Disentangling the relationship between sex-biased gene expression and X-linkage

Richard P. Meisel,[1,4,5] John H. Malone,[2,3,4] and Andrew G. Clark[1]

[1]*Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York 14853, USA;* [2]*Laboratory of Cellular and Developmental Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland 20892, USA*

X chromosomes are preferentially transmitted through females, which may favor the accumulation of X-linked alleles/genes with female-beneficial effects. Numerous studies have shown that genes with sex-biased expression are under- or over-represented on the X chromosomes of a wide variety of organisms. The patterns, however, vary between different animal species, and the causes of these differences are unresolved. Additionally, genes with sex-biased expression tend to be narrowly expressed in a limited number of tissues, and narrowly expressed genes are also non-randomly X-linked in a taxon-specific manner. It is therefore unclear whether the unique gene content of the X chromosome is the result of selection on genes with sex-biased expression, narrowly expressed genes, or some combination of the two. To address this problem, we measured sex-biased expression in multiple *Drosophila* species and at different developmental time points. These data were combined with available expression measurements from *Drosophila melanogaster* and mouse to reconcile the inconsistencies in X-chromosome content among taxa. Our results suggest that most of the differences between *Drosophila* and mammals are confounded by disparate data collection/analysis approaches as well as the correlation between sex bias and expression breadth. Both the *Drosophila* and mouse X chromosomes harbor an excess of genes with female-biased expression after controlling for the confounding factors, suggesting that the asymmetrical transmission of the X chromosome favors the accumulation of female-beneficial mutations in X-linked genes. However, some taxon-specific patterns remain, and we provide evidence that these are in part a consequence of constraints imposed by the dosage compensation mechanism in *Drosophila*.

[Supplemental material is available for this article.]

Sexual dimorphism is widespread among animal species. Because males and females share nearly identical genomes, differences in phenotypes are in part the result of the biased expression of certain genes in one sex or the other (Ellegren and Parsch 2007); these genes are said to have "sex-biased expression." Selection pressures can differ between males and females, and these sex-specific or sexually antagonistic selection pressures are expected to act on genes with sex-biased expression (Vicoso and Charlesworth 2006; Innocenti and Morrow 2010). X chromosomes are found in females two-thirds of the time, whereas autosomes are equally transmitted via the two sexes. The net selection pressure acting on X-linked genetic variation may therefore be biased in favor of female-beneficial alleles (Rice 1984; Connallon and Clark 2010; Fry 2010). Assuming that genes with sex-biased expression perform sex-specific-beneficial functions (Innocenti and Morrow 2010), the X-linkage of genes with sex-biased expression can thus be a valuable readout of historical sex-specific selection pressures (Vicoso and Charlesworth 2006).

The degree to which the sexes are dimorphic also varies across tissues. For example, some somatic tissues are highly homogeneous between the sexes, other somatic tissues are unique to one sex (e.g., male accessory glands in insects, the prostate in mammals), and the gonads are extremely differentiated between males and females. Each tissue expresses a subset of the genes in a genome—the more sexually dimorphic the tissue, the more differentiated is the pattern of female versus male gene expression in that tissue (Jin et al. 2001; Arbeitman et al. 2002, 2004; Parisi et al. 2003, 2004). In addition, the expression breadth of individual genes can range from broad (i.e., the gene is expressed in multiple tissues/developmental stages/environments) to narrow (expressed in very few tissues/stages/environments). Sex-specific selection pressures may differ if a gene is narrowly expressed in a sex-limited reproductive tissue or a tissue shared by both sexes (Meisel 2011). Furthermore, sexually antagonistic selection pressures are expected to be greatest in sexually mature adults and minimal in sexually immature juveniles (Chippindale et al. 2001; Rice and Chippindale 2001; Gibson et al. 2002). The fitness of a genotype is a composite of the performance of that genotype integrated across all tissues, developmental stages, and environments in which the gene(s) is/are expressed. Therefore, if one is to interpret the selection pressures acting on genes based on their X-linkage, one must consider both when and where each gene is expressed in the organism.

Multiple experiments have revealed that genes with male- or testis-biased expression are under-represented on the *Drosophila* X chromosome, while there is an excess of X-linked genes with female- and ovary-biased expression (e.g., Parisi et al. 2003; Ranz et al. 2003; Sturgill et al. 2007; Mikhaylova and Nurminsky 2011). Additionally, the right arm (XR or neo-X) of the *Drosophila pseudoobscura* X chromosome (which is autosomal in most other species) also contains a significant deficiency of genes with male-biased expression (Sturgill et al. 2007), suggesting that the paucity

[3]Present address: Department of Biological Science, Florida State University, Tallahassee, FL 32306, USA.
[4]These authors contributed equally to this work.
[5]Corresponding author
E-mail meisel@cornell.edu
Article published online before print. Article, supplemental material, and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.132100.111.

of X-linked genes with male-biased expression is an emergent property of the *Drosophila* X chromosome and not intrinsic to the proto-X.

Three hypotheses have been put forth to explain the deficiency of genes with male-biased expression on the *Drosophila* X chromosome. First, sexually antagonistic selection may prevent the accumulation of X-linked male-beneficial mutations because the X chromosome is preferentially transmitted through females (Rice 1984; Connallon and Clark 2010). Second, male meiotic sex-chromosome inactivation (MSCI) may prevent the expression of X-linked genes during spermatogenesis (Betrán et al. 2002; Hense et al. 2007; Vibranovski et al. 2009), which could decrease the proportion of X-linked genes with testis-biased expression. Third, dosage compensation, via hyperexpression, of the hemizygous X chromosome in males may limit the degree to which X-linked genes can acquire male-biased expression (Vicoso and Charlesworth 2009; Bachtrog et al. 2010). It is still unclear, however, the extent to which sexual antagonism, MSCI, and "dosage limits" each contributes to the deficiency of X-linked genes with male-biased expression.

In mammals, the X chromosome also harbors a unique collection of gene types, but in a way that is thought to differ from the patterns in *Drosophila* (Vicoso and Charlesworth 2006; Gurbich and Bachtrog 2008). For example, the *Drosophila* accessory gland and the mammalian prostate are both male-limited reproductive organs that contribute seminal proteins to the ejaculate. The *Drosophila melanogaster* X chromosome is depauperate in genes encoding accessory gland proteins (ACPs) (Wolfner et al. 1997; Swanson et al. 2001; Mueller et al. 2005) and genes with accessory-gland-biased expression (Mikhaylova and Nurminsky 2011), while there is an excess of X-linked human genes with prostate-biased expression (Lercher et al. 2003).

In addition, unlike in *Drosophila*, the mammalian X chromosome has an excess of genes with testis-biased expression (Wang et al. 2001; Lercher et al. 2003; Mueller et al. 2008). A closer examination, however, has revealed that this excess is limited to genes with pre-meiotic expression, whereas post-meiotically expressed genes are under-represented on the X chromosome because of MSCI (Khil et al. 2004). Similarly, the paucity of X-linked testis-expressed *D. melanogaster* genes may be limited to genes expressed during later stages of spermatogenesis such as meiosis (Vibranovski et al. 2009; Gan et al. 2010). Furthermore, new genes with testis- or male-biased expression tend to be preferentially located on the X or Z chromosome in *Drosophila* (Zhang et al. 2010a), mammals (Zhang et al. 2010b), and chicken (Ellegren 2011). Therefore, the X- (or Z-) linkage of testis-expressed genes may not differ substantially between taxa.

Analyses of non-reproductive tissues have also revealed similarities in the X-linkage of genes with sex-biased expression in *Drosophila* and mammals. For example, genes with male-biased expression in mouse brain are over-represented on the X chromosome (Yang et al. 2006), as are genes with male-biased expression in the *D. melanogaster* head (Chang et al. 2011). Also, genes with female-biased expression in brain, liver, and muscle are over-represented on the mouse X chromosome (Yang et al. 2006), and there is evidence for an excess of X-linked *D. melanogaster* genes with female- or ovary-biased expression (Parisi et al. 2003; Ranz et al. 2003; Mikhaylova and Nurminsky 2011). These results suggest that, while there are some differences in gene content between the *Drosophila* and mammalian X chromosomes, they may be more similar than previously thought (Vicoso and Charlesworth 2006; Gurbich and Bachtrog 2008). The remaining differences between the X chromosomes could be the result of differences in dosage compensation (Vicoso and Charlesworth 2006), different dominance coefficients of sexually antagonistic genetic variants (Rice 1984), or artifacts of how expression data are collected and analyzed in each taxon. To distinguish between these hypotheses, we must analyze comparable data from *Drosophila* and mouse using the same approaches.

In addition to the X chromosome harboring a non-random proportion of genes with sex-biased expression, the expression breadth of X-linked genes also differs from that of autosomal genes: The *D. melanogaster* X chromosome contains a deficiency of narrowly expressed genes (Mikhaylova and Nurminsky 2011), while there is an excess of X-linked narrowly expressed human genes (Lercher et al. 2003). Genes with sex-biased expression in non-reproductive tissues tend to be more narrowly expressed in both *D. melanogaster* and mouse (Mank et al. 2008; Meisel 2011). Additionally, *D. melanogaster* genes with female-biased expression are often highly expressed in ovary, and ovary-expressed genes tend to be broadly expressed (Parisi et al. 2004; Meisel 2011). It is therefore not clear whether the selection pressures that drive differences in gene content between the X chromosome and the autosomes are acting on genes with sex-biased expression, narrowly expressed genes, or some combination of the two (Mikhaylova and Nurminsky 2011).

The X-linkage of genes with sex-biased expression is an informative metric of sex-specific selection pressures (Vicoso and Charlesworth 2006), but to properly interpret these patterns and compare across taxa, one must control for when and where genes are expressed. To address the limitations of previous studies, we measured sex-biased expression in the heads of four *Drosophila* species (*D. melanogaster*, *D. pseudoobscura*, *Drosophila willistoni*, and *Drosophila mojavensis*) using a combination of microarrays and high-throughput RNA sequencing (RNA-seq), in whole flies from three species (*D. melanogaster*, *D. pseudoobscura*, and *D. mojavensis*) using RNA-seq, in thorax and abdomen of *D. willistoni* using RNA-seq, and in larvae of *D. melanogaster* using microarrays. *D. pseudoobscura* and *D. willistoni* are particularly informative because they have independently derived X-autosome fusions (creating neo-X chromosome arms) that involved the same ancestral autosome (see Meisel et al. 2009). We combined these expression measurements with available data on sex-biased and tissue-specific expression from both *D. melanogaster* and *Mus musculus* to more comprehensively examine how and why gene content differs between the X chromosome and autosomes in *Drosophila* and mammals.

Results

X-linkage and sex-biased expression in *Drosophila*

To confirm the results of microarray analyses (Parisi et al. 2003; Ranz et al. 2003; Sturgill et al. 2007), we measured sex-biased expression in whole flies using RNA-seq. As in the microarray experiments, our RNA-seq data show that the *Drosophila* X chromosome and *D. pseudoobscura* neo-X contain a dearth of genes with male-biased expression. We also find that genes with female-biased expression are over-represented on the X and neo-X chromosomes, and there is an excess of genes with male-biased expression on the arm homologous to *D. melanogaster* chromosome 2L (Muller element B) in all species (Fig. 1A), confirming patterns observed previously only in *D. melanogaster* (Parisi et al. 2003; Ranz et al. 2003). Most sex-biased expression is the result of genes

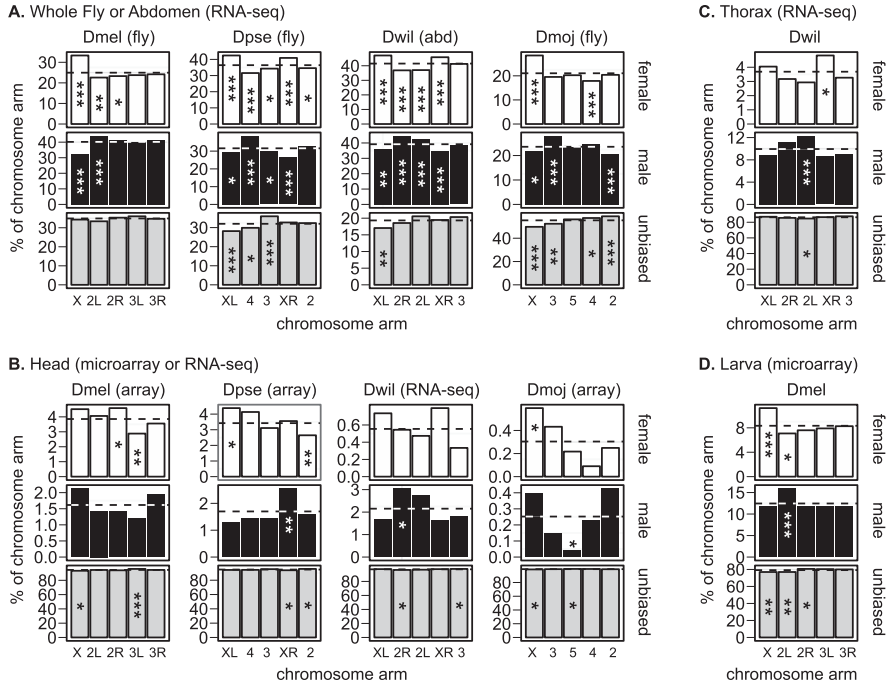

Figure 1. Frequency of genes with sex-biased expression on each chromosomal arm of *Drosophila* genomes. The percentages of genes on each chromosome arm that have female-biased (white), male-biased (black), or unbiased (gray) expression are indicated by the height of the bars. (Dashed lines) Genome-wide percent of genes in each sex-bias class. (*) Significant deviations between observed and expected frequencies (as determined by a permutation test); (*) $P < 0.05$, (**) $P < 0.01$, (***) $P < 0.001$. Chromosome arms are sorted such that homologous arms are in the same order for each species with the dot chromosome (Muller element F) omitted: (Dmel) *D. melanogaster*; (Dpse) *D. pseudoobscura*; (Dwil) *D. willistoni*; (Dmoj) *D. mojavensis*. Expression was measured with RNA-seq using RNA from whole flies (Dmel, Dpse, and Dmoj) or abdomen (Dwil) (*A*), microarrays or RNA-seq from head (*B*), RNA-seq from thorax (*C*), and microarrays using RNA from third instar larva (*D*). In the case of the RNA-seq data, reads were aligned to the genome using BWA, and differential expression was determined using edgeR (see Methods). Species in panels *A* and *B* are sorted by increasing evolutionary distance from *D. melanogaster*.

expressed in reproductive tissues (Parisi et al. 2003, 2004), so we measured expression in abdomen from male and female *D. willistoni* using RNA-seq to determine if the ancestral and neo-X chromosome arms have an excess/deficiency of genes with sex-biased expression. Genes with male-biased expression in *D. willistoni* abdomen are, indeed, under-represented, and genes with female-biased expression are over-represented on both arms of the X chromosome (Fig. 1A).

To test whether genes with sex-biased expression in non-reproductive tissues are over- or under-represented on the X chromosome, we measured expression in heads of males and females of four *Drosophila* species using microarrays and RNA-seq. A previous analysis of RNA-seq data found an excess of X-linked genes with male-biased expression in *D. melanogaster* head (Chang et al. 2011), and we confirmed this pattern with our RNA-seq data (Supplemental Fig. 1). While we detect a similar trend in our microarray data, the excess is not significant (Fig. 1B). However, the X-linkage of genes with male-biased expression in non-reproductive tissues in other species is not consistent with the pattern in *D. melanogaster*. For example, genes with male-biased expression in head are not significantly over-represented on the *D. mojavensis* X chromosome when we analyze our microarray data (Fig. 1B), and they are under-represented on the X chromosome in our RNA-seq data (Supplemental Fig. 1). The ancestral X chromosome in *D. pseudoobscura* is similarly depauperate in genes with male-biased expression in head (Fig. 1B; Supplemental Fig. 1). Interestingly, we detect an excess of genes with male-biased expression on the *D. pseudoobscura* neo-X based on our microarray data (Fig. 1B), but the RNA-seq data show the opposite pattern (Supplemental Fig. 1). In addition, genes with male-biased expression in head are neither significantly over- nor under-represented on the *D. willistoni* X or neo-X (Fig. 1B; Supplemental Fig. 1), and genes with male-biased expression in thorax may be under-represented on the *D. willistoni* X chromosome (Fig. 1C; Supplemental Fig. 2). Therefore, unlike the pattern observed when reproductive tissues are included, there is not a consistent under-representation of X-linked genes with male-biased expression in non-reproductive tissues. On the other hand, genes with female-biased expression in head or thorax tend to be over-represented on the X chromosome (Fig. 1B,C; Supplemental Figs. 1, 2), consistent with the pattern we observe when reproductive tissues are included (Fig. 1A).

We next tested whether ontogenic changes in sex-specific selection pressures (Chippindale et al. 2001; Rice and Chippindale 2001; Gibson et al. 2002) affect the X-linkage of genes with sex-biased expression by measuring gene expression in female and male wandering third instar larvae of *D. melanogaster* using microarrays. We failed to detect a significant deficiency of X-linked genes with male-biased expression in larvae (Fig. 1D), suggesting that the factors responsible for the paucity of X-linked genes with male-biased expression do not on average act on larval-expressed genes. However, genes with male-biased expression in both larvae and adults may be under-represented on the X chromosome (Supplemental Fig. 3), indicating that the dearth of X-linked genes with male-biased expression may not be solely the result of genes specifically expressed in adults. We also find that genes with male-biased expression in larvae are over-represented on chromosome arm 2L, and genes with female-biased expression are over-represented on the X chromosome, consistent with the patterns observed in adults (Fig. 1). These trends are also observed when we consider genes that have the same sex-biased expression in both larvae and adults (Supplemental Fig. 3).

On the X-linkage of *Drosophila* genes with testis-biased expression

We confirmed that *D. melanogaster* genes with enriched expression in testis relative to ovary are under-represented on the X chromosome (Supplemental Fig. 4; Sturgill et al. 2007; Vibranovski et al. 2009). Genes with testis-biased expression in mammals are identified by comparing across multiple tissues (e.g., Lercher et al. 2003). To parallel that analysis, we tested for a deficiency of X-linked *D. melanogaster* genes with testis-biased expression using

microarray measurements of expression from 14 adult tissues (Chintapalli et al. 2007).

First, we applied an approach in which the minimal ratio of testis expression level (microarray signal intensity) to the expression level in other tissues is used as a measure of the degree of testis-biased expression (Mikhaylova and Nurminsky 2011), and we failed to detect a significant deficiency of X-linked genes with testis-biased expression (Supplemental Fig. 5). Next, we calculated τ (Yanai et al. 2005), a metric of expression breadth that ranges from 0 (for broadly expressed genes) to 1 (for narrowly expressed genes). We applied τ cutoffs to call genes as narrowly expressed in a single tissue or broadly expressed in multiple tissues, and we did not observe a significant deficiency of X-linked genes with testis-biased expression (Fig. 2), even at the highest τ cutoffs (Supplemental Fig. 6). This result is not because of low power to detect significant differences—genes with testis-biased expression are the largest class of narrowly expressed genes, which means that we have the greatest power to detect significant differences from the expectation. In addition, X-linked genes with testis-biased expression have statistically indistinguishable τ values when compared with autosomal genes with testis-biased expression (Fig. 3A; Supplemental Fig. 7), demonstrating that there is not a difference in the degree of testis-biased expression between X-linked and autosomal genes.

Interestingly, genes with testis-biased expression that encode components of the sperm proteome (Dorus et al. 2006; Wasbrough et al. 2010) are under-represented on the X chromosome (Fig. 2; Supplemental Fig. 8). Therefore, while genes with enriched expression in testis relative to ovary are under-represented on the X chromosome (Sturgill et al. 2007; Vibranovski et al. 2009), the only genes with testis-biased expression that are deficient on the X chromosome when additional tissues are considered are those that encode proteins incorporated in the ejaculate.

Dosage compensation, sex-specific selection, and the X-linkage of *Drosophila* genes with accessory-gland-biased expression

We detect a significant deficiency of genes with accessory-gland-biased expression on the *D. melanogaster* X chromosome (Fig. 2; Supplemental Figs. 5, 6), consistent with the paucity of X-linked ACP genes (Wolfner et al. 1997; Swanson et al. 2001; Mueller et al. 2005). X-linked genes with accessory-gland-biased expression also tend to have broader expression profiles than autosomal genes with accessory-gland-biased expression (Fig. 3B; Supplemental Fig. 7). Therefore, genes with accessory-gland-biased expression are under-represented on the X chromosome, and those that are X-linked are not as narrowly expressed in accessory gland as are autosomal genes. This suggests that the exclusion of accessory-gland-expressed genes from the X chromosome may be responsible for the dearth of X-linked genes with male-biased expression. However, if we exclude genes with accessory-gland-biased expression, the X chromosome is still depauperate in genes with male-biased expression (Supplemental Fig. 9), demonstrating that the deficiency of X-linked accessory-gland-biased genes is not solely responsible.

X-linked genes with accessory-gland-biased expression have lower expression levels in accessory gland than similarly narrowly expressed autosomal genes (Fig. 4A; Supplemental Fig. 10), suggesting that limits imposed by dosage compensation may be responsible for the dearth of genes with accessory-gland-biased expression on the X chromosome (Vicoso and Charlesworth 2009; Bachtrog et al. 2010). In comparison, the accessory-gland-expression levels of broadly expressed genes do not differ significantly between the X chromosome and the autosomes (Fig. 4A). If dosage limits in accessory gland were responsible for the paucity of X-linked genes with male-biased expression, we would expect to see the greatest deficiency of X-linked genes with male-biased expression among those genes that are highly expressed in accessory gland. Using sex-biased expression calls from our RNA-seq data, we found that genes with male-biased expression that are expressed in accessory gland at low, moderate, and high levels are all under-represented on the X chromosome (Fig. 4B). This suggests that dosage limits in accessory gland are not solely responsible for the deficiency of X-linked genes with male-biased expression. However, genes with male-biased expression that are

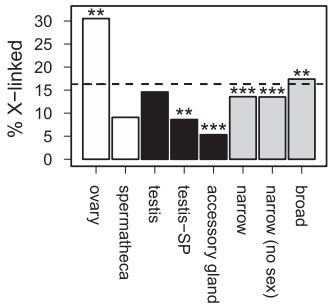

Figure 2. X-linkage of *D. melanogaster* broadly and narrowly expressed genes. (Barplots) Percent of *D. melanogaster* genes with narrow expression ($\tau > 0.7$) in each of four sex-limited tissues, with testis-biased expression and detectable in the sperm proteome (testis-SP), narrowly expressed in any of 14 tissues (narrow), narrowly expressed in one of 10 non-sex-limited tissues (no sex) or broadly expressed genes ($\tau \leq 0.4$) that are X-linked. (Dashed lines) Percent of the entire genome that is X-linked. (*) Observed values that significantly differ from the expectation based on the size of the X chromosome as determined by a permutation test; (**) $P < 0.01$, (***) $P < 0.001$.

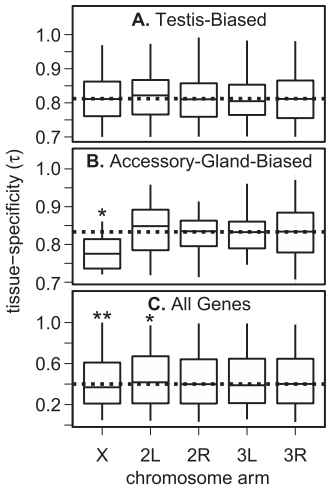

Figure 3. Distribution of τ for each *D. melanogaster* chromosome arm. (Boxes) Interquartile range; (horizontal line in the *middle* of the boxes) median value; (whiskers) 1.5× the interquartile range (outliers were omitted), (dashed line) genome-wide average. Significant differences between τ for a given chromosome and the rest of the genome were assessed using a Mann-Whitney test; (*) $P < 0.05$; (**) $P < 0.005$. Distributions of τ were calculated for (*A*) genes with testis-biased expression ($\tau > 0.7$), (*B*) genes with accessory-gland-biased expression ($\tau > 0.7$), and (*C*) all genes.

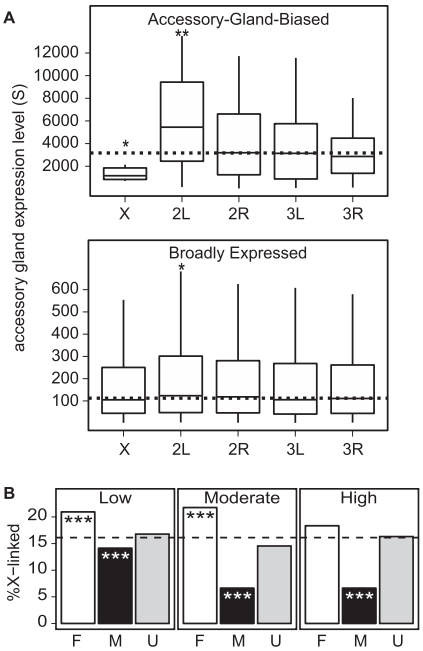

Figure 4. Accessory-gland expression level and X-linkage. (*A*) Boxplots show the distribution of the accessory gland expression level (*S*) for genes on each chromosome arm in *D. melanogaster* (chromosome 4 was omitted). (Boxes) Interquartile range; (horizontal line in the *middle* of the boxes) median value; (whiskers) 1.5× the interquartile range (outliers were omitted), (dashed line) genome-wide average. Genes were divided into those with accessory-gland-biased expression ($\tau > 0.7$) or those that are broadly expressed ($\tau \leq 0.4$). Significant differences in the expression level for a given chromosome and the rest of the genome were assessed using a Mann-Whitney test; (*) $P < 0.05$; (**) $P < 0.005$. (*B*) *D. melanogaster* genes were divided into those with low expression in accessory gland ($S < 100$), moderate expression in accessory gland ($100 \leq S < 1000$), and highly expressed in accessory gland ($S \geq 1000$). Within each panel is shown the percent of genes with female-biased (F), male-biased (M), and unbiased (U) expression that are X-linked. Genes were assigned to sex-biased expression classes using RNA-seq data (BWA alignments and edgeR to call differential expression). (*) Observed values that significantly differ from the expectation based on the size of the X chromosome as determined by a permutation test; (*) $P < 0.05$, (**) $P < 0.01$, (***) $P < 0.001$.

moderately or highly expressed in accessory gland are more underrepresented on the X chromosome than genes with male-biased expression that are lowly expressed in accessory gland ($P < 10^{-7}$, Fisher's exact test, FET). These results are robust to using different data to make the sex-biased expression calls (Supplemental Fig. 11). We therefore conclude that dosage limits in the accessory gland are partially responsible for the dearth of X-linked genes with male-biased expression.

Expression breadth, sex-biased expression, and X-linkage in *Drosophila*

We confirmed that the *D. melanogaster* X chromosome has a dearth of narrowly expressed genes and an excess of broadly expressed genes (Fig. 2; Supplemental Figs. 5, 6), and we also found that X-linked genes tend to have broader expression profiles than autosomal genes (Fig. 3C). Additionally, the deficiency of narrowly expressed genes on the X chromosome remains when genes that are narrowly expressed in sex-limited tissues are excluded (Fig. 2; Supplemental Figs. 5, 6), indicating that genes with accessory-gland-biased expression are not solely responsible for the deficiency of X-linked narrowly expressed genes.

If the paucity of X-linked genes with male-biased expression is a by-product of the correlation between expression breadth and sex bias (Meisel 2011; Mikhaylova and Nurminsky 2011), we do not expect to observe a deficiency of genes with male-biased expression on the X chromosome when we control for expression breadth. However, there is a significant deficiency of X-linked genes in *D. melanogaster* with male-biased expression among both broadly and narrowly expressed genes (Fig. 5A; Supplemental Figs. 12, 13), suggesting that the dearth of X-linked genes with male-biased expression is independent of expression breadth. The excess of X-linked genes with female-biased expression is also independent of expression breadth (Fig. 5A; Supplemental Figs. 12, 13), further demonstrating that the non-random X-linkage of genes with sex-biased expression is not a by-product of the correlation with expression breadth.

On the other hand, narrowly expressed genes with non-sex-biased (unbiased) expression are not significantly underrepresented on the X chromosome (Fig. 5; Supplemental Figs. 13, 14). Additionally, there is not a significant difference in the proportion of narrowly expressed genes with unbiased expression that are X-linked when compared with the proportion of more broadly expressed genes with unbiased expression ($P = 0.102$, FET) (Fig. 5B), and this result is robust to most τ cutoffs (Supplemental Table 1). We therefore conclude that the deficiency of X-linked narrowly expressed genes is an artifact of the correlation between expression breadth and sex bias, but the deficiency of X-linked genes with male-biased expression and excess with female-biased expression are independent of expression breadth.

Reconciling the differences in gene content between the *Drosophila* and mouse X chromosomes

To examine the apparent differences in gene content between the *Drosophila* and mammalian X chromosomes (Vicoso and Charlesworth 2006; Gurbich and Bachtrog 2008), we analyzed expression measurements from mouse in a similar manner as the *Drosophila* data described above. First, using microarray data (Yang et al. 2006), we confirmed that the mouse X chromosome has an excess of genes with female-biased expression in brain, liver, and muscle (Fig. 6A), much like the excess of X-linked genes with female-biased expression in *Drosophila* (Fig. 1). In addition, there is an excess of X-linked genes with male-biased expression in brain and a deficiency of X-linked genes with male-biased expression in muscle (Fig. 6A), demonstrating that in both *Drosophila* (Fig. 1) and

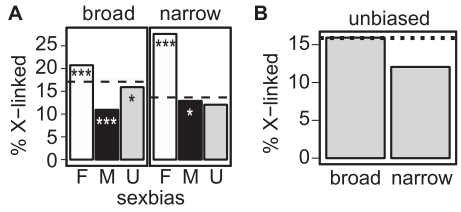

Figure 5. X-linkage, expression breadth, and sex-biased expression of *D. melanogaster* genes. (Barplots) Percent of *D. melanogaster* genes in various expression classes that are X-linked. (Dashed lines) Percent of all genes within that panel that are X-linked. (*) Observed values that significantly differ from the expectation based on the size of the X chromosome as determined by a permutation test; (*) $P < 0.05$, (***) $P < 0.001$. (*A*) Percent of broadly or narrowly expressed genes with either female-biased (F), male-biased (M), or unbiased (U) expression is plotted. (*B*) Percent of genes with unbiased expression that are either broadly ($\tau \leq 0.4$) or narrowly ($\tau > 0.7$) expressed is plotted. Genes were assigned to sex-biased expression classes using RNA-seq data (BWA alignments and edgeR to call differential expression).

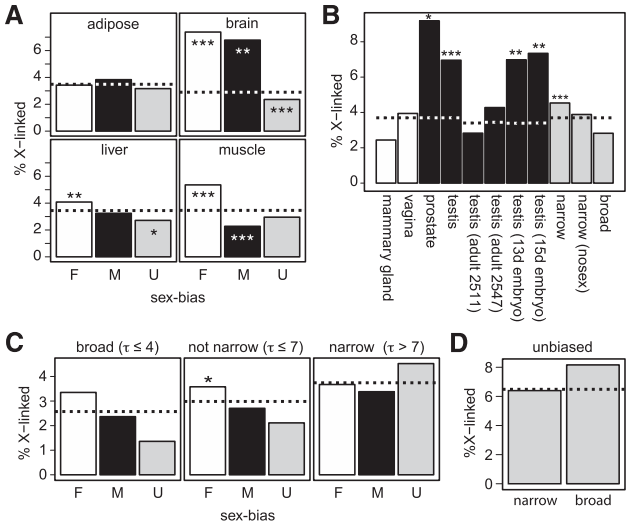

Figure 6. Sex-biased expression, expression breadth, and X-linkage in mouse. (Barplots) Percent of mouse genes in various expression classes that are X-linked. (Dashed lines) Percent of the entire genome that is X-linked. (*) Observed values that significantly differ from the expectation based on the size of the X chromosome as determined by a permutation test; (*) $P < 0.05$, (**) $P < 0.01$, (***) $P < 0.001$. (*A*) Percent of genes with female-biased (F), male-biased (M), and unbiased (U) expression that are X-linked is shown for data from four mouse tissues. (*B*) Percent of genes with narrow expression ($\tau > 0.7$) in each of four sex-limited tissues (white and black), with narrow expression in one of 25 tissues (gray), narrow expression in one of 21 non-sex-limited tissues (gray, no sex) or broad expression ($\tau \leq 0.4$; white) that are X-linked is plotted. Genes with testis-biased expression were identified using all testis EST libraries (testis), EST libraries from adult testis (UniGene IDs 2511 and 2547), or EST libraries from 13-d and 15-d embryonic testis (UniGene IDs 2555 and 2594). (*C*) Percent of X-linked broadly expressed ($\tau \leq 0.4$), not narrowly expressed ($\tau \leq 0.7$), or narrowly expressed ($\tau > 0.7$) genes with either female-biased (F), male-biased (M), or unbiased (U) expression that are X-linked is plotted. (*D*) Percent of X-linked genes with unbiased expression that are either broadly or narrowly expressed is plotted.

mammals, the X-linkage of genes with male-biased expression is tissue-dependent.

We used mapped expressed sequence tags (ESTs) from adult mouse testes and unfertilized mouse ovaries to identify genes with enriched expression in testis relative to ovary. As in *D. melanogaster* (Supplemental Fig. 4; Sturgill et al. 2007; Vibranovski et al. 2009), there is a deficiency of X-linked genes with testis-enriched expression (Supplemental Fig. 15). We next used mapped ESTs from 25 mouse tissues to calculate τ, and we found that there is an excess of X-linked genes with testis-biased expression (Fig. 6B; Supplemental Fig. 16). However, the testis samples include a mixture of cell types from different developmental stages, including embryos (which should contain only pre-meiotic cells unaffected by MSCI) and adults (with pre-meiotic, meiotic, and post-meiotic cells) (Khil et al. 2004). To adequately mirror our analysis of the *Drosophila* data (Fig. 2), we only included EST libraries from adult testis, and we failed to detect a significant excess or deficiency of X-linked genes with testis-biased expression (Fig. 6B). This is true for most EST libraries and τ cutoffs (Supplemental Figs. 16, 17). On the other hand, genes with biased expression in embryonic testis are over-represented on the X chromosome (Fig. 6B; Supplemental Fig. 16). Therefore, when similar analyses are performed on comparable tissue samples, genes with testis-biased expression in adults are found at their expected frequencies on both the *Drosophila* and mouse X chromosomes.

The mammalian prostate and *Drosophila* accessory gland are analogous male-specific organs that secrete seminal proteins necessary for male fertility. The *D. melanogaster* X chromosome has a paucity of genes with accessory-gland-biased expression (Fig. 2), and, similar to the pattern observed in humans (Lercher et al. 2003), there is an excess of X-linked mouse genes with prostate-biased expression (Fig. 6B; Supplemental Fig. 16). We hypothesize as to why the patterns differ between *Drosophila* and mammals in the Discussion.

While the *Drosophila* X chromosome appears to be deficient in narrowly expressed genes (Fig. 1; Mikhaylova and Nurminsky 2011), the mammalian X chromosome is enriched for narrowly expressed genes (Fig. 6B; Supplemental Fig. 16; Lercher et al. 2003; Deng et al. 2011). Sex bias and expression breadth are correlated in mouse (Mank et al. 2008; Meisel 2011), so we tested whether the non-random X-linkage of genes with sex-biased or narrow expression is an artifact of this correlation. When we exclude genes that are narrowly expressed in sex-limited tissues, there is no longer significant excess of X-linked narrowly expressed mouse genes (Fig. 6B; Supplemental Fig. 16). Additionally, narrowly expressed genes with non-sex-biased expression are not significantly over-represented on the mouse X chromosome (Fig. 6C,D; Supplemental Figs. 18, 19). Therefore, as in *Drosophila*, narrowly expressed mouse genes merely appear to be non-randomly X-linked because of the correlation between sex bias and expression breadth. Lastly, the mouse X chromosome has an excess of genes with female-biased expression that are not narrowly expressed (Fig. 6C; Supplemental Fig. 18), much like the excess of X-linked *Drosophila* genes with female-biased expression (Fig. 5A).

Discussion

Our analyses combined measurements of sex-biased expression and expression breadth to determine more precisely what types of genes are over- or under-represented on the *Drosophila* and mouse X chromosomes so that we can infer the historical effects of sex-specific selection pressures. We have attempted to consider all possible confounding factors and correlations exhaustively, and our results suggest the following conclusions (summarized in Table 1).

X-linkage and sex-biased expression in non-reproductive tissues

Consistent with previous results (e.g., Parisi et al. 2003; Yang et al. 2006; Chang et al. 2011), we showed that genes with sex-biased expression in non-reproductive tissues in both *Drosophila* and mammals are often non-randomly X-linked. Genes with female-biased expression are consistently over-represented on the X chromosome in both taxa (Figs. 1, 5A, 6), possibly because the female-biased transmission of the X chromosome favors the retention of alleles/genes with female-beneficial functions (Rice 1984; Connallon and Clark 2010). The X-linkage of genes with male-biased expression varies across tissues, changes throughout development, differs between taxa, and can depend on the methodology used to measure expression (Figs. 1, 5A, 6; Supplemental Fig. 1). The proportion of X-linked genes with male-biased expression in *Drosophila* head, for example, seems particularly labile (Fig. 1B; Supplemental Fig. 1). We do, however, find some support for the hypothesis that stronger net selection in females prevents the accumulation of X-linked genes with male-biased expression in non-reproductive tissues (Figs. 1, 5A; Supplemental Figs. 2, 9; Parisi et al. 2003).

Table 1. Factors responsible for non-random X-linkage

	Drosophila		Mouse	
Expression	**Pattern**	**Driven by**	**Pattern**	**Driven by**
Male-biased	Tissue-dependent	Sexual conflict in soma Acp genes on autosomes	Tissue-dependent	Sexual conflict in soma Prostate-biased on X
Female-biased	$X > A$	Sexual conflict in soma Ovary-biased on X	$X > A$	Sexual conflict in soma
Narrow	$X < A$	Correlation with sex bias	$X > A$	Correlation with sex bias
Testis	$X = A$	$X < A$ from comparing testis and ovary	$X = A$	$X > A$ if exclude MSCI
Accessory gland or prostate	$X < A$	Dosage limits	$X > A$	Male-specific selection

Interestingly, we detect an excess of X-linked genes with female-biased expression in larvae (Fig. 1C). If this excess is the result of female-beneficial substitutions being favored because of the X-chromosome's transmission bias, it would contradict the prediction that sexually antagonistic selection pressures are minimal in juveniles (Chippindale et al. 2001; Rice and Chippindale 2001; Gibson et al. 2002). However, third instar larvae have developing gonads (Hartenstein 1993), and this pattern could be driven by the excess of X-linked genes with ovary-biased expression (Fig. 2). In addition, while genes with male-biased expression in larvae are found on the X chromosome at the expected proportion (Fig. 1C), genes with male-biased expression in both larvae and adults may be under-represented on the X chromosome (Supplemental Fig. 3). The male-biased expression profile found in adults begins to manifest in the larval stages (Graveley et al. 2011), suggesting that selection on genes expressed in adults could affect the X-linkage of some larval-expressed genes. Additional data should be collected from larvae lacking gonadal tissues and from earlier stages of development to further interrogate how ontogenic changes in sexually antagonistic selection can affect the X-linkage of genes encoding proteins that perform sex-beneficial functions.

X-linkage and biased expression in male reproductive tissues

The *D. melanogaster* X chromosome has a deficiency of ACP genes (Wolfner et al. 1997; Swanson et al. 2001; Mueller et al. 2005) and genes with accessory-gland-biased expression (Fig. 2; Mikhaylova and Nurminsky 2011). We hypothesize that this is the result of dosage limits on X-linked genes in males, and these dosage limits in accessory gland appear to be partially responsibly for the paucity of genes with male-biased expression on the *Drosophila* X chromosome (Fig. 4B). Our expression data from larvae support this hypothesis. Male third instar larvae contain gonads with cells that have begun spermatogenesis (Fuller 1993), but they have yet to develop accessory glands (Hartenstein 1993). There is not a significant deficiency of X-linked genes with male-biased expression in larvae (Fig. 1D), suggesting that genes expressed in accessory gland make an important contribution to the dearth of X-linked genes with male-biased expression in adults. Genes with testis-biased expression that encode components of the sperm proteome are also under-represented on the *Drosophila* X chromosome (Fig. 2). Both ACPs and sperm proteins must be synthesized at high levels (Wolfner et al. 2005), which could exacerbate the dosage limits imposed by being X-linked.

The *Drosophila* accessory gland plays an important role in male fertility, secreting proteins that are incorporated into the seminal fluid, which mediate post-mating inter- and intrasexual conflict (Swanson and Vacquier 2002; Ravi Ram and Wolfner 2007). While it is tempting to invoke sexually antagonistic selection models (e.g., Rice 1984; Patten and Haig 2009; Connallon and Clark 2010) to explain the paucity of X-linked ACP genes, these models assume intralocus sexual antagonism. *Drosophila* male reproductive proteins, such as ACPs, are thought to mediate intersexual conflict by interacting with proteins expressed in the female reproductive tract (e.g., Yapici et al. 2008). Therefore, new models will be required to examine whether interlocus sexual conflict can explain the non-random X-linkage of male reproductive proteins.

Mammalian genes with prostate-biased expression, unlike *Drosophila* ACPs, are over-represented on the X chromosome (Fig. 6B; Lercher et al. 2003), as are mouse genes with testis-biased expression prior to meiosis (Fig. 6B; Khil et al. 2004). X-chromosome dosage in eutherian mammals is equilibrated by up-regulation of X-linked expression in both sexes followed by random inactivation of one X in each female cell (Payer and Lee 2008), which does not appear to impose dosage limits on highly expressed X-linked genes in males.

The excess of genes with prostate-biased and early spermatogenic expression on the mammalian X chromosome (Fig. 6B; Lercher et al. 2003; Khil et al. 2004) could be the result of sex-specific selection pressures. Unlike *Drosophila* ACPs, mammalian seminal proteins do not alter female physiology at a large scale (Pitnick et al. 2009), but they do mediate male–male sexual selection (Pizzari and Parker 2009). Therefore, the selection pressures on genes with prostate- and testis-biased expression likely act in a male-limited manner. If this is the case, the female-biased transmission of the X chromosome will not affect the X-linkage of these genes. We hypothesize that the hemizygous transmission of the X chromosome in males allows for the selective fixation of male-beneficial recessive mutations in X-linked mammalian genes (Charlesworth et al. 1987), promoting the accumulation of genes expressed in male reproductive tissues on the X chromosome.

MSCI and the X-linkage of genes with testis-biased expression

We detect a paucity of X-linked *D. melanogaster* genes with enriched expression in testis relative to ovary (Supplemental Fig. 4), but not when we examine genes with testis-biased expression compared with multiple tissues (Fig. 2). The previously noted deficiency of X-linked genes with testis-enriched expression (Sturgill et al. 2007; Vibranovski et al. 2009) is likely a methodological artifact for the following reasons. First, genes with ovary-biased expression are over-represented on the X chromosome (Fig. 2). Second, testis-expressed genes tend to be narrowly expressed (Meisel 2011), and narrowly expressed genes are under-represented on the X chromosome (Fig. 2). Furthermore, ovary-expressed genes are broadly expressed (Meisel 2011), and there is an excess of X-linked broadly expressed genes (Fig. 3C). While the relationship between expression breadth and X-linkage appears to be an artifact of the

correlation between expression breadth and sex-biased expression (Fig. 5), direct comparisons between testis and ovary expression do not control for this. Therefore, by comparing testis and ovary expression, one artificially obtains evidence for a paucity of X-linked testis-enriched genes.

The lack of a deficiency of X-linked genes with testis-biased expression does not, however, reveal insights into the debate over MSCI in *Drosophila* (Hense et al. 2007; Vibranovski et al. 2009; Meiklejohn et al. 2011; Mikhaylova and Nurminsky 2011). Mouse genes with testis-biased expression are also not under-represented on the X chromosome even when meiotic tissues are included (Fig. 6B; Khil et al. 2004), and MSCI is well documented in mouse (Turner 2007). On the other hand, our results as well as others—e.g., the X chromosome is deficient in genes with male-biased expression even when testis-expressed genes are excluded (Sturgill et al. 2007)—suggest that, even if MSCI occurs in *Drosophila*, it is not an important proximate cause of the paucity of X-linked genes with male-biased expression.

Conclusions

In summary, our results suggest that most of the observed differences in X-chromosome content between *Drosophila* and mammals are confounded by the correlation between sex bias and expression breadth, or they are affected by taxon-specific data collection and approaches to the analysis. The female-biased transmission of the X chromosome appears to favor the accumulation of female-beneficial alleles/genes on the X chromosome in both taxa, but the effects on male-beneficial alleles/genes are not consistent between taxa or across tissues. While models assuming intralocus sexual antagonism may explain many of these patterns, further work is required to develop theoretical models of the effects of interlocus conflict on the X-linkage of genes that are under sex-specific selection pressures.

Methods

Samples

Flies were grown under the following conditions: *D. melanogaster* (y^1 w^{67c} for whole fly, OreR for head and larvae) at 22°C on cornmeal media; *D. pseudoobscura* (*Drosophila* Species Stock Center [DSSC] 14011-0121.94) at 22°C on banana-opuntia media; *D. willistoni* (DSSC 14030-0811.25) at 18°C on cornmeal media; *D. mojavensis* (DSSC 15081-1352.22) at 22°C on banana-opuntia media. For the whole flies from *D. melanogaster*, *D. pseudoobscura*, and *D. mojavensis*, 5- to 7-d-old females and males were sampled. For the head samples, newly eclosed *D. melanogaster*, *D. pseudoobscura*, and *D. mojavensis* were aged in vials (males and females together) for 7 d. The sexes were separated on day 7 under CO_2 and stored in separate vials for 24 h. Flies were flash-frozen on dry ice after 24 h and stored at −80°C; heads were dissected from frozen flies using forceps on ceramic plates chilled on dry ice. Males and females of newly eclosed *D. willistoni* were separated under CO_2 and aged for 5–7 d; whole flies were placed live in Ringers solution, and heads, thoraxes, and abdomens were dissected; samples from each segment were flash-frozen in liquid N_2 and stored at −80°C. *D. melanogaster* wandering third instar larva were collected and sexed based on the size of the visible gonad (Demerec 1950). Males and females were separately flash-frozen in liquid N_2 and stored at −80°C. RNA was extracted from the samples using TRIzol (Invitrogen) following the manufacturer's instructions.

Microarrays

Microarray measurements of gene expression were performed on RNA samples from the heads of *D. melanogaster* using the stock Nimblegen 60-mer array (Dmel r. 4.2.1), heads of *D. pseudoobscura* and *D. mojavensis* using custom Nimblegen 50-mer arrays (Zhang et al. 2007), and larvae of *D. melanogaster* using Agilent 4-sector 60-mer arrays (G2519F). Four biological replicates were collected of the head microarray data, and three biological replicates were collected of the larval microarray data. Total RNA was reverse-transcribed to cDNA, samples were dye-labeled, and hybridization was performed per the manufacturers' instructions. For the head data, RMA normalization was performed using Nimblescan v 2.4 software (Roche Nimblegen Inc.). We tested the null hypothesis that a gene is expressed at the same level in females and males using a moderated *t*-test implemented in the LIMMA package of Bioconductor in the R statistical software environment (R Development Core Team 2009) with the empirical Bayes function to pool sample variances toward a common value (Smyth 2004). For the larval data, the LIMMA package was used to normalize channel intensities between arrays (Smyth and Speed 2003) and test for differential expression between samples (Smyth 2004). *P*-values from statistical tests were corrected for the false discovery rate (FDR) using the Benjamini and Hochberg (1995) adjustment. Additional information can be found in GEO accessions GSE23309 and GSE31722.

RNA-seq

Illumina libraries were prepared using RNA extracted from males and females from the following samples: *D. melanogaster* whole fly, *D. melanogaster* head, *D. pseudoobscura* whole fly, *D. pseudoobscura* head, *D. willistoni* head, *D. willistoni* thorax, *D. willistoni* abdomen, *D. mojavensis* whole fly, *D. mojavensis* head. Standard Illumina library sample preparation procedures were used: Poly-adenylated mRNA was isolated using either the QIAGEN Oligotex mRNA kit or oligo-dT Dynabeads (Invitrogen); mRNA was fragmented using alkaline hydrolysis; double-stranded cDNA was created using random hexamer primers; sequencing adaptors were ligated to the cDNA; fragments were size-selected by gel purification; and the size-selected fragments were amplified by PCR. Samples were run on the following instrumentation: Illumina Genome Analyzer I (one lane each of male and female *D. melanogaster*, *D. pseudoobscura*, and *D. mojavensis* head samples), Illumina Genome Analyzer II (one lane each of male and female *D. melanogaster*, *D. pseudoobscura*, and *D. mojavensis* whole fly samples; three lanes each of male and female *D. melanogaster* head samples; three lanes each of male and female *D. pseudoobscura* head samples; four lanes each of male and female *D. mojavensis* head samples; one lane each of male and female *D. melanogaster*, *D. pseudoobscura*, and *D. mojavensis* whole fly samples; one lane each of male and female *D. willistoni* head, thorax, and abdomen samples), Illumina HiSeq (one lane each of male and female *D. melanogaster*, *D. pseudoobscura*, and *D. mojavensis* whole fly samples; one lane each of male and female *D. willistoni* head samples). Additional details can be found in GEO accessions GSE20348, GSE20882, GSE19989, GSE28078, and GSE31723 (Graveley et al. 2011; Malone and Oliver 2011). A full list of the NCBI Sequence Read Archive runs can be found in Supplemental Table 2.

Reads were processed using the standard Illumina pipeline (see GEO accessions), and all reads longer than 36 bp were shortened to 36 bp to avoid biases introduced by different read lengths and to eliminate low-quality sequences at the 3′ end of longer reads. For paired-end runs, we only used the reads in the forward direction. The reads were aligned to the annotated protein-coding

sequences using BWA (Li and Durbin 2009) and Bowtie (Langmead et al. 2009). The following annotations were used: *D. melanogaster* r5.36, *D. pseudoobscura* r2.19, *D. willistoni* r1.3, and *D. mojavensis* r1.3. Only the longest annotated protein-coding sequence was used for each gene. Reads that mapped to multiple locations were not counted, and genes with fewer than 50 mapped reads (males and females combined) were excluded from subsequent analyses. We then tested for differential expression between male and female samples using DESeq (Anders and Huber 2010), edgeR (Robinson et al. 2010), and Cufflinks (Trapnell et al. 2010), using *P*-values corrected for an FDR of 0.05 (Benjamini and Hochberg 1995). Our results are robust to the alignment algorithm and statistical analysis used (Supplemental Fig. 1) unless otherwise mentioned above.

Additional sex-biased expression data sets

We obtained microarray measurements of sex-biased expression in whole *D. melanogaster* and head from SEBIDA (Gnad and Parsch 2006). We also obtained microarray measurements of sex-biased gene expression from four mouse tissue samples (Yang et al. 2006). For the mouse data, we used the results of a Mann-Whitney test to assign genes to sex-biased expression categories (Mank et al. 2008), and we applied an FDR cutoff of 0.05. This was done separately for each tissue. We called genes as sex-biased in mouse (Fig. 6C,D) if they were assigned to a sex-biased expression class in one tissue and were either put in the same sex-biased expression class or not called as sex-biased in all other tissues.

Expression breadth

Microarray signal intensities (S) from 14 adult *D. melanogaster* tissues (brain, eye, thoracicoabdominal ganglion, salivary gland, crop, midgut, tubule, hindgut, heart, fat body, ovary, testis, accessory gland, and spermatheca) were obtained from FlyAtlas (Chintapalli et al. 2007). Measurements from spermatheca of virgin and mated females were averaged to obtain a single spermatheca signal because they are well correlated (Meisel 2011). Genes were said to be moderately expressed in a tissue if $S \geq 100$ and highly expressed if $S \geq 1000$, per the curators' instructions (Chintapalli et al. 2007).

The number of ESTs mapping to each gene in the mouse genome was obtained for 25 tissues (adipose, bladder, blood, bone, bone marrow, brain, dorsal root ganglion, eye, vagina, intestine, heart, inner ear, liver, lung, lymph node, mammary gland, muscle, pancreas, prostate, skin, spinal cord, spleen, sympathetic ganglion, testis, and thymus) from UniGene build 190. Additionally, the following individual EST libraries were extracted from adult or embryonic testis: 254 (Yuan et al. 1995), 483, 1336 (Marra et al. 1999), 1784, 2511, 2555, 2594, 18007 (Okazaki et al. 2002), 2547 (Carninci et al. 2003), and 7009 (Strausberg et al. 2002).

Expression breadth was calculated two ways. In the first approach (applied only to the *D. melanogaster* data), we calculated the minimal ratio of the expression level in a particular tissue relative to all other tissues (Mikhaylova and Nurminsky 2011). In the second approach, we calculated a metric of expression breadth as follows:

$$\tau = \frac{\sum_{i=1}^{N} 1 - \frac{\log S_i}{\log S_{max}}}{N-1},$$

where N is the number of tissues analyzed, S_i is the expression level in tissue i (signal intensity in FlyAtlas and number of ESTs mapped to the gene standardized by the total number of mapped ESTs in that tissue for the UniGene data), and $S_{\max}$ is the maximum expression level across all tissues (Yanai et al. 2005; Larracuente et al. 2008; Mank et al. 2008; Meisel 2011). In the FlyAtlas data, all $S_i < 1$ were set to 1 so that $\log(S_i) \geq 0$. If multiple probes were present for a given gene, the median S across all probes was used for that gene. In the UniGene data, all genes with fewer than three mapped ESTs across all 25 tissues were removed, and $S_i < 2$ were set to 2 for the remaining genes. The τ value for each gene was used to determine if the gene is narrowly or broadly expressed using various cutoffs.

Sperm proteome

Genes were said to encode proteins in the sperm proteome if they were detected in at least one of two mass spectrometry analyses of *D. melanogaster* sperm proteins (Dorus et al. 2006; Wasbrough et al. 2010).

Statistical analysis

We permuted the chromosomal locations of the genes in our data sets 1000 times to estimate a null distribution of the number of genes in each expression class on each chromosome arm in *Drosophila* and each chromosome in mouse. Depending on the comparison being made, this was either done for all genes or a subset of genes under consideration. All permutations were performed in the R statistical package (R Development Core Team 2009).

Data access

All microarray and RNA-seq data presented in this manuscript have been submitted to the NCBI Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) under accession numbers GSE23309, GSE31722, and GSE31723.

Acknowledgments

This work was supported in part by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases. R.P.M. was supported by National Institutes of Health fellowship F32GM087611, and J.H.M. completed this work while in the laboratory of Brian Oliver. We would like to thank Caroline Sartain for assisting in the sexing of larvae, Amanda Manfredo for preparing RNA-seq libraries, and the Cornell University Genomics Core Laboratory for running RNA-seq and microarray samples. This manuscript benefited from discussions with Brian Oliver, Colin Meiklejohn, and members of the Clark laboratory, as well as comments from three anonymous reviewers.

References

Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biol* **11:** R106. doi: 10.1186/gb-2010-11-10-r106.

Arbeitman MN, Furlong EEM, Imam F, Johnson E, Null BH, Baker BS, Krasnow MA, Scott MP, Davis RW, White KP, et al. 2002. Gene expression during the life cycle of *Drosophila melanogaster*. *Science* **297:** 2270–2275.

Arbeitman MN, Fleming AA, Siegal ML, Null BH, Baker BS. 2004. A genomic analysis of *Drosophila* somatic sexual differentiation and its regulation. *Development* **131:** 2007–2021.

Bachtrog D, Toda NR, Lockton S. 2010. Dosage compensation and demasculinization of X chromosomes in *Drosophila*. *Curr Biol* **20:** 1476–1481.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol* **57:** 289–300.

Betrán E, Thornton K, Long M. 2002. Retroposed new genes out of the X in *Drosophila*. *Genome Res* **12:** 1854–1859.

Carninci P, Waki K, Shiraki T, Konno H, Shibata K, Itoh M, Aizawa K, Arakawa T, Ishii Y, Sasaki D, et al. 2003. Targeting a complex

transcriptome: The construction of the mouse full-length cDNA encyclopedia. *Genome Res* **13:** 1273–1289.
Chang PL, Dunham JP, Nuzhdin SV, Arbeitman MN. 2011. Somatic sex-specific transcriptome differences in *Drosophila* revealed by whole transcriptome sequencing. *BMC Genomics* **12:** 364. doi: 10.1186/1471-2164-12-364.
Charlesworth B, Coyne JA, Barton NH. 1987. The relative rates of evolution of sex chromosomes and autosomes. *Am Nat* **130:** 113–146.
Chintapalli VR, Wang J, Dow JAT. 2007. Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. *Nat Genet* **39:** 715–720.
Chippindale AK, Gibson JR, Rice WR. 2001. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. *Proc Natl Acad Sci* **98:** 1671–1675.
Connallon T, Clark AG. 2010. Sex linkage, sex-specific selection, and the role of recombination in the evolution of sexually dimorphic gene expression. *Evolution* **64:** 3417–3442.
Demerec M. 1950. *Biology of* Drosophila. Wiley, New York.
Deng X, Hiatt JB, Nguyen DK, Ercan S, Sturgill D, Hillier LW, Schlesinger F, Davis CA, Reinke VJ, Gingeras TR, et al. 2011. Evidence for compensatory upregulation of expressed X-linked genes in mammals, *Caenorhabditis elegans* and *Drosophila melanogaster*. *Nat Genet* **43:** 1179–1185.
Dorus S, Busby SA, Gerike U, Shabanowitz J, Hunt DF, Karr TL. 2006. Genomic and functional evolution of the *Drosophila melanogaster* sperm proteome. *Nat Genet* **38:** 1440–1445.
Ellegren H. 2011. Emergence of male-biased genes on the chicken Z-chromosome: Sex-chromosome contrasts between male and female heterogametic systems. *Genome Res* **21:** 2082–2086.
Ellegren H, Parsch J. 2007. The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev Genet* **8:** 689–698.
Fry JD. 2010. The genomic location of sexually antagonistic variation: Some cautionary comments. *Evolution* **64:** 1510–1516.
Fuller MT. 1993. Spermatogenesis. In *The development of* Drosophila melanogaster (ed. M Bate and A Martinez-Arias), Vol. I, pp. 71–147. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Gan Q, Chepelev I, Wei G, Tarayrah L, Cui K, Zhao K, Chen X. 2010. Dynamic regulation of alternative splicing and chromatin structure in *Drosophila* gonads revealed by RNA-seq. *Cell Res* **20:** 763–783.
Gibson JR, Chippindale AK, Rice WR. 2002. The X chromosome is a hot spot for sexually antagonistic fitness variation. *Proc R Soc Lond B Biol Sci* **269:** 499–505.
Gnad F, Parsch J. 2006. Sebida: A database for the functional and evolutionary analysis of genes with sex-biased expression. *Bioinformatics* **22:** 2577–2579.
Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, Artieri CG, van Baren MJ, Boley N, Booth BW, et al. 2011. The developmental transcriptome of *Drosophila melanogaster*. *Nature* **471:** 473–479.
Gurbich TA, Bachtrog D. 2008. Gene content evolution on the X chromosome. *Curr Opin Genet Dev* **18:** 493–498.
Hartenstein V. 1993. *Atlas of* Drosophila *development*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Hense W, Baines JF, Parsch J. 2007. X chromosome inactivation during *Drosophila* spermatogenesis. *PLoS Biol* **5:** e273. doi: 10.1371/journal.pbio.0050273.
Innocenti P, Morrow EH. 2010. The sexually antagonistic genes of *Drosophila melanogaster*. *PLoS Biol* **8:** e1000335. doi: 10.1371/journal.pbio.1000335.
Jin W, Riley RM, Wolfinger RD, White KP, Passador-Gurgel G, Gibson G. 2001. The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*. *Nat Genet* **29:** 389–395.
Khil PP, Smirnova NA, Romanienko PJ, Camerini-Otero RD. 2004. The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. *Nat Genet* **36:** 642–646.
Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10:** R25. doi: 10.1186/gb-2009-10-3-r25.
Larracuente AM, Sackton TB, Greenberg AJ, Wong A, Singh ND, Sturgill D, Zhang Y, Oliver B, Clark AG. 2008. Evolution of protein-coding genes in *Drosophila*. *Trends Genet* **24:** 114–123.
Lercher MJ, Urrutia AO, Hurst LD. 2003. Evidence that the human X chromosome is enriched for male-specific but not female-specific genes. *Mol Biol Evol* **20:** 1113–1116.
Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25:** 1754–1760.
Malone JH, Oliver B. 2011. Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biol* **9:** 34. doi: 10.1186/1741-7007-9-34.
Mank JE, Hultin-Rosenberg L, Zwahlen M, Ellegren H. 2008. Pleiotropic constraint hampers the resolution of sexual antagonism in vertebrate gene expression. *Am Nat* **171:** 35–43.
Marra M, Hillier L, Kucaba T, Allen M, Barstead R, Beck C, Blistain A, Bonaldo M, Bowers Y, Bowles L, et al. 1999. An encyclopedia of mouse genes. *Nat Genet* **21:** 191–194.
Meiklejohn CD, Landeen EL, Cook JM, Kingan SB, Presgraves DC. 2011. Sex chromosome-specific regulation in the *Drosophila* male germline but little evidence for chromosomal dosage compensation or meiotic inactivation. *PLoS Biol* **9:** e1001126. doi: 10.1371/journal.pbio.1001126.
Meisel RP. 2011. Towards a more nuanced understanding of the relationship between sex-biased gene expression and rates of protein coding sequence evolution. *Mol Biol Evol* **28:** 1893–1900.
Meisel RP, Han MV, Hahn MW. 2009. A complex suite of forces drives gene traffic from *Drosophila* X chromosomes. *Genome Biol Evol* **1:** 176–188.
Mikhaylova LM, Nurminsky DI. 2011. Lack of global meiotic sex chromosome inactivation, and paucity of tissue-specific gene expression on the *Drosophila* X chromosome. *BMC Biol* **9:** 29. doi: 10.1186/1741-7007-9-29.
Mueller JL, Ram KR, McGraw LA, Bloch Qazi MC, Siggia ED, Clark AG, Aquadro CF, Wolfner MF. 2005. Cross-species comparison of *Drosophila* male accessory gland protein genes. *Genetics* **171:** 131–143.
Mueller JL, Mahadevaiah SK, Park PJ, Warburton PE, Page DC, Turner JMA. 2008. The mouse X chromosome is enriched for multicopy testis genes showing postmeiotic expression. *Nat Genet* **40:** 794–799.
Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, Kondo S, Nikaido I, Osato N, Saito R, Suzuki H, et al. 2002. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* **420:** 563–573.
Parisi M, Nuttall R, Naiman D, Bouffard G, Malley J, Andrews J, Eastman S, Oliver B. 2003. Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. *Science* **299:** 697–700.
Parisi M, Nuttall R, Edwards P, Minor J, Naiman D, Lu J, Doctolero M, Vainer M, Chan C, Malley J, et al. 2004. A survey of ovary-, testis-, and soma-biased gene expression in *Drosophila melanogaster* adults. *Genome Biol* **5:** R40. doi: 10.1186/gb-2004-5-6-r40.
Patten MM, Haig D. 2009. Maintenance or loss of genetic variation under sexual and parental antagonism at a sex-linked locus. *Evolution* **63:** 2888–2895.
Payer B, Lee JT. 2008. X chromosome dosage compensation: How mammals keep the balance. *Annu Rev Genet* **42:** 733–772.
Pitnick S, Wolfner MF, Suarez S. 2009. Sperm–female interactions. In *Sperm biology: An evolutionary perspective* (ed. TR Birkhead et al.), pp. 247–304. Academic Press, London.
Pizzari T, Parker GA. 2009. Sperm competition and sperm phenotype. In *Sperm biology: An evolutionary perspective* (ed. TR Birkhead et al.), pp. 207–245. Academic Press, London.
R Development Core Team. 2009. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
Ranz JM, Castillo-Davis CI, Meiklejohn CD, Hartl DL. 2003. Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. *Science* **300:** 1742–1745.
Ravi Ram K, Wolfner MF. 2007. Seminal influences: *Drosophila* Acps and the molecular interplay between males and females during reproduction. *Integr Comp Biol* **47:** 427–445.
Rice WR. 1984. Sex chromosomes and the evolution of sexual dimorphism. *Evolution* **38:** 735–742.
Rice WR, Chippindale AK. 2001. Intersexual ontogenetic conflict. *J Evol Biol* **14:** 685–693.
Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: A Bioconductor package for differential expression analysis of digital expression data. *Bioinformatics* **26:** 139–140.
Smyth GK. 2004. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* **3:** 1544–6115.
Smyth GK, Speed TP. 2003. Normalization of cDNA microarray data. *Methods* **31:** 265–273.
Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, et al. 2002. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc Natl Acad Sci* **99:** 16899–16903.
Sturgill D, Zhang Y, Parisi M, Oliver B. 2007. Demasculinization of X chromosomes in the *Drosophila* genus. *Nature* **450:** 238–241.
Swanson WJ, Vacquier VD. 2002. The rapid evolution of reproductive proteins. *Nat Rev Genet* **3:** 137–144.
Swanson WJ, Clark AG, Waldrip-Dail HM, Wolfner MF, Aquadro CF. 2001. Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. *Proc Natl Acad Sci* **98:** 7375–7379.
Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. 2010. Transcript assembly and

quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28:** 511–515.
Turner JMA. 2007. Meiotic sex chromosome inactivation. *Development* **134:** 1823–1831.
Vibranovski MD, Lopes HF, Karr TL, Long M. 2009. Stage-specific expression profiling of *Drosophila* spermatogenesis suggests that meiotic sex chromosome inactivation drives genomic relocation of testis-expressed genes. *PLoS Genet* **5:** e1000731. doi: 10.1371/journal.pgen.1000731.
Vicoso B, Charlesworth B. 2006. Evolution on the X chromosome: Unusual patterns and processes. *Nat Rev Genet* **7:** 645–653.
Vicoso B, Charlesworth B. 2009. The deficit of male-biased genes on the *D. melanogaster* X chromosome is expression-dependent: A consequence of dosage compensation? *J Mol Evol* **68:** 576–583.
Wang PJ, McCarrey JR, Yang F, Page DC. 2001. An abundance of X-linked genes expressed in spermatogonia. *Nat Genet* **27:** 422–426.
Wasbrough ER, Dorus S, Svenja H, Howard-Murkin J, Lilley K, Wilkin E, Polpitiya A, Petritis K, Karr TL. 2010. The *Drosophila melanogaster* sperm proteome-II (DmSP-II). *J Proteomics* **73:** 2171–2185.
Wolfner MF, Harada HA, Bertram MJ, Stelick TJ, Kraus KW, Kalb JM, Lung YO, Neubaum DM, Park M, Tram U. 1997. New genes for male accessory gland proteins in *Drosophila melanogaster. Insect Biochem Mol Biol* **27:** 825–834.
Wolfner MF, Applebaum S, Heifetz Y. 2005. Insect gonadal glands and their gene products. In *Comprehensive insect physiology, biochemistry, pharmacology and molecular biology* (ed. L Gilbert et al.), pp. 179–212. Elsevier, New York.
Yanai I, Benjamin H, Shmoish M, Chalifa-Caspi V, Shklar M, Ophir R, Bar-Even A, Horn-Saban S, Safran M, Domany E, et al. 2005. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* **21:** 650–659.
Yang X, Schadt EE, Wang S, Wang H, Arnold AP, Ingram-Drake L, Drake TA, Lusis AJ. 2006. Tissue-specific expression and regulation of sexually dimorphic genes in mice. *Genome Res* **16:** 995–1004.
Yapici N, Kim Y-J, Ribeiro C, Dickson BJ. 2008. A receptor that mediates the post-mating switch in *Drosophila* reproductive behavior. *Nature* **451:** 33–37.
Yuan L, Liu JG, Höög C. 1995. Rapid cDNA sequencing in combination with RNA expression studies in mice identifies a large number of male germ cell-specific sequence tags. *Biol Reprod* **52:** 131–138.
Zhang Y, Sturgill D, Parisi M, Kumar S, Oliver B. 2007. Constraint and turnover in sex-biased gene expression in the genus *Drosophila*. *Nature* **450:** 233–237.
Zhang YE, Vibranovski MD, Krinsky BH, Long M. 2010a. Age-dependent chromosomal distribution of male-biased genes in *Drosophila*. *Genome Res* **20:** 1526–1533.
Zhang YE, Vibranovski MD, Landback P, Marais GA, Long M. 2010b. Chromosomal redistribution of male-biased genes in mammalian evolution with two bursts of gene gain on the X chromosome. *PLoS Biol* **8:** e1000494. doi: 10.1371/journal.pbio.1000494.

Received September 16, 2011; accepted in revised form April 6, 2012.

Disentangling the relationship between sex-biased gene expression and X-linkage

Richard P. Meisel, John H. Malone and Andrew G. Clark

Genome Res. 2012 22: 1255-1265 originally published online April 12, 2012
Access the most recent version at doi:10.1101/gr.132100.111