

# The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation

Pavel P Khil<sup>1,3</sup>, Natalya A Smirnova<sup>1,3</sup>, Peter J Romanienko<sup>1,2</sup> & R Daniel Camerini-Otero<sup>1</sup>

**Sex chromosomes are subject to sex-specific selective evolutionary forces<sup>1,2</sup>. One model predicts that genes with sex-biased expression should be enriched on the X chromosome<sup>2–5</sup>. In agreement with Rice's hypothesis<sup>3</sup>, spermatogonial genes are over-represented on the X chromosome of mice<sup>6</sup> and sex- and reproduction-related genes are over-represented on the human X chromosome<sup>7,8</sup>. Male-biased genes are under-represented on the X chromosome in worms and flies<sup>9–11</sup>, however. Here we show that mouse spermatogenesis genes are relatively under-represented on the X chromosome and female-biased genes are enriched on it. We used *Spo11*<sup>-/-</sup> mice blocked in spermatogenesis early in meiosis<sup>12</sup> to evaluate the temporal pattern of gene expression in sperm development. Genes expressed before the *Spo11* block are enriched on the X chromosome, whereas those expressed later in spermatogenesis are depleted. Inactivation of the X chromosome in male meiosis may be a universal driving force for X-chromosome demasculinization.**

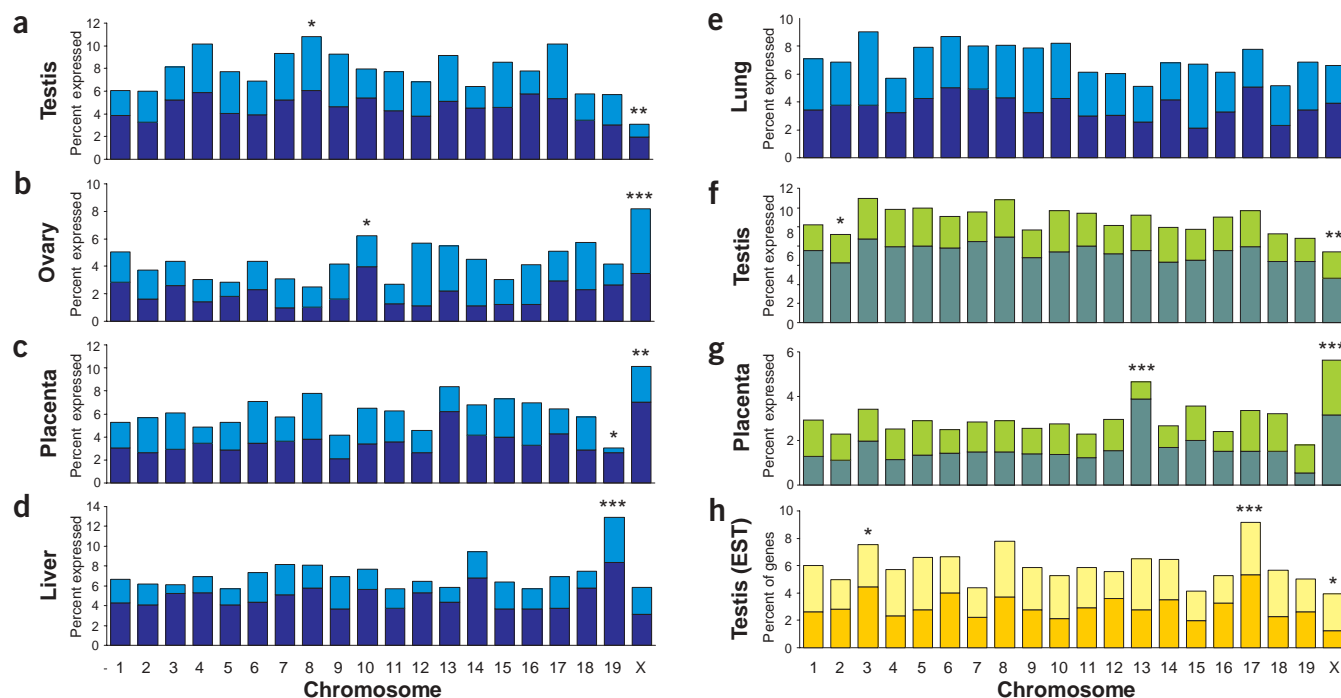
To investigate the genomic distribution of mouse genes with patterns of sex-biased expression, we analyzed the chromosomal distribution of genes expressed in sexually dimorphic tissues. We consider genes that are preferentially expressed in such tissues as testis, ovary and placenta to be sex-biased genes. We used two approaches to define the tissue specificity of genes expressed in mice: analysis of large-scale microarray-based gene expression data and analysis of the distribution of expressed-sequence tags (ESTs) in different cDNA libraries. Instead of focusing on transcripts restricted to a single tissue, our analysis targeted the identification of tissue-enriched transcripts using a 'preferential expression measure' (PEM)<sup>13</sup>. As a uniform source of expression data, we used publicly available data sets of mouse RNA samples representing 49 tissues hybridized to Affymetrix mouse U74A microarrays (GNF data set)<sup>14</sup> and 20 tissues analyzed with RIKEN cDNA microarrays (RIKEN data set)<sup>15</sup>.

In the GNF data set<sup>14</sup>, expression of 624 genes (~7% of all analyzed) in testis was three times greater than the median expression in other

tissues, and expression of 361 genes in testis was five times greater (PEM values >1.58 and >2.32, respectively). These genes are considered testis-enriched genes. The chromosomal distribution of testis-enriched genes was significantly different from uniform ( $\chi^2 = 30.8$ ,  $P = 0.05$ ; **Fig. 1a** and **Supplementary Table 1** online). The most significant deviation from the average gene density was on the X chromosome, where testis-enriched genes were 2.5 times less dense (the ratio of the density of testis-enriched genes on the X chromosome to the average density on all chromosomes,  $\rho_X$ , was 0.40;  $\chi^2 = 7.7$ ,  $P = 0.006$ ). The depletion of testis-enriched genes on the X chromosome was independent of the cut-off (three or five times greater expression) used in the selection process (**Table 1**). In addition, genes that were expressed most highly in testis and all the genes detected in testis were under-represented on the X chromosome (**Table 1**).

In contrast to testes-enriched genes, the chromosomal distribution of most tissue-enriched genes was not significantly different from random (**Fig. 1d,e**;  $\chi^2 = 23.1$  and 13.4 for liver- and lung-enriched genes, respectively;  $P > 0.1$ ). The chromosomal distribution of ovary-specific genes, however, was far from uniform ( $\chi^2 = 31.9$ ,  $P = 0.03$ ; **Fig. 1b**). Again, the chromosome with the most profound difference in gene density was the X chromosome, but unlike testis-enriched genes, ovary-enriched genes were more abundant on the X ( $\rho_X = 1.95$ ;  $\chi^2 = 10.58$ ,  $P = 0.001$ ; **Table 1**). We also investigated the chromosome distribution for genes preferentially expressed in placenta (**Fig. 1c**), which was markedly similar to that of ovary-enriched genes ( $\rho_X = 1.66$ ;  $\chi^2 = 7.24$ ,  $P = 0.007$ ). As with testis-enriched genes, the enrichment was independent of the cut-off expression level for both sets of female-biased genes (**Table 1**). In addition to tissue-enriched genes, genes that are highly expressed in placenta or ovary were also enriched on the X chromosome ( $\rho_X = 1.5$ ,  $P = 0.03$  and  $\rho_X = 1.7$ ,  $P = 0.01$ , respectively; **Table 1**). An analysis of the independent RIKEN set of microarray expression data<sup>15</sup> confirmed the trends observed with the GNF set: testis-enriched genes were under-represented on the X chromosome ( $\rho_X = 0.62$ ;  $\chi^2 = 7.92$ ,  $P = 0.005$ ) and placenta-enriched genes were over-represented ( $\rho_X = 2.12$ ;  $\chi^2 = 12.1$ ,  $P < 0.001$ ; **Fig. 1f,g**).

<sup>1</sup>Genetics and Biochemistry Branch, NIDDK, National Institutes of Health, 5 Memorial Drive, Bethesda, Maryland 20892, USA. <sup>2</sup>Present address: Developmental Biology Program, Mouse Genetics Core, Memorial Sloan-Kettering Cancer Center, 1275 York Ave, New York, New York 10021, USA. <sup>3</sup>These authors contributed equally to this work. Correspondence should be addressed to R.D.C.-O. (camerini@ncifcrf.gov).



**Figure 1** Sex-specific genes are differentially represented on the X chromosome. (a–e) The percentages of genes residing on individual chromosomes and satisfying given selection criteria at different stringency levels are shown. Dark blue, genes that are five times more abundant than median of all tissues; light blue, genes that are three times more abundant. (f,g) The chromosomal distribution of tissue-enriched genes based on an analysis of the RIKEN data set. Dark green, genes whose expression is five times greater than the median; light green, genes whose expression is three times greater. (h) The chromosomal distribution of transcripts enriched in testis as found in our EST analysis using R statistics. Dark yellow,  $R > 10$ , expression in testis is two times greater than median; light yellow,  $R > 20$ , expression in testis is two times greater. Chromosomes that have gene densities significantly deviating from the average are indicated (\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ).

To complement and extend the microarray expression data, we analyzed the distribution of ESTs in different cDNA libraries available in GenBank. We selected genes that satisfied two criteria: (i) a statistically significant preferential expression in testis and (ii) expression at least two times higher in testis than in other tissues. The chromosomal distribution of testis-enriched genes based on the EST data confirmed the under-representation of these genes on the X chromosome ( $\rho_X = 0.45$ ;  $\chi^2 = 5.0$ ,  $P = 0.02$ ; **Fig. 1h**).

The X chromosome is not the only chromosome with a skewed distribution of tissue-enriched genes. Analysis of both the GNF data set and EST data showed that testis-enriched genes are over-represented on chromosome 17. Approximately half (18 of 38) of the testis-enriched genes on chromosome 17 are located in the t-complex region that contains many male fertility genes<sup>16</sup>. In addition, we identified another cluster located centromeric to the t-complex (**Supplementary Table 2** online).

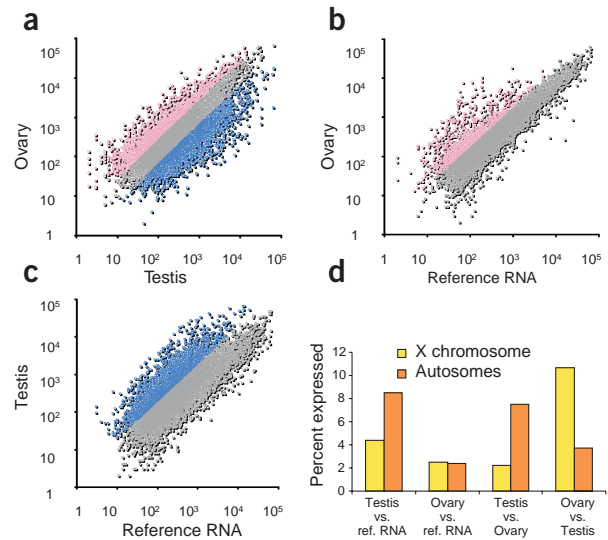
To further define testis- and ovary-enriched genes, we compared the RNAs expressed in mouse gonads using direct competitive hybridization on cDNA microarrays. We identified 749 genes whose expression was at least three times higher in testes than in ovaries and 404 genes that were preferentially expressed in ovaries compared with testes (**Fig. 2a**). Again we found that the X chromosome is a disfavored location for testis-enriched genes ( $\rho_X = 0.28$ ;  $P < 0.001$ ) but a preferred site for ovary-enriched genes ( $\rho_X = 2.88$ ;  $P < 0.001$ ; **Fig. 2d**). We also compared the RNAs expressed in testes and ovaries with a reference RNA population, which gives a measure of the preferential expression in a tissue that is similar to PEM. We found that the expression of 858 genes was at least three times higher in testes than in reference RNA and that 243 genes were overexpressed in ovaries (**Fig. 2b,c**). As in the direct comparison, testis-expressed genes were under-represented on the X chromosome ( $\rho_X = 0.50$ ;  $P = 0.01$ ; **Fig. 2d**). The distribution of

**Table 1** The proportion of genes on the X chromosome that are preferentially expressed in different tissues

Tissue	Two times higher		Three times higher		Five times higher		Highly expressed		Expressed in tissue	
	$\rho_X$	$P$	$\rho_X$	$P$	$\rho_X$	$P$	$\rho_X$	$P$	$\rho_X$	$P$
Testis	<b>0.38</b>	<b>&lt;0.001</b>	<b>0.40</b>	<b>0.006</b>	<b>0.42</b>	<b>0.046</b>	0.75	0.234	<b>0.73</b>	<b>0.004</b>
Ovary	<b>1.59</b>	<b>&lt;0.001</b>	<b>1.95</b>	<b>0.001</b>	1.91	0.063	<b>1.76</b>	<b>0.007</b>	1.03	0.725
Placenta	1.28	0.168	<b>1.65</b>	<b>0.007</b>	<b>2.06</b>	<b>0.002</b>	<b>1.54</b>	<b>0.047</b>	1.02	0.742
Lung	0.94	0.670	0.97	0.736	1.04	0.942	1.26	0.299	0.90	0.284
Liver	0.80	0.288	0.85	0.460	0.71	0.229	0.90	0.640	0.96	0.593

Selection criteria shown represent genes whose expression in that tissue is two, three or five times greater than the median expression level. Data for genes highly expressed in a tissue (~500 genes with the highest level of expression in a given tissue) and for all genes expressed in a tissue (above a cut-off level corresponding to three to five copies per cell) are also shown. Statistically significant results are indicated in bold.

**Figure 2** The X chromosome is enriched for genes expressed in female gonads but depleted of genes expressed in male gonads. Scatter plots of gene expression values (in arbitrary units) in testis versus ovary (a), ovary versus reference RNA (b) and testis versus reference RNA (c). Each gene is represented by a dot and the axes are the associated intensity values. Genes whose expression was at least three times greater in a given tissue are shown in pink (ovary-enriched) and blue (testis-enriched). (d) The percentages of expressed genes residing on the X chromosome and the autosomes.



genes enriched in ovaries, however, was closer to random ( $\rho_X = 1.10$ ).

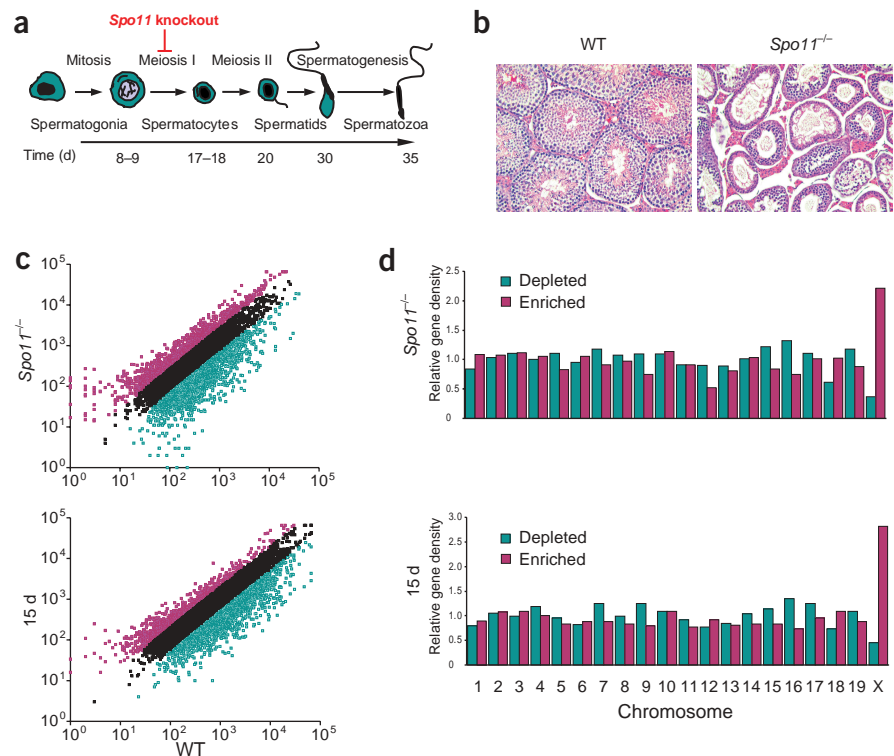
Our analysis of several independent data sets consistently identified deviations in the chromosomal distribution of sex-biased genes. The PEM values calculated for sexually dimorphic tissues from the different data sets were positively correlated on a gene-by-gene basis (Supplementary Fig. 1 online).

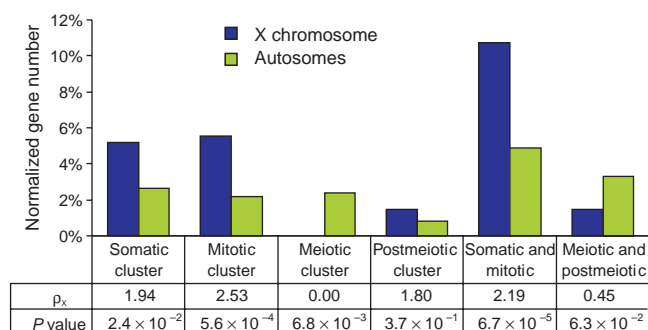
Testes contain more than 10 different cell types, most of which participate in spermatogenesis; over 90% are germ cells, most of which are past early meiosis (Fig. 3a)<sup>17,18</sup>. Sex chromosomes are condensed and transcriptionally repressed at the onset of male meiosis (meiotic sex chromosome inactivation, MSCI)<sup>19–21</sup>. To elucidate the effect of MSCI on the genomic distribution of male-biased genes, we analyzed the transcripts of testes from *Spo11*<sup>-/-</sup> mice which have a genetic ablation of cells past the early stages of spermatogenesis. The disruption of *Spo11* in mice blocks meiosis before pachytene and interrupts normal testis development (Fig. 3a,b)<sup>12,22</sup>. In agreement with the marked changes in testis morphology in *Spo11*<sup>-/-</sup> mice (Fig. 3b), the expression of almost 30% of the transcripts changed by a factor of more than two (Fig. 3c). Consistent with the relative proportion of meiotic and postmeiotic cells in testis, most of these testis-enriched transcripts were depleted in testes of mature *Spo11*<sup>-/-</sup> mice (Supplementary Fig. 2 online). Like genes encoding testis-enriched transcripts, genes encoding transcripts depleted in the testes of *Spo11*<sup>-/-</sup> mice were strongly under-represented on the X chromosome ( $\rho_X = 0.34$ ;  $\chi^2 = 12.9$ ,  $P <$

0.001; Fig. 3d). Next, we analyzed transcripts that were more abundant in testes of *Spo11*<sup>-/-</sup> mice than in testes of wild-type mice. Because testes of *Spo11*<sup>-/-</sup> mice are relatively enriched with premeiotic cells, the transcripts over-represented in these testes should include spermatogonial transcripts, reported to be over-represented on the X chromosome<sup>6</sup>. Like spermatogonial genes, *Spo11*<sup>-/-</sup>-enriched genes (which are not limited to spermatogonial genes) were significantly more dense on the X chromosome ( $\rho_X = 2.23$ ;  $\chi^2 = 47.8$ ,  $P < 0.001$ ; Fig. 3d).

The first wave of meiosis in male mice is synchronous and starts at birth<sup>17</sup> (Fig. 3a). Therefore, another approach to identify genes transcribed mostly in premeiotic cells is to work with testes from juvenile mice. At 15 d of age in wild-type mice, meiotic progression

**Figure 3** Genes expressed early and late in spermatogenesis are differentially represented on the X chromosome. (a) Time line of spermatogenesis. (b) Hematoxylin and eosin-stained cross-section of seminiferous tubules from wild-type (WT) and *Spo11*<sup>-/-</sup> mice. Sperm development proceeds towards the lumen with the earliest cells with most dense nuclei at the periphery and later cells including elongated spermatids near the center. The seminiferous tubules of *Spo11*<sup>-/-</sup> mice are devoid of all but the earliest cells. (c) Gene expression profiles in testes of *Spo11*<sup>-/-</sup> and juvenile wild-type (15 d) mice. Testis RNA samples from adult (8–10 weeks) wild-type, adult *Spo11*<sup>-/-</sup> and 15-d-old wild-type mice were competitively hybridized to cDNA microarrays in four independent experiments in the combinations indicated. Intensities beyond a two-fold cutoff are shown in red (enriched in *Spo11*<sup>-/-</sup> or juvenile testis) and green (depleted in *Spo11*<sup>-/-</sup> or juvenile testis). Most genes reported by Wang and coworkers<sup>6</sup> map to the group of ~1,000 genes enriched in testes of *Spo11*<sup>-/-</sup> mice (13 of 20, compared with only 2 of 20 that mapped to the group of genes depleted in testes of *Spo11*<sup>-/-</sup> mice). (d) The chromosomal distributions of differentially expressed genes.





**Figure 4** Genes expressed in rat spermatogonia or Sertoli cells are enriched on the X chromosome. Premeiotic, meiotic and postmeiotic germ cells and somatic Sertoli cells were purified from rat testes and the level of gene expression in individual cell fractions was measured using Affymetrix microarrays<sup>24</sup>. The percentage of genes on autosomes and on the X chromosome belonging to each cluster corresponding to a given cell population is plotted. Below the chart the relative gene density and the probability for observing such a distribution by chance are shown.

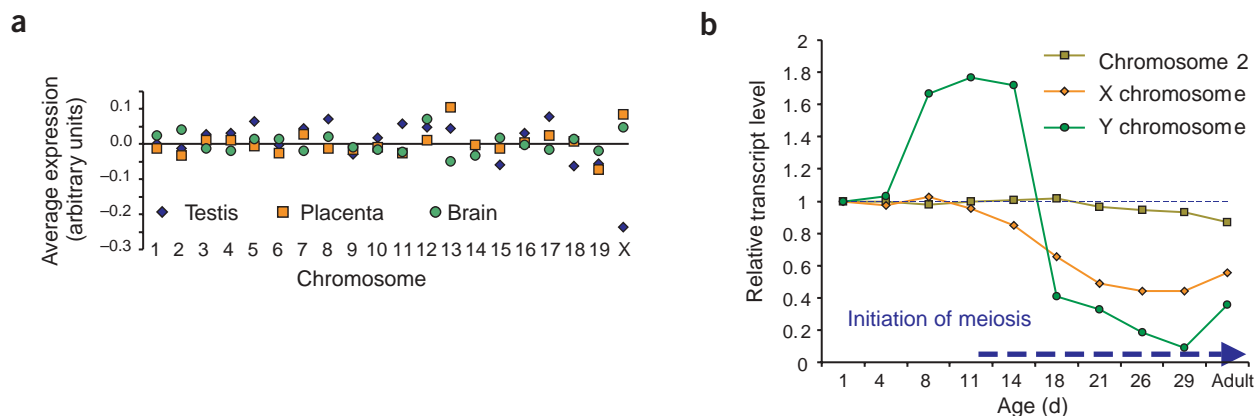
reaches approximately the stage of the block in *Spo11*<sup>-/-</sup> mice. We found that the expression profiles of 95% of the genes that were differentially expressed in testes of 15-d-old mice were similar to those of the *Spo11* mutant (**Supplementary Fig. 2** online). The chromosomal distribution of genes encoding transcripts that were enriched or depleted in juvenile mice followed the same trend as in the *Spo11*<sup>-/-</sup> mice (**Fig. 3d**;  $\rho_x = 0.44$ ,  $\chi^2 = 9.96$ ,  $P = 0.002$  and  $\rho_x = 3.2$ ,  $\chi^2 = 73.4$ ,  $P < 0.001$ , respectively). Thus, both genetic ablation studies and analysis of testes from juvenile mice indicate that early genes are relatively abundant but that late genes are sparse on the mouse X chromosome.

A similar dissection of gene expression in spermatogenesis is possible using gene expression profiles of developing testes in mice<sup>23</sup>. We compared expression in samples from 1-, 4-, 8- and 11-d-old mice with samples from 18–26-d-old mice and found that the chromosome distribution of these genes was more skewed than that of *Spo11*<sup>-/-</sup> mice. Genes expressed at the earlier stages were abundant on the X chromosome ( $\rho_x = 2.32$ ;  $\chi^2 = 125.9$ ,  $P < 0.001$ ) whereas late genes

were nearly absent ( $\rho_x = 0.08$ ;  $\chi^2 = 31.9$ ,  $P < 0.001$ ; **Supplementary Table 1** online). We examined the differential genomic distribution of spermatogenic genes using existing data on the expression of these genes in the rat<sup>24</sup>. This analysis showed that meiotic-specific genes were completely absent from the X chromosome, whereas genes expressed in spermatogonia were significantly more abundant on the X chromosome (**Fig. 4**). Genes expressed in somatic cells of testis were also more abundant on the X chromosome.

Next, we estimated the global effect of MSCI on transcription. Both the GNF and RIKEN data sets showed that genes located on the X chromosome had lower average expression in adult testes (**Fig. 5a** and **Supplementary Fig. 3** online). Exploration of the mouse spermatogenesis data set<sup>23</sup> showed that the decrease in net transcription from the X coincides with the onset of meiosis in developing testes (**Fig. 5b**). Although the X chromosome is inactivated in cells going through meiosis, several X-linked genes were reliably detected in more than one data set as testis-enriched (see **Supplementary Note** and **Supplementary Table 3** online).

In this study, we found that sex-biased genes had a nonrandom chromosomal distribution in the mouse. The accumulation of both male- and female-biased genes on the X chromosome is consistent with the predictions of W. Rice<sup>3</sup> and previous reports on mice and humans<sup>6–8</sup>. We believe that the paucity of testis-enriched genes on the mouse X is mainly attributable to the effects of MSCI<sup>19,20</sup>. The under-representation of male-biased genes on the mouse X chromosome is similar to that reported for male-specific genes in *Drosophila melanogaster* and *Caenorhabditis elegans*<sup>9,10</sup>, where (in *C. elegans*, at least) MSCI is present<sup>25,26</sup>. The evolutionary roots of MSCI are deep, and MSCI is tightly linked to the evolution of the sex chromosomes. MSCI probably preceded somatic X-chromosome inactivation and may be the origin of paternal X inactivation in marsupials and in preimplantation mouse embryos<sup>27</sup>. Thus, there seem to be two main counteracting forces in evolution that affect the distribution of male-biased genes. One force removes male-biased genes from the X chromosome (MSCI), and the other adds them (the predicted and observed accumulation of male beneficial genes on the X chromosome). We believe that this hypothesis most parsimoniously accounts for and reconciles the previously published data on this subject from flies, worms, mice and humans.



**Figure 5** X-linked genes have a lower average expression in testis. **(a)** Average expression level of genes residing on different chromosomes in testis, placenta and brain. The analysis was done on the RIKEN data set. **(b)** The transcriptional activity of chromosome 2 and the X and Y chromosomes during mouse spermatogenesis<sup>23</sup>. The average expression values were normalized to day 1 and plotted. An analysis of gene expression from other autosomes gave similar results to those obtained with chromosome 2.



## METHODS

**Determining tissue-specificity from microarray data.** We defined genes as tissue-enriched or preferentially expressed if their expression in that tissue was several times greater than the median expression level calculated for all other tissues (the PEM is essentially the  $\log_2$  relative excess over the median<sup>13</sup>). We obtained individual gene expression level data from data sets on hybridization intensity values of RNA samples from 49 different tissues on Affymetrix mouse U74A microarrays<sup>14</sup> (GNF data set) and 20 tissues analyzed with RIKEN cDNA microarrays<sup>15</sup> (RIKEN data set). The degree of correlation of PEM values from independent data sets is similar to that reported previously<sup>13</sup>. In the GNF data set, we defined genes expressed in a given tissue as those genes with an expression level higher than 200, estimated to correspond to three to five transcripts per cell<sup>14</sup>. We defined highly expressed genes in the GNF data set as those genes with an expression level higher than 1,200 in a tissue without regard to the expression in other tissues. We set this threshold so as to give ~500 genes in every tissue.

**EST analysis.** We counted the number of individual ESTs mapped to each Unigene cluster in two pools of cDNA libraries, one consisting of 7 testis cDNA libraries (pool A; 74,672 ESTs) and another a pool of 24 libraries from other tissues (pool B; 264,355 ESTs). The statistical likelihood analysis was based on a calculation of R statistics<sup>28</sup> for each gene. We selected the genes according to two criteria: (i) representation in testis cDNA libraries (pool A) was two times higher than representation in the other libraries (pool B) and (ii) the R likelihood score was above a preset limit.

**Analysis of chromosomal distributions.** We defined the gene density as the fraction of genes satisfying a given selection criterion of the total number of genes present on a chromosome (normalized gene number). We determined relative gene density ( $\rho_i$ ) by dividing the gene density on chromosome  $i$  by the average gene density on all chromosomes. We used the  $\chi^2$  test with 19 degrees of freedom for the genomic distribution as a whole or with one degree of freedom for individual chromosomes. We compared the observed number of genes on a given chromosome with the expected number based on a random distribution of all the genes in a data set.

**Microarray procedures.** All the animal procedures were carried out according to the National Institutes of Health Guide for the Care and Use of Laboratory Animals. We surgically removed testes and ovaries from mice and stored them at  $-80^\circ\text{C}$  until use. We isolated RNA using a Trizol solution (Invitrogen). We used a mixture of total RNAs from 11 different cell lines from several different tissues (universal reference RNA, Stratagene) as reference RNA. We synthesized amino-allyl-modified cDNA from 10–20  $\mu\text{g}$  of DNase-treated RNA using SuperScript II Reverse Transcriptase (Invitrogen), oligo(dT) (Invitrogen) and amino-allyl-modified dUTP (Sigma). We generated fluorescent probes by labeling amino-allyl-modified cDNA with Cy3 or Cy5 (Amersham). We carried out hybridizations on glass slide microarrays containing 15,000 genes from the NIA 15K mouse gene collection<sup>29</sup> (Keck Biotechnology Resource Lab at Yale University). All microarray experiments were done in quadruplicate with different mouse pairs and dye-reversed hybridizations. We scanned the chips using the GenePix 4000A scanner (Axon Instruments) and analyzed primary data using Genepix 3.0 software (Axon Instruments). We carried out statistical analysis using a modified  $t$ -test implemented in SAM software<sup>30</sup>. We defined differentially expressed genes at a 1% false discovery rate confidence level and a cutoff for differential expression equal to three for the analysis of tissue specificity and two for *Spo11* experiments. We identified early and late spermatogenic genes from a mouse spermatogenesis data set<sup>23</sup> as those genes differentially expressed between groups of samples from 1-, 4-, 8-, 11- and 18-, 21-, 26-, 29-d-old mice ( $t$ -test,  $P < 0.01$ ; 3,102 genes are differentially expressed at this confidence level).

**GEO accession number.** Microarray data are available under the accession number GSE1138.

Note: Supplementary information is available on the Nature Genetics website.

## ACKNOWLEDGMENTS

We thank P. Hsieh and B. Oliver for critically reading the manuscript, the members of the Camerini-Otero lab for discussions, L. Robinson and L. Moore for their help and M. Primig for communicating results before publication.

## COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 12 December 2003; accepted 25 March 2004

Published online at <http://www.nature.com/naturegenetics/>

- Charlesworth, B. The evolution of chromosomal sex determination. *Novartis Found. Symp.* **244**, 207–219 (2002).
- Vallender, E.J. & Lahn, B.T. How mammalian sex chromosomes acquired their peculiar gene content. *Bioessays* **26**, 159–169 (2004).
- Rice, W.R. Sex chromosomes and the evolution of sexual dimorphism. *Evolution* **38**, 735–742 (1984).
- Fisher, R.A. The evolution of dominance. *Biol. Rev.* **6**, 345–368 (1931).
- Rice, W.R. Sexually antagonistic genes: experimental evidence. *Science* **256**, 1436–1439 (1992).
- Wang, P.J., Mccarrey, J.R., Yang, F. & Page, D.C. An abundance of X-linked genes expressed in spermatogonia. *Nat. Genet.* **27**, 422–426 (2001).
- Saifi, G.M. & Chandra, H.S. An apparent excess of sex- and reproduction-related genes on the human X chromosome. *Proc. R. Soc. Lond. B Biol. Sci.* **266**, 203–209 (1999).
- Lercher, M.J., Urrutia, A.O. & Hurst, L.D. Evidence that the human X chromosome is enriched for male-specific but not female-specific genes. *Mol. Biol. Evol.* **20**, 1113–1116 (2003).
- Parisi, M. *et al.* Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. *Science* **299**, 697–700 (2003).
- Reinke, V. *et al.* A global profile of germline gene expression in *C. elegans*. *Mol. Cell* **6**, 605–616 (2000).
- Ranz, J.M., Castillo-Davis, C.I., Meiklejohn, C.D. & Hartl, D.L. Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. *Science* **300**, 1742–1745 (2003).
- Romanienko, P.J. & Camerini-Otero, R.D. The mouse *Spo11* gene is required for meiotic chromosome synapsis. *Mol. Cell* **6**, 975–987 (2000).
- Humiecki, L., Lloyd, A.T. & Wolfe, K.H. Congruence of tissue expression profiles from Gene Expression Atlas, SAGEmap and TissueInfo databases. *BMC Genomics* **4**, 31 (2003).
- Su, A.I. *et al.* Large-scale analysis of the human and mouse transcriptomes. *Proc. Natl. Acad. Sci. USA* **99**, 4465–4470 (2002).
- Bono, H. *et al.* Systematic expression profiling of the mouse transcriptome using RIKEN cDNA microarrays. *Genome Res.* **13**, 1318–1323 (2003).
- Fraser, L.R. & Dudley, K. New insights into the t-complex and control of sperm function. *Bioessays* **21**, 304–312 (1999).
- Eddy, E.M. Male germ cell gene expression. *Recent Prog. Horm. Res.* **57**, 103–128 (2002).
- Bellve, A.R. Purification, culture, and fractionation of spermatogenic cells. *Methods Enzymol.* **225**, 84–113 (1993).
- Solari, A.J. The behavior of the XY pair in mammals. *Int. Rev. Cytol.* **38**, 273–317 (1974).
- Mckee, B.D. & Handel, M.A. Sex chromosomes, recombination, and chromatin conformation. *Chromosoma* **102**, 71–80 (1993).
- Lifschytz, E. & Lindsley, D.L. The role of X-chromosome inactivation during spermatogenesis. *Proc. Natl. Acad. Sci. USA* **69**, 182–186 (1972).
- Baudat, F., Manova, K., Yuen, J.P., Jasin, M. & Keeney, S. Chromosome synapsis defects and sexually dimorphic meiotic progression in mice lacking *Spo11*. *Mol. Cell* **6**, 989–998 (2000).
- Schultz, N., Hamra, F.K. & Garbers, D.L. A multitude of genes expressed solely in meiotic or postmeiotic spermatogenic cells offers a myriad of contraceptive targets. *Proc. Natl. Acad. Sci. USA* **100**, 12201–12206 (2003).
- Schlecht, U. *et al.* Expression profiling of Mammalian male meiosis and gametogenesis identifies novel candidate genes for roles in the regulation of fertility. *Mol. Biol. Cell* **15**, 1031–1043 (2004).
- Goldstein, P. The synaptonemal complexes of *Caenorhabditis elegans*: pachytene karyotype analysis of male and hermaphrodite wild-type and *him* mutants. *Chromosoma* **86**, 577–593 (1982).
- Kelly, W.G. *et al.* X-chromosome silencing in the germline of *C. elegans*. *Development* **129**, 479–492 (2002).
- Huynh, K.D. & Lee, J.T. Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos. *Nature* **426**, 857–862 (2003).
- Stekel, D.J., Git, Y. & Falciani, F. The comparison of gene expression from multiple cDNA libraries. *Genome Res.* **10**, 2055–2061 (2000).
- Tanaka, T.S. *et al.* Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray. *Proc. Natl. Acad. Sci. USA* **97**, 9127–9132 (2000).
- Tusher, V.G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA* **98**, 5116–5121 (2001).