNA-binding protein important for translational repression of target transcripts in the germline (Zhang et al., 1997). Three germline-intrinsic genes, F54C9.8, B0273.2, and Y45F10A.2, encode proteins similar to Pumilio and FBF, suggesting that these genes may also be involved in regulating germline translation. RNA helicases have been demonstrated to be components of germ granules; *glh-1* and *glh-2* in *C. elegans* encode two predicted RNA helicases similar to *vasa* in *Drosophila*, and all three have been demonstrated to affect germline fertility (Lasko and Ashburner, 1988; Gruidl et al., 1996). Although *glh-1* and *glh-2* were not present on our arrays, there were four other germline-enriched genes that encode RNA helicases, and these genes might encode protein components of P granules in *C. elegans*. Finally, many of the target transcripts of these translational control proteins should be present in our list of germline-enriched genes.

In addition to the RNA-binding proteins discussed above, 19 of the germline-intrinsic genes are likely to function in mRNA processing and splicing, and 12 may mediate ribosomal RNA processing. Controlling chromatin conformation is important for maintaining the germline; mutations in genes that affect germline chromatin, such as *mes-2*, which encodes a Polycomb group protein, result in degeneration of the germline (Holdeman et al., 1998). Genes that encode proteins involved in chromatin or chromosome structure are overrepresented among the germline-intrinsic genes relative to the rest of the genome (Table 1).

Gene Expression Differences in Sperm and Oocytes

Our experiments have identified the 650 most highly regulated sperm-enriched and 258 most highly regulated oocyte-enriched genes regulated at the 99.9% confidence level. In addition, many more genes are likely to differ in abundance but at a lower level, such that they are not included in the 99.9% confidence interval. To estimate the number of such genes, we used a scatter plot to compare the level of regulation for each gene for each pair of independent *fem-1/fem-3* microarray experiments. The five sperm/oocyte microarray experiments can be compared in ten pairwise combinations, and one representative comparison is shown in Figure 7A. Oocyte-enriched genes appear in the upper right quadrant, whereas sperm-enriched genes appear in the lower left quadrant. Genes expressed equally in *fem-1* and *fem-3* are distributed at random around the origin, appearing equally in all four quadrants. We calculated the number of genes that are randomly scattered about the origin and found an average of approximately 7770 genes (see Experimental Procedures). The oocyte-enriched and the sperm-enriched quadrants contain approximately 4150 genes in excess over the number that would be predicted to appear due to random scatter. This result indicates that there are a large number of

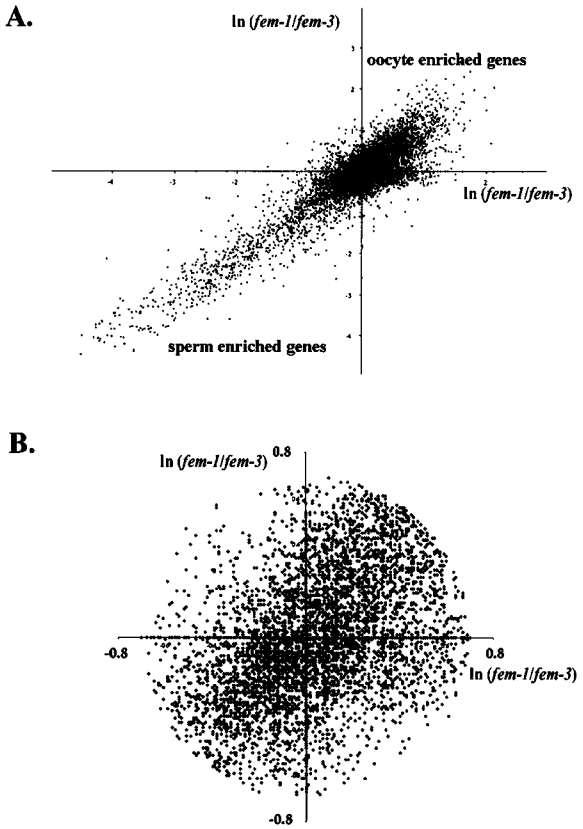

Figure 7. Gene Expression Differences in Sperm and Oocytes

(A) Scatter plot of representative comparison between two *fem-1/fem-3* experiments. For each experiment, the natural logarithm of the *fem-1/fem-3* ratio is plotted. RNAs expressed at different levels in the *fem-1* and *fem-3* strains appear along a diagonal axis. RNAs that are not expressed at different levels appear randomly scattered around the origin. There are approximately 3885 genes present in the upper left and lower right quadrant from all five experiments, approximately 8035 genes present in the upper right and lower left quadrant, and thus approximately 4150 more genes than would be predicted by random scatter.

(B) Scatter plot of RNAs that differ by less than 2-fold between the same two *fem-1/fem-3* experiments shown in (A), showing nonrandom scatter near the origin. There are approximately 1600 more genes in the upper right and lower left quadrants than would be predicted by random scatter.

differences in RNA expression between sperm and oocytes (about 34% of all of the genes on the microarray). Furthermore, many of the 4150 differentially expressed RNAs are regulated by only a slight amount; about 1600 differ by less than 2-fold in the *fem-3(gf)* sperm and *fem-1(lf)* oocyte RNA samples (Figure 7B).

Lack of Sperm-Enriched and Germline-Intrinsic Genes on the X Chromosome

By examining the chromosomal locations of all the germline-enriched genes, we discovered a striking bias in the distribution of these genes throughout the genome (Figure 8). Almost none of the germline-intrinsic or sperm-enriched genes reside on the X chromosome, but the oocyte-enriched genes are equally represented on the X chromosome and the autosomes. If the 650 sperm-enriched genes were distributed evenly throughout the genome, 102 would be expected to be X linked, and we found only six. Five of these X-linked sperm-enriched genes are at least 90% identical at the nucleotide level to a gene sequence located on an autosome. To determine whether any of these X-linked sperm-enriched genes are false positives due to cross-hybridization with autosomal genes, we designed gene-specific primers for four of the six (the remaining two genes did not contain any unique gene-specific sequences). We then used these primers in RT–PCR experiments and confirmed that two genes were sperm-enriched (F48E3.8 and F41G4.4), but that two genes (C04H5.1 and F11A1.1) were not expressed in *fem-3(gf)* worms and so are likely to be false positives (H. E. S. and J. N., data not shown).

Likewise, the 508 germline-intrinsic genes were underrepresented on the X chromosome; if the genes were randomly distributed, we would expect about 95 genes to be on the X chromosome, and we found only four. Three of the genes show a high degree of similarity to an autosomal gene. We used gene-specific primers in RT–PCR experiments for all four X-linked genes, and found that two were germline enriched (F08F1.9 and F16H11.3) and that two (C52B11.4 and C40H5.6) were false positives (H. E. S. and J. N., data not shown). Thus, of the ten X-linked germline-intrinsic and sperm-enriched genes identified by the DNA microarray experiments, we determined that at least four were false positives due to cross-hybridization, leaving no more than six on the X chromosome. This number is much less than the 197 genes than we would expect from a random distribution of genes on all of the chromosomes.

We expected about 40 oocyte genes to be on the X chromosome and found 36 (90%), distributed fairly evenly along the length of the X chromosome. Therefore, in contrast to the sperm-enriched and germline-intrinsic genes, the oocyte-enriched genes are evenly distributed on all chromosomes.

Discussion

Similar to a saturation genetic screen to identify genes based on mutant phenotypes, we have defined the component parts of a biological system based on gene expression patterns. By scanning 63% of the genome using DNA microarrays to compare gene expression between wild-type animals and three strains defective in some aspect of germline development, we have identified 1416 germline-enriched genes at the 99.9% confidence level, which corresponds to gene expression ratios with a range of 1.6- to 104-fold. These genes were further subdivided into those that exhibit sperm-enriched (650 genes), oocyte-enriched (258 genes), and germline-intrinsic (508 genes) expression patterns. This list forms a partial description of the germline at the molecular genetic level and contains genes likely to account for many germline functions, including the ability to undergo unrestricted proliferation, to generate all the tissues in a new individual, to control genetic heredity via meiosis and recombination, to form spermatozoa and oocytes, and to regulate early embryonic patterning with maternal transcripts.

Several observations validate the use of DNA microarrays to analyze germline gene expression patterns. First, the DNA microarray results are reproducible using multiple independent preparations of RNA. Second, 66 out

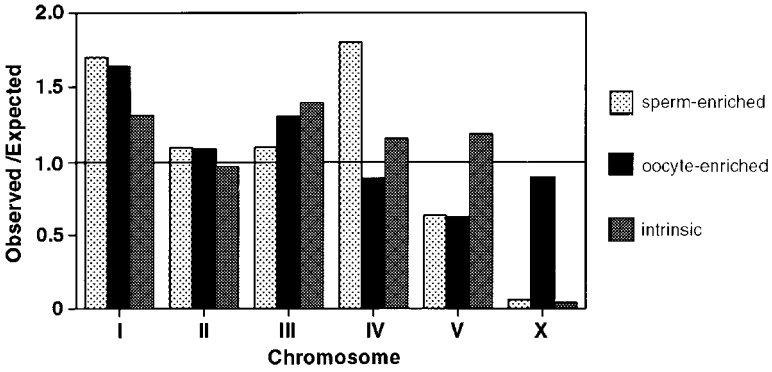

Figure 8. Intrinsic and Sperm-Enriched Genes Are Excluded from the X Chromosome

The number of germline-enriched genes expected to be on each chromosome, given the number of genes per chromosome on the array, was determined for oocyte-enriched, sperm-enriched, and germline-intrinsic genes. The ratio of observed to expected genes for each class of genes was plotted for each chromosome.

of 71 genes previously known to be expressed in the germline were also identified in the microarray experiments. Third, we independently tested 25 sperm-enriched genes identified by the microarray experiments by PCR screening of *fem-1(lf)* and *fem-3(gf)* cDNA libraries and confirmed that 21 were sperm enriched (expression from the remaining four was not detected).

Although nearly all of the 1416 genes identified in these microarray experiments are likely to be germline enriched, the list is not complete since the microarrays do not contain all of the genes in the genome. In addition, some genes may be expressed at levels that are too low to be measured by the DNA microarrays, although the germline comprises over 50% of the mass of the adult. For genes expressed in both the germline and the soma, changes in germline expression might not be detected due to somatic expression. In addition to missing some genes, the list of germline-enriched genes might also contain false positives. Some gene family members share high sequence identity and could cross-hybridize. Also, somatic genes might appear to be germline enriched if the germline induces the soma to express certain genes; these somatic genes would appear germline enriched as they would not be expressed in the *glp-4* mutant. However, we examined yolk transcripts, which are normally transcribed and translated in the intestine and transported into the germline, and determined that yolk transcripts are equally expressed in the *glp-4* mutant as they are in wild type (data not shown).

Germline Expression Patterns

In addition to the 43 sperm-specific genes on the array that were previously known, our experiments identified over 600 genes with sperm-enriched expression. The large number of protein kinases and protein phosphatases among these genes might reflect the unique properties of the developing *C. elegans* sperm. Because ribosomes are discarded prior to terminal cellular differentiation, the maturation of nonmotile symmetrical spermatids into crawling spermatozoa occurs in the absence of new protein synthesis. Phosphorylation and dephosphorylation could provide the means to modulate protein activities and promote the radical rearrangement of preexisting cellular components necessary for pseudopod extension and movement.

In hermaphrodites, sperm must crawl continuously in the spermatheca to fertilize incoming oocytes. Since fatty acids provide an efficient energy source, enhanced expression of β-oxidation genes in the sperm may provide energy for crawling. Interestingly, we found that sperm express a high level of *sod-2*, which encodes a mitochondrial superoxide dismutase (Hunter et al., 1997). Since β-oxidation of fatty acids produces free radicals, increased SOD-2 activity might protect sperm against oxidative damage that results from high metabolic activity, ensuring continued mitochondrial function and protecting the germline DNA from mutation.

The oocyte contains maternal RNAs that pattern the embryo prior to the onset of transcription. The *glp-1* Notch signaling pathway specifies the fates of anterior blastomeres. The *mom-2* Wnt signaling pathway polarizes a posterior blastomere, permitting it to undergo an asymmetric cell division. In addition to identifying known components in these early embryonic signaling pathways, the microarray experiments identify three genes that might act in the *glp-1* (*emb-5*, *sup-17*, and ZK154.7) and two genes that might act in the *mom-2* (*mig-5* and C27A2.6) embryonic pathways. The microarray results identified two genes related to Kuzbanian (*sup-17* and ZK154.7) and two related to Dishevelled (*mig-5* and C27A2.6). An important advantage of using microarrays to screen for genes based on expression patterns is that this approach can identify genes with partially redundant functions. In contrast, such genes may elude genetic screens, as mutations in only a single gene may not cause a mutant phenotype.

Germline stem cells proliferate in the adult when all somatic cells are quiescent, and they are totipotent as they generate all the cells of the next generation. Genes that function in stem cells should have a germline intrinsic expression pattern and be expressed early in germline development. Several representatives of a large protein family related to *Drosophila piwi* have an intrinsic, early-onset expression pattern. Piwi-related proteins are large, novel nuclear proteins required for maintenance of stem cells in *Drosophila*, *Arabidopsis*, and *C. elegans*. In *C. elegans*, two *piwi*-related genes (*prg-1* and *prg-2*) were previously demonstrated to function in germline proliferation (Cox et al., 1998). We found that eight of 18 *piwi*-related genes (including *prg-1* and *prg-2*) have germline-enriched expression patterns. These results suggest that an entire family of *piwi*-related genes are involved in either establishing or maintaining the germline of *C. elegans.*

The germline-intrinsic set of 508 genes includes eight genes previously known to function in meiosis and recombination, and preliminary results have already shown that the list of germline-intrinsic genes has identified genes involved in segregation of homologous chromosomes during meiosis. RNA interference experiments

indicate that functional depletion of two germline-intrinsic genes that encode proteins with coiled-coil domains results in a defect in homologous chromosome synapsis (Reddy et al., unpublished data).

Some aspects of germ cell differentiation and function are regulated by RNA-binding proteins that may control translation of target mRNAs (reviewed by Goodwin and Evans, 1997). Twenty-four germline-intrinsic genes encode RNA-binding proteins: *gld-1*, *pgl-1,* and 22 that have not been previously studied. RNA helicases are often components of germline P granules and may be important for germline specification (Lasko and Ashburner, 1988; Gruidl et al., 1996), and four germline-intrinsic genes encode RNA helicases. Gamete sex determination and differentiation are controlled by RNA binding proteins; in addition to *gld-1* and *pgl-1,* there are three germline-enriched genes that encode RNA-binding proteins related to the Pumilio/FBF family (Zhang et al., 1997).

The development of the germline is under tight spatial and temporal control by diverse types of regulatory pathways. RNA-binding proteins regulate RNA localization or translation. Signaling pathways such as the *glp-1* Notch and the *mpk-1* MAP kinase pathways control mitotic proliferation of germline stem cells and progression through pachytene during meiosis, respectively. DNA microarrays could be used to identify the downstream targets of these RNA control proteins or signaling pathways; the list of germline-enriched genes should include genes that are targets of the *glp-1* and *mpk-1* signaling pathways and genes that produce RNA transcripts bound by germline RNA-binding proteins.

Large Number of RNA Differences between Sperm and Oocytes

In addition to determining whether each gene is sperm or oocyte enriched, we were also able to analyze the entire set of genes to estimate the total number of transcripts that differ in abundance between sperm and oocytes. We compared the results of each individual sperm/oocyte microarray experiment to each of the others and found that approximately 34% of all the genes on the microarray differ between sperm and oocytes, many of which are predicted to differ by less than 2-fold. These results indicate that the differences between sperm and oocytes are not only characterized by strongly regulated genes, but also involve a large number of transcripts that may differ by only a small amount.

Germline-Intrinsic and Sperm-Enriched Genes Do Not Reside on the X Chromosome

Our DNA microarray and RT–PCR results also showed that there are almost no sperm-enriched and germline-intrinsic genes on the X chromosome. We identified at most four sperm-enriched genes and two germline-intrinsic genes on the X chromosome, which is much less than the 197 genes that would be expected from a random distribution given the chromosomal location of the genes on the microarray. In contrast, the oocyte-enriched genes are distributed at random on all six chromosomes.

What could explain the near absence of sperm-enriched and germline-intrinsic genes from the X chromosome? These genes would be expressed in the male germline in addition to the hermaphrodite germline, whereas the oocyte-enriched genes would be expressed primarily in the hermaphrodite germline. One possibility is that these two sets of genes may not reside on the X chromosome due to condensation of the X chromosome in the male germline. Serial reconstruction from electron micrographs of *C. elegans* meiotic nuclei suggested that the unpaired X chromosome in males forms a highly condensed chromosome early in pachytene, whereas the autosomal chromosomes all form synaptonemal complexes (Goldstein, 1982). If genes on this condensed X chromosome were expressed poorly, then selection would act to prevent the genes required for sperm formation from residing on this chromosome.

Another possible explanation is that the absence of sperm-enriched and germline-intrinsic genes from the X chromosome may be due to a lack of dosage compensation. In the soma, dosage compensation equalizes expression levels from the X chromosome in XO males and XX hermaphrodites (reviewed in Meyer, 1997). Dosage compensation may not be present in the germline, however, since proteins required for dosage compensation in the soma are either undetectable or not restricted to the germline X chromosomes (Lieb et al., 1996). If so, unequal levels of X-linked germline gene expression in XO males and XX hermaphrodites could result in improper germ cell function and hence reduced fitness. Reduced fitness would result in the apparent exclusion of genes expressed in both male and hermaphrodite germlines from the X chromosome.

The ancestral mode of sex determination in nematodes is most likely XX females and XO males, while hermaphroditism, such as in *C. elegans*, probably derived from this ancestral state (Triantaphyllou, 1983; Fitch and Thomas, 1997). One or both of the above possibilities could have acted in an ancestral male/female species to exclude the germline-intrinsic genes from the X chromosome. However, chromosome condensation but not a lack of dosage compensation could act to exclude the sperm-specific genes from the X chromosome in the ancestral male/female species. In this ancestral species, a lack of dosage compensation alone would not prevent sperm-specific genes from residing on the X chromosome, since these genes would only function in the male. Therefore, the evolution of a hermaphrodite from an ancestral female would involve the removal of over 100 sperm-specific genes from the X chromosome.

X chromosome condensation likely occurred in the male of the ancestral male/female species, as X chromosome condensation has been observed in another nematode species evolutionarily distant from *C. elegans* (Goldstein and Moens, 1976). This observation suggests that the X chromosome may be condensed in all XO nematodes. Furthermore, condensation of unpaired sex chromosomes is common in insects and other organisms, suggesting that sex chromosome condensation is widespread in evolution (White, 1973). In an ancestral male/female nematode, chromosome condensation could act to exclude the sperm-specific genes from the X chromosome and hermaphroditism could emerge without requiring a selective force to remove a large number of X-linked sperm-specific genes.

Conclusion

Expression analysis on a whole-genome scale using DNA microarrays has provided an overview of the development and functioning of the *C. elegans* germline at the

molecular level. This overview can be used to compare global profiles of germline gene expression to those from other tissues and to search for interesting global control mechanisms of gene regulation. This microarray analysis has provided a list of most of the genes that participate in the biological processes that make the germline unique. This list can now be searched for genes that act in diverse germline functions, such as downstream targets of signaling pathways or RNA-binding proteins, new recombination and meiosis genes, or new genes with sperm-specific functions. The list can be used in positional cloning experiments to rapidly identify candidate transcripts that correspond to interesting germline genes. In addition, the list of germline genes can be used as the basis for future functional genomics experiments, such as efforts to determine the loss-of-function phenotype for all of the germline genes or to determine all of the binding interactions between the proteins encoded by these genes.

Experimental Procedures

See http://cmgm.stanford.edu/~kimlab/germline for a complete description of the experimental procedures.

Construction of DNA Microarrays

We designed primer pairs corresponding to 13,323 genes, and then used these primers to generate PCR fragments from genomic DNA that were 1–2 kb in length containing at least 700 bp of predicted protein coding sequence or which covered 90% or more of the predicted coding sequence of the gene. Primers were designed not to span known *C. elegans* repetitive elements (*C. elegans* Sequencing Consortium, 1998). Primer sequences were determined using PRIMER3 (Rozen and Skaletsky 2000). Of the PCR reactions, 11,990 (89%) yielded a single band of the predicted size. These PCR fragments of genomic DNA were used to prepare DNA microarrays by Rosetta Inpharmatics as previously described (DeRisi et al., 1997). Seventy-three DNA spots were excluded due to errors in producing the DNA microarrays, resulting in DNA microarrays containing 11,917 genes.

RNA Isolation, cDNA Synthesis, and Microarray Hybridization

Strains used: wild-type is *C. elegans* variety Bristol strain N2. Linkage group I: *glp-4(bn2)* (Beanan and Strome, 1992); Linkage group IV: *fem-3(q23gf)* (Barton et al., 1987) and *fem-1(hc17ts)* (Nelson et al., 1978).

The reference RNA used in the developmental time course is from a mixed stage population of wild-type worms grown at 20°C. Wild-type worms were grown at 20°C, and mutant worms were grown at 15°C in liquid culture (*glp-4*) or on peptone plates (*fem-1 and fem-3*). Worms were isolated by floating on sucrose, and then large quantities of embryos were prepared by treatment with sodium hypochlorite. For wild-type and *glp-4* worms, mid-L1 stage worms were obtained by hatching eggs in liquid culture without food. For all strains, the staged worms were grown at 25°C on nematode plates seeded with a lawn of bacteria. Worms were harvested at L2, L3, L4, and young adult stages for wild-type and *glp-4* isolations by washing the plates twice with M9 at 25°C and then suspending in 4 ml Trizol/ml worms. At each harvest, the stage of development of the synchronized population was verified using Nomarski microscopy to directly observe vulval development. Wild-type young adults were collected prior to the initiation of embryogenesis. *fem-1* and *fem-3* young adults were harvested off plates in M9 buffer, pelleted in 15% Ficoll, floated on 35% Ficoll, washed twice in M9 before freezing, and then suspended in 3 ml Trizol (GIBCO, Grand Island, NY) per 1 ml packed worms (Lewis and Fleming, 1995). Total RNA and poly(A)$^+$ RNA was isolated as described (http://cmgm.stanford.edu/~kimlab/germline).

Labeled cDNA probe for DNA microarray hybridizations was prepared from 5 μg of poly(A)$^+$ RNA as described (DeRisi et al., 1997; see website). Wild-type, *glp-4* and *fem-1* cDNAs were labeled with Cy3, whereas the reference and *fem-3* cDNAs were labeled with Cy5. Microarrays were hybridized for 20 hr as described (DeRisi et al., 1997), scanned using an Axon scanner, and the expression levels for each gene were determined using GenePix software.

Data Analysis

(See http://cmgm.stanford.edu/~kimlab/germline for a complete description). The natural log (ln) of the ratio of Cy3/Cy5 for each gene was determined for each microarray hybridization. Each experiment was repeated between three and five times, and the average level of regulation for each gene for each experiment was calculated as the mean natural logarithm of the Cy3/Cy5 ratios. To determine the average ln(wild-type/*glp-4*) for each gene, ln(*glp-4*/ref) was subtracted from ln(wild-type/ref). The standard error in our measurements was determined by measuring the error for each gene from repetitions of 15 different experiments involving 61 separate microarray hybridizations. Genes regulated at the 99.9% confidence level in the *fem-1/fem-3* or in the wild-type/*glp-4* experiments were determined using a Student's t test for one or two populations, respectively. Descriptions of the proteins encoded by each of the genes were provided by Proteome.

Characteristics of 43 genes previously determined to be sperm specific by differential Northern analysis or RT–PCR were used to establish criteria for defining sperm-specific genes. From this set, we calculated the mean background level $\pm$ SD in the three most similar *fem-1/fem-3* microarray hybridizations. Genes with an average *fem-1* expression level that was less than two standard deviations above background ($<$0.1 normalized gene expression units) and had an expression level of 0 in at least one experiment were defined as sperm specific. Equivalent criteria defined the oocyte-specific genes.

The number of RNAs that differ in the *fem-1/fem-3* microarray experiments is equal to the number of genes that are consistently higher or lower in expression in both experiments minus the expected number due to random chance. In Figure 7A, genes that appear in the upper left and lower right quadrants are due to random scatter, and there should be an equal number of genes in the remaining two quadrants. We counted the number of genes with increased expression in one experiment but decreased expression in the other experiment for each of the five experiments. We then calculated the number of genes that consistently have higher or lower expression in both experiments and subtracted the number predicted from random scatter. Finally, we averaged the results from all five repeats of the *fem-1/fem-3* experiment, yielding approximately 4150 genes that are non-randomly distributed in the upper-right and lower-left quadrants. This number is an estimate of the total number of genes expressed at different levels between the sperm and oocyte samples.

Acknowledgments

The authors would like to thank the Worm Genome Consortium for providing the sequence of the *C. elegans* genome, and Proteome for its rapid annotation. We also thank Research Genetics and Ron Davis for assisting in primer synthesis, and Ethan Wolf and John Newman for helping to identify *C. elegans* coiled-coil proteins. We appreciate the assistance of Rosetta Inpharmatics in producing the microarrays. We are grateful to Kevin White, Anne Villeneuve, and Abby Dernburg for helpful discussions and to the Kim lab for critical reading of the manuscript. We especially thank the entire community of worm germline researchers for their willingness to share unpublished data regarding possible germline genes. This work was supported by a National Institutes of Health (NIH) National Research Service Award fellowship (V. R.), NIH, National Center for Research Resources, Merck, and Rhone-Poulenc Rorer grants (S. K. K.), and NIH GM25243 (S. W.) and T32-CA09213 (J. N. and H. S.).

Received November 23, 1999; revised July 18, 2000.

References

Achanzar, W.E., and Ward, S. (1997). A nematode gene required for sperm vesicle fusion. J. Cell Sci. *110*, 1073-1081.

Austin, J., and Kimble, J. (1989). Transcript analysis of *glp-1* and

lin-12, homologous genes required for cell interactions during development of *C. elegans*. Cell *58*, 565-571.

Barton, M.K., Schedl, T., and Kimble, J. (1987). Gain-of-function mutations of *fem-3*, a sex-determination gene in *Caenorhabditis elegans*. Genetics *115*, 107-119.

Beanan, M., and Strome, S. (1992). Characterization of a germline proliferation mutantion in *Caenorhabditis elegans*. Development *116*, 755-766.

Boxem, M., Srinivasan, D.G., and van den Heuvel, S. (1999). The *Caenorhabditis elegans* gene *ncc-1* encodes a *cdc-2*-related kinase required for M phase in meiotic and mitotic cell divisions, but not for S phase. Development *126*, 2227-2239.

Burke, D.J., and Ward, S. (1983). Identification of a large multigene family encoding the major sperm protein of *Caenorhabditis elegans*. J. Mol. Biol. *171*, 1-19.

C. elegans Sequencing Consortium. (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. Science *282*, 2012-2018.

Chase, D., Serafina, C., Ashcroft, N., Kosinsky, M., Longo, D., Ferris, D.K., and Golden, A. (2000). The Polo-like kinase PLK-1 is required for nuclear envelope breakdown and the completion of meiosis in *Caenorhabditis elegans*. Genesis, *26*, 26-41.

Chen, P.J., Singal, A., Kimble, J., and Ellis, R.E. (2000). A novel member of the tob family of proteins controls sexual fate in *Caenorhabditis elegans* germ cells. Dev. Biol. *217*, 77-90.

Christensen, C., Kodoyianni, V., Bosenberg, M., Friedman, L., and Kimble, J. (1996). *lag-1*, a gene required for *lin-12* and *glp-1* signaling in *Caenorhabditis elegans*, is homologous to human CBF1 and *Drosophila* Su(H). Development *122*, 1373-1383.

Clark-Maguire, S. and Mains, P.E. (1994). *mei-1*, a gene required for meiotic spindle formation in *Caenorhabditis elegans*, is a member of a family of ATPases. Genetics *136*, 533-546.

Cox, D.N., Chao, A., Baker, J., Chang, L., Qiao, D., and Lin, H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. Genes Dev. *12*, 3715-3727.

Dernburg, A.F., McDonald, K., Moulder, G., Barstead, R., Dresser, M., and Villeneuve, A.M. (1998). Meiotic recombination in *C. elegans* initiates by a conserved mechanism and is dispensable for homologous chromosome synapsis. Cell *94*, 387-398.

Dierick, H., and Bejsovec, A. (1999). Cellular mechanisms of wingless/Wnt signal transduction. Curr. Top. Dev. Biol. *43*, 153-190.

DeRisi, J., Iyer, V.R., and Brown, P.O. (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. Science *278*, 680-686.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. USA *95*, 14863-14868.

Fitch, D.H.A., and Thomas, W.K. (1997). Evolution. In *C. elegans* II, D.L. Riddle, T. Blumenthal, B.J. Meyer, and J.R. Priess, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 815-850.

Goldstein, P. (1982). The synaptonemal complexes of *Caenorhabditis elegans*: pachytene karyotype analysis of male and hermaphrodite wild-type and *him* mutants. Chromosoma *86*, 577-593.

Goldstein, P. and Moens, P.B. (1976). Karyotype analysis of *Ascaris lumbricoides* var suum. Chromosoma *58*, 101-111.

Goodwin, E.B., and Evans, T.C. (1997). Translational control of development in *C. elegans*. Semin. Cell Dev. Biol *8*, 551-559.

Grant, B., and Hirsh, D. (1999). Receptor-mediated endocytosis in the *Caenorhabditis elegans* oocyte. Mol. Biol. Cell. *10*, 4311-4326.

Gruidl, M.E., Smith, P.A., Kuznicki, K.A., McCrone, J.S., Kirchner, J., Roussell, D.L., Strome, S., and Bennett, K.L. (1996). Multiple potential germ-line helicases are components of the germ-line-specific P granules of *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA *93*, 13837-13842.

Guedes, S. and Priess, J.R. (1997). The *C. elegans* MEX-1 protein is present in germline blastomeres and is a P granule component. Development *124*, 731-739.

Hill, D.P., Shakes, D.C., Ward, S., and Strome, S. (1989). A sperm-supplied product essential for initiation of normal embryogenesis in *Caenorhabditis elegans* is encoded by the paternal-effect embryonic-lethal gene, *spe-11*. Dev. Biol. *136*, 154-166.

Holdeman, R., Nehrt, S., and Strome, S. (1998). MES-2, a maternal protein essential for viability of the germline in *Caenorhabditis elegans*, is homologous to a *Drosophila* Polycomb group protein. Development *125*, 2457-2467.

Hubbard, E.J., Dong, Q., and Greenwald, I. (1996). Evidence for physical and functional association between EMB-5 and LIN-12 in *Caenorhabditis elegans*. Science *273*, 112-114.

Hunter, T., Bannister W.H., and Hunter G.J. (1997). Cloning, expression and characterization of two manganese superoxide dismutases from *Caenorhabditis elegans*. J. Biol. Chem. *272*, 28652-28659.

Jones, A.R., and Schedl, T. (1995). Mutations in *gld-1*, a female germ cell-specific tumor suppressor gene in *Caenorhabditis elegans*, affect a conserved domain also found in Src-associated protein Sam68. Genes Dev. *9*, 1491-1504.

Kawasaki, I., Shim, Y.-H., Kirchner, J., Kaminker, J., Wood, W.B., and Strome, S. (1998). PGL-1, a predicted RNA-binding component of germ granules, is essential for fertility in *C. elegans*. Cell *94*, 635-645.

Kimble, J., and Ward, S. (1988). Germ-line development and fertilization. In The Nematode *Caenorhabditis elegans*, W.B. Wood, ed. (Plainview, NY: Cold Spring Harbor Laboratory Press), pp. 191-214.

Kreutzer, M.A., Richards, J.P., De Silva-Udawatta, M.N., Temenak, J.J., Knoblich, J.A., Lehner, C.F., and Bennett, K.L. (1995). *Caenorhabditis elegans* cyclin A- and B-type genes: a cyclin A multigene family, an ancestral cyclin B3 and differential germline expression. J. Cell Sci. *108*, 2415-2424.

Lasko , P.F., and Ashburner, M. (1988). Posterior localization of vasa protein correlates with, but is not sufficient for, pole cell development. Genes Dev. *4*, 905-921.

Lewis, J.A., and Fleming, J.T. (1995). Basic culture methods. In *Caenorhabditis elegans*: Modern Biological Analysis of an Organism, H.F. Epstein and D.C. Shakes, eds. (San Diego: Academic Press, Inc.), pp. 3-30.

Lieb, J. D., Capowski, E.E., Meneely, P., and Meyer, B.J. (1996). DPY-26, a link between dosage compensation and meiotic chromosome segregation in the nematode. Science *274*, 1732-1736.

L'Hernault, S.W. (1997). Spermatogenesis. In *C. elegans* II, D.L. Riddle, T. Blumenthal, B.J. Meyer, and J.R. Priess, eds. (Plainview, NY: Cold Spring Harbor Laboratory Press), pp. 271-294.

L'Hernault, S.W., Benian, G.M., and Emmons, R.B. (1993). Genetic and molecular characterization of the *Caenorhabditis elegans* spermatogenesis-defective gene *spe-17*. Genetics *134*, 769-780.

Mello, C.C., Draper, B.W., and Priess, J.R. (1994). The maternal genes *apx-1* and *glp-1* and establishment of dorsal-ventral polarity in the early *C. elegans* embryo. Cell *77*, 95-106.

Meuwissen, R.L., Offenberg, H.H., Dietrich, A.J., Riesewijk, A., van Iersel, M., and Heyting C. (1992). A coiled-coil related protein specific for synapsed regions of meiotic prophase chromosomes. EMBO J. *11*, 5091-5100.

Meyer, B. (1997). Sex determination and X chromosome dosage compensation. In *C. elegans* II, D.L. Riddle, T. Blumenthal, B.J. Meyer, and J.R. Priess, eds. (Plainview, NY: Cold Spring Harbor Laboratory Press), pp. 209-240.

Moussian, B., Schoof, H., Haecker, A., Jurgens, G., and Laux, T. (1998). Role of the ZWILLE gene in the regulation of central shoot meristem cell fate during Arabidopsis embryogenesis. EMBO J. *17*, 1799-1809.

Nelson, G.A., Lew, K.K., and Ward, S. (1978). *Intersex*, a temperature-sensitive mutant of the nematode *C. elegans*. Dev. Biol. *66*, 386-409.

Pan, D., and Rubin, G.M. (1997). Kuzbanian controls proteolytic processing of Notch and mediates lateral inhibition during Drosophila and vertebrate neurogenesis. Cell *90*, 271-280.

Priess, J.R., Schnabel, H., and Schnabel, R. (1987). The *glp-1* locus and cellular interactions in early *C. elegans* embryos. Cell *51*, 601-611.

Rocheleau, C.E., Downs, W.D., Lin, R., Wittmann, C., Bei, Y., Cha, Y.-H., Ali, M., Priess, J.R., and Mello, C.C. (1997). Wnt signaling and an APC-related gene specify endoderm in early *C. elegans* embryos. Cell *90*, 707-716.

Rosenquist, T.A. and Kimble, J. (1988). Molecular cloning and transcript analysis of *fem-3*, a sex-determination gene in *Caenorhabditis elegans*. Genes Dev. *2*, 606-616.

Rozen, S., and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. Methods Cell Biol. *132*, 365-386.

Schedl, T. (1997). Developmental genetics of the germline. In *C. elegans* II, D.L. Riddle, T. Blumenthal, B.J. Meyer, and J.R. Priess, eds. (Plainview, NY: Cold Spring Harbor Laboratory Press), pp. 241-269.

Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science *270*, 467-470.

Schumacher, J.M., Golden, A., and Donovan, P.J. (1998). AIR-2: an Aurora/Ipl1-related protein kinase associated with chromosomes and midbody microtubules is required for polar body extrusion and cytokinesis in *Caenorhabditis elegans* embryos. J. Cell Biol. *143*, 1635-1646.

Shalon, D., Smith, S.J., and Brown, P.O. (1996). A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. Genome Res. *6*, 639-645.

Sym, M., Engebrecht, J.M., and Roeder, G.S. (1993). ZIP1 is a synaptonemal complex protein required for meiotic chromosome synapsis. Cell *72*, 365-378.

Tax, F.E., Thomas, J.H., Ferguson, E.L., and Horvitz, H.R. (1997). Identification and characterization of genes that interact with *lin-12* in *Caenorhabditis elegans*. Genetics *147*, 1675-1695.

Thorpe, C.J., Schlesinger, A., Carter, J.C., and Bowerman, B. (1997). Wnt signaling polarizes an early *C. elegans* blastomere to distinguish endoderm from mesoderm. Cell 90, 695-705.

Triantaphyllou, A.C. (1983). Cytogenetic aspects of nematode evolution. In Concepts in Nematode Systematics, A.R. Stone, H.M. Platt, and L.F. Khalil, eds. (London: Academic Press), pp. 55-71.

Varkey, J.P., Muhhad, A.D., Minniti, A.N., Do, B., and Ward, S. (1995). The *Caenorhabditis elegans* gene *spe-26* is necessary to form spermatids and encodes a protein similar to the actin-associated proteins kelch and scruin. Genes Dev. *9*, 1074-1086.

Ward, S., Burke, D.J., Sulston, J.E., Coulson, A.R., Albertson, D.G., Ammons, D., Klass, M., and Hogan, E. (1988). Genomic organization of major sperm protein genes and pseudogenes in the nematode *Caenorhabditis elegans*. J. Mol. Biol. *5*, 1-13.

Wen, C., Metzstein, M.M., and Greenwald, I. (1997). SUP-17, a *Caenorhabditis elegans* ADAM protein related to *Drosophila* KUZBANIAN, and its role in LIN-12/NOTCH signaling. Development *124*, 4759-4767.

White, M.J.D. (1973). Animal Cytology and Evolution, Third edition (Cambridge, U.K.: Cambridge University Press).

Wolf, E., Kim, P.S., and Berger, B. (1997). Multicoil: a program for predicting two- and three-stranded coiled coils. Protein Sci. *6*, 1179-1189.

Zetka, M.C., Kawasaki, I., Strome, S., and Muller, F. (1999). Synapsis and chiasma formation in *Caenorhabditis elegans* requires HIM-3, a meiotic chromosome core component that functions in chromosome segregation. Genes Dev. *13*, 2258-2270.

Zhang, B., Gallegos, M., Puoti, A., Durkin, E., Fields, S., Kimble, J., and Wickins, M.P. (1997). A conserved RNA-binding protein that regulates sexual fates in the *C. elegans* hermaphrodite germline. Nature *390*, 477-484.