A long term demasculinization of X-linked intergenic noncoding RNAs in *Drosophila melanogaster*

Ge Gao, Maria D. Vibranovski, Li Zhang, et al.

Genome Res. published online January 9, 2014
Access the most recent version at doi:10.1101/gr.165837.113

A long term demasculinization of X-linked intergenic noncoding RNAs in *Drosophila melanogaster*

Ge Gao*[1], Maria D Vibranovski * [2,3], Li Zhang[1,2], Zheng Li[4], Ming Liu[4], Yong E Zhang[2,5], Xinmin Li[6♤], Wenxia Zhang[4], Qichang Fan[4], Nicholas W. VanKuren[2,7], Manyuan Long[2,#], Liping Wei[1,#]

1. *Center for Bioinformatics, State Key Laboratory of Protein and Plant Gene Research, College of Life Sciences, Peking University, Beijing, 100871, P.R. China;* 2. *Department of Ecology and Evolution, University of Chicago, Chicago, IL USA 60637; 3. Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo, Brazil 05508; 4. Center of Developmental Biology and Genetics, College of Life Sciences, Peking University, Beijing, P.R. China 100871; 5. Key Laboratory of the Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing P.R. China 100101; 6. Genome Facility, University of Chicago, Chicago, IL USA 60637; 7. Committee on Genetics, Genomics, and Systems Biology, University of Chicago, Chicago, IL USA 60637*

* These authors contributed equally to this work.

♤ Current address: Department of Pathology & Laboratory Medicine, University of California at Los Angeles, Los Angeles, CA 90095

To whom correspondence should be addressed. E-mail: weilp@mail.cbi.pku.edu.cn; mlong@uchicago.edu

Manuscript Type: Research Letter

Running Title: Demasculinization of X-linked incRNAs in *Drosophila*

Keywords: noncoding RNAs, transcriptome profiling, male-biased expression, demasculinization

ABSTRACT

Recent studies revealed key roles of non-coding RNAs in sex-related pathways, but little is known about the evolutionary forces acting on these non-coding RNAs. Profiling the transcriptome of *Drosophila melanogaster* with whole-genome tiling arrays found that 15% of male-biased transcribed fragments are intergenic non-coding RNAs (incRNAs), suggesting a potentially important role for incRNAs in sex-related biological processes. Statistical analysis revealed a paucity of male-biased incRNAs and coding genes on the X chromosome, suggesting that similar evolutionary forces could be affecting the genomic organization of both coding and non-coding genes. Expression profiling across germline and somatic tissues further suggested that both male meiotic sex chromosome inactivation (MSCI) and sexual antagonism could contribute to the chromosomal distribution of male-biased incRNAs. Comparative sequence analysis showed that the evolutionary age of male-biased incRNAs is a significant predictor of their chromosomal locations. In addition to identifying abundant sex-biased incRNAs in fly genome, our work unveils a global picture of the complex interplay between non-coding RNAs and sexual chromosome evolution.

[Supplemental material is available online at http://www.genome.org.]

INTRODUCTION

Sex chromosomes are major targets for sex-related selection, and many genome-wide studies have revealed differences between sex chromosomes and autosomes with respect to divergence rate, gene content and gene expression patterns (Vicoso and Charlesworth 2006; Ellegren and Parsch 2007; Mank 2009; Qvarnstrom and Bailey 2009). Male-biased coding genes, *i.e.* genes that are more highly expressed in males than in females, are unevenly distributed between the sex chromosomes and autosomes in *Drosophila*, mammals, and worms (Betran et al. 2002; Parisi et al. 2003; Ranz et al. 2003; Khil et al. 2004; Reinke et al. 2004; Wang et al. 2005). In mammals, coding genes expressed in male meiotic or post-meiotic cells are under-represented on the X chromosome, whereas coding genes expressed in pre-meiotic stem cells are over-represented on the X chromosome (Khil et al. 2004). Similar results were obtained from transcriptional profiling of samples enriched with *D. melanogaster* meiotic cells (Vibranovski et al. 2009a). In *D. melanogaster* and in *Caenorhabditis elegans* gonads, male-biased coding genes are found to be under-represented on the X chromosome (Betran et al. 2002; Parisi et al. 2003; Ranz et al. 2003; Reinke et al. 2004). In *Drosophila* this under-representation is observed only for older X-linked coding genes. Young male-biased coding genes, those that emerged after the split of the *melanogaster* subgroup (<3-6 million years ago, Russo et al. 1995), are found to be over-represented on the X chromosome (Zhang et al. 2010a).

Directional movement of male-biased genes out of the X chromosome is one evolutionary process that could contribute to such an uneven chromosomal distribution of male-biased genes in these taxa. New *Drosophila* retrogenes tend to escape from the X chromosome and are more likely to be expressed in testis (Betran et al. 2002), and

excessive male-biased retrogene traffic has been observed on the mammalian X chromosome (Emerson et al. 2004). Further studies showed that new DNA-based duplicate coding genes exhibit a similar chromosomal distribution pattern to retrogenes (Betran et al. 2002; Emerson et al. 2004; Meisel et al. 2009; Vibranovski et al. 2009b), suggesting that the uneven chromosomal distribution of male-biased genes might not depend on a specific molecular mechanism, but rather is the product of natural selection acting on genes with male-related functions (Meisel et al. 2009; Vibranovski et al. 2009b). This hypothesis is supported by independent evidence from population genomic analysis of copy number variation of *Drosophila* retrogenes (Schrider et al. 2011).

Several different hypotheses invoking natural selection could explain the paucity of X-linked male-biased genes. First, inactivation of X-linked genes during male meiosis (meiotic sex chromosome inactivation, MSCI; (Lifschytz and Lindsley 1972)) may favor the accumulation of testis-expressed genes in autosomes (Betran et al. 2002; Emerson et al. 2004). Second, the sexual antagonism hypothesis predicts that the chromosomal distribution of sex-biased genes is driven by sexually antagonistic forces (Rice 1984), such that dominant alleles with beneficial fitness effects in males, but detrimental effects in females have a higher probability of being fixed on the autosomes (Charlesworth et al. 1987). Another dominance-independent hypothesis of sexual antagonism driving germline X inactivation, predicted demasculinization of X-chromosome based on different sojourning times of X, 1/3 of the time in males and 2/3 of the time in females, which would result in an excess male-biased genes out of X-chromosome and enrichment of X-linked female-biased genes (Wu and Xu 2003). Third, the dosage compensation hypothesis predicts that hypertranscription of the *Drosophila* X chromosome in males prevents further up-regulation of X-linked male-biased genes, thus favoring their relocation to an autosome (Vicoso and Charlesworth 2009; Bachtrog et al. 2010). While

evidence for all three hypotheses has been demonstrated using coding genes in *Drosophila* (Parisi et al. 2003; Hense et al. 2007; Kemkemer et al. 2011; Kemkemer et al. 2013; Vibranovski et al. 2009a, Bachtrog et al. 2010), no systematic experimental study on non-coding RNAs has been done so far to test these hypotheses. Non-coding RNAs (ncRNAs) play important roles in many reproductive processes (Mattick and Makunin 2005; Prasanth and Spector 2007). If selection governs the chromosomal distribution of sex-biased genes, we expect male-biased ncRNAs to exhibit a chromosomal distribution similar to that observed for coding genes.

In this report, we tested these hypotheses by experimentally identifying male-biased ncRNAs in *D. melanogaster* and analyzing their chromosomal distribution. Whole-transcriptome profiling revealed a large number of intergenic noncoding RNAs (incRNAs) with male-biased expression in both whole body and reproductive organs, which we confirmed with RT-PCR. We demonstrate that these incRNAs are unevenly distributed between the autosomes and X chromosome. Comparisons of germline and somatic tissue transcriptional profiles suggest that sexual antagonism and male germline MSCI both could be contributing to the peculiar chromosomal distributions of male-biased incRNAs. In concordance with previous studies on coding genes, comparative genomics analyses revealed that male-biased incRNAs that originated during different evolutionary periods have different chromosomal distribution patterns, indicating that evolutionary time has a significant effect on their chromosomal locations (Zhang et al. 2010a). As for coding genes (Zhang et al. 2010a), we found that old male-biased incRNAs (>6 my old) are enriched on autosomes, while new male-biased incRNAs are enriched on the X chromosome.

In addition, our analyses shed some light in the current debate about the

demasculinization of the X chromosome in *Drosophila*. More specifically, our analyses clarify why recent studies have shown that *Drosophila* testis-specific genes are not under-represented in the X chromosome, while male-biased genes are (Meiklejohn and Presgraves 2012; Meisel et al. 2012). Gene age is positively correlated with expression breadth (Zhang et al. 2012). Therefore testis-specific genes (narrowly expressed genes) are enriched with very young genes, which were previously shown to be over-represented in the X chromosome (Zhang et al. 2010a). In order to evaluate the chromosomal distribution of testis-specific genes in an unbiased way, we analyzed the old testis-specific coding genes and found that they are under-represented in the X chromosome. These results corroborate our previous findings that the process of desmasculinization is an evolutionary process that appears only over evolutionary time in both *Drosophila* and mammals (Zhang et al. 2010a; Zhang et al. 2010b).

Results and Discussion

Transcriptome profiling reveals abundant incRNAs with sex-biased expression

We first profiled the transcriptomes of whole male and female *D. melanogaster* adults using Affymetrix whole-genome tiling arrays. The arrays used 3,116,816 25-nt probe pairs to assay transcription of 109,088,560 bp of repeat-free euchromatic genome. We detected a total of 35,884,625 bp (~32.89%) in male and 32,921,857 bp (~30.18%) in female as being transcribed. Of those nucleotides transcribed, 7,738,215 bp (~21.56%) exhibited male-biased expression and 3,649,022 bp (~11.08%) exhibited female-biased expression with >2-fold increase. In addition to whole body samples, we profiled the transcriptome of adult reproductive tracts (gonads: testis, ovaries; and accessory glands) and non-reproductive tracts (gut and thorax). We identified 29,188,400 bp (~26.76%), 25,316,156 bp (~23.21%) and 25,843,450 bp (~23.69%) as being transcribed in testis, ovaries and accessory glands, respectively. Of those nucleotides transcribed, 10,066,819 bp (~34.49%) in testis and 7,633,727 bp (~29.54%) in accessory glands were identified as male-biased with >2-fold increase over expression in ovaries. As expected, a relatively low proportion of male-biased transfrags (*trans*cribed *frag*ments, see *Material and Methods* for further details) was found in gut (7.36%, 2,605,069 bp out of 35,390,445 bp) and thorax (9.51%, 3,043,360 bp out of 32,001,727 bp) (detailed transfrag counting and proportions could be found in Supplementary Figure S1 and Table S1).

According to FlyBase annotation release 5.46, 10-20% of male-biased and female-biased transfrags belong to intergenic regions (Figure 1). Analysis of coding potential with CPC

(Coding Potential Calculator) (Kong et al. 2007) suggested that more than 98% of these intergenic transfrags are truly noncoding transcripts. We now focus on these intergenic noncoding RNAs, or incRNAs. Intronic noncoding transfrags were excluded from most of our analyses because current annotations might not distinguish true intronic noncoding RNAs from novel exons of protein-coding genes or non-degraded intronic transcripts (we provide the chromosomal distribution of intronic ncRNAs in Supplementary Figure S2).

In all three comparisons performed using male/female adult transcriptome profiles derived from either whole body or reproductive tracts organs (Figure 1a-i), the fraction of incRNAs is significantly higher in sex-biased comparisons than in non-sex-biased comparisons (p < 2.2e-16, Chi square test; Figure 1a-b *vs*. Figure 1c, Figure 1d-e *vs*. Figure 1f, and Figure 1g-h *vs*. Figure 1i), suggesting as expected that incRNA expression differences between males and females tend to be associated with sex-related biological processes. As expected, comparisons between reproductive organs (testis and accessory glands vs. ovaries) identified significantly more sex-biased incRNAs than in the non-reproductive organs comparisons (32%~35% vs. 12%~16%, Chi square test, p < 2.2e-16, Supplementary Figure S3), suggesting that incRNAs, like any other transcription unit, are more prone to be involved with sex-related functions in the reproductive organs. Two examples of sex-biased intergenic noncoding RNAs are shown in Supplementary Figure S4.

Sex-biased incRNAs are non-randomly distributed between the autosomes and X chromosome

We tested the hypothesis that selection governs the chromosomal distribution of

sex-biased genes by comparing the distributions of sex-biased coding and non-coding RNAs. If male-biased ncRNAs are randomly distributed along the chromosomes, there is probably no selection forces acting on non-coding regions and this is a unique property of the coding genes. As predicted by our hypothesis, an excess of autosomal male-biased incRNAs were identified in whole body (21% excess) and testis (21%) (Figure 2, Odds ratios (ORs) = 1.21, $p \leq 0.01$, Fisher exact test, for each comparison. We applied the Fisher exact test to assess the uneven chromosomal distribution using the estimated Odds ratio (OR) as an intuitive measurement. Odds ratio is the ratio between male-biased genes (autosomal/X-linked) and non-male biased genes (autosomal/X-linked). Thus, Odds ratio greater than 1 indicates male-biased genes are enriched on autosomes, and less than 1 indicates X enrichment. See *Materials and Methods* for more details). A similar trend was found for coding transfrags, consistent with previous studies (Parisi et al. 2003; Ranz et al. 2003; Emerson et al. 2004; Sturgill et al. 2007) (Figure 2). In contrast, we found that female-biased incRNAs are significantly over-represented on the X chromosome in comparisons of ovary *vs*. testis and ovary *vs*. accessory gland (OR range = 0.77-0.86, $p <$ 6.53e-03, Figure 2b and 2c). No significant departure from random chromosomal distribution, however, was observed for female-biased incRNAs derived from whole body female *vs.* male comparison (OR = 1.01, p = 0.44) (Figure 2a).

We verified sex-biased expression of incRNAs using tissue-specific RT-PCR. We previously detected 528 incRNAs in adult flies in a random sample of *D. melanogaster* intergenic regions using the RT-PCR approach (Li et al. 2009). Using the same primers, we detected about half of these incRNAs (261 out of 528) in testes, ovaries, and heads in *Drosophila* adults (RT-PCR primers are listed in Supplementary Table S2). We considered testis-biased incRNAs those that were detected in testis but not in ovaries. The reverse logic was applied to ovary-biased incRNAs. We considered non-sex biased

incRNAs those that were detected in both testis and ovary and those that were detected only in head. The chromosomal distribution of testis-biased and ovary-biased incRNAs differed significantly from the distribution of non-sex biased incRNAs (Supplementary Table S3, p = 0.0287, Fisher exact test), with a significant deficiency of X-linked testis-biased incRNAs (OR = 0, p = 0.0223 for testis vs. non-sex biased incRNAs) and a marginal enrichment of X-linked female-biased incRNAs (OR=1.35, n.s., probably due to small sample size). Thus the PCR-based independent test revealed the same robust chromosomal distribution patterns for *Drosophila* male-biased incRNAs, verifying that the observed under-representation of male-biased incRNAs in the X chromosome is not a methodological artifact.

Disentangling the contribution of sexual antagonism and MSCI to the demasculinization of the X chromosome

X chromosome demasculinization is the evolutionary process by which selective forces drive male-biased genes off of the X chromosome, either relocating them to the autosomes or eliminating them from the genome entirely. We investigated the contribution of MSCI and sexual antagonism to the observed X chromosome demasculinization for incRNAs in *Drosophila*. It is not trivial to separate the effects of sexual antagonism and MSCI, as MSCI only occurs in male meiosis, but sexual antagonism may occur in any tissue or cell. However, most sex-biased expression is found in testes and ovaries, especially in the meiotic phase (Parisi et al. 2003; Vibranovski et al. 2009a). MSCI could therefore be assessed by analyzing testis-expressed genes with biased expression in meiosis but not in mitosis, thus including the effect of inactivation of X-linked genes in meiotic cells but ignoring sexual antagonistic effects present in mitotic cells (Vibranovski et al. 2009a). However, meiotic cells could also be under the effect of sexual antagonistic forces

preventing the complete separation of those two processes. We thus identified male-biased RNAs involved in meiosis as those that are testis-biased but not accessory-gland biased, as accessory glands only contain mitotic cells. Accessory glands produce proteins and compounds that comprise seminal fluid and affect the reproductive capacity of both sexes (Ravi Ram and Wolfner, 2007). Accessory-gland biased genes are therefore potential sexually antagonistic genes. We observed a statistically significant over-representation of strictly testis-biased incRNAs on the autosomes (OR = 1.20; p = 0.0069), suggesting that MSCI contributes to the desmaculinization of the X chromosome despite the effect of accessory gland expressed genes with sexual antagonistic effects. It is possible that there are sexually antagonistc genes expressed in testis mitotic cells that are not expressed in accessory gland mitotic cells. Therefore our data only suggest the role of X-inactivation in producing the paucity of X-linked male-biased incRNAs. Nevertheless, the role of MSCI observed for incRNAs was also observed for coding exons (OR = 1.50, p = 1.07e-31) and is consistent with previous observations derived from protein-coding genes expressed in meiosis (Vibranovski et al. 2009a).

Conversely, we assessed the effects of sexual antagonism by comparing incRNAs with biased expression in accessory gland (somatic) but unbiased in testis (spermatogenesis). Statistical tests showed no significant X demasculinization for these incRNAs (OR = 1.01, p = 0.464). Although we found no X demasculinization of accessory gland-biased incRNAs, it should be noted that accessory gland-biased coding exons are significantly under-represented on the X-chromosome even after removing genes that are also testis-biased (OR = 1.41, p = 2.82e-12). Moreover, sex-biased genes in other male-specific tissues may be sexually antagonistic; accessory gland is not the sole male-specific somatic tissue in *Drosophila* (e.g. male genitalia (Liu et al. 1996)).

To further investigate the effects of sexual antagonism in other somatic tissues, we performed additional transcriptome profiling on non-reproductive organs: gut and thorax of male and female adults. No significant chromosomal distribution imbalance for either male-biased or female-biased coding exons was found (Figure 2d and 2e). Female-biased incRNAs expressed in the thorax and gut do not deviate from the random chromosomal distribution in both tissues. Male-biased incRNAs are also randomly distributed on chromosomes, except those expressed in guts which are significantly over represented on the X chromosome (Figure 2d), the opposite pattern expected for demasculinization. One possible explanation for the excess of X-linked male-biased genes found only in guts is a higher proportion of young genes which are known to be found in excess in the X chromosome (Zhang et al 2010). Indeed, 10% of male-biased genes expressed in guts were originated less than 3 million years ago in comparison to 7% of those expressed in thorax (Fisher Exact test, $p = 0.066$). The generally random chromosomal distribution of sex-biased genes in gut and thorax, except for the excess of male-biased incRNAs in the X, suggests that the demasculinization of genes is more often associated with reproductive organs.

Moreover, our entire data, which combines reproductive and non-reproductive organs, suggests that the X demasculinization effects of sexual antagonism are limited to accessory gland-biased coding genes. These results support the hypothesis that sexual antagonism probably contributes less than MSCI to the non-random chromosomal distribution of male-biased genes in *Drosophila*.

Our results from non-reproductive organs do not support the involvement of dosage compensation in generating the paucity of male-biased protein coding genes observed in *Drosophila* X chromosome (Vicoso and Charlesworth 2009; Bachtrog et al. 2010). That

is, although thorax and gut also experience dosage compensation, we observed no paucity of X-linked male-biased protein coding genes in those tissues.

X chromosome demasculinization

Two recent studies have examined the demasculinization of the X chromosome in *Drosophila* and the contribution of MSCI to the phenomenon (Meiklejohn and Presgraves 2012; Meisel et al. 2012). In both papers, it was claimed that a deficit of male-biased genes on the X chromosome is attributable solely to lower average expression of genes on the X relative to the autosomes in *Drosophila* testes, most likely due to an absence of dosage compensation in the germline (Meiklejohn and Presgraves 2012; Meisel et al. 2012). In both studies, this argument is mainly based on the random chromosomal distribution of testis-specific genes (*i.e.* those that are expressed in testis, but not in any other tissue) as opposed to the deficit of testis-biased genes (*i.e.* those that are expressed more in testis than ovary) on the X chromosome. The interpretations of those results are the following: i) MSCI is not a factor that contributes to the non-random chromosomal distribution testis-biased genes because the X chromosome shows no deficit of testis-specific genes that, in theory, should be under meiotic inactivation; ii) deficit of testis-biased genes should be attributed to lack of dosage compensation in the male germline as comparisons between testis and ovaries are the only analysis that present the deficit; iii) direct comparisons between testis and ovary expression do not control for correlation between expression breadth and sex-biased expression. Therefore, by comparing those sex-specific tissues, one might obtain evidence for paucity of X-linked testis expressed genes.

However, those studies did not take into account the importance of gene age when

looking to the random chromosomal distribution of testis-specific genes (Meiklejohn and Presgraves 2012; Meisel et al. 2012). It is now well established that young male-biased genes both in *Drosophila* and mammals tend to be more frequently found in the X chromosome, whereas the opposite pattern is found for older male-biased genes (Zhang et al. 2010a; Zhang et al. 2010b). Therefore, the process of desmasculinization is an evolutionary process that appears over evolutionary time. In order to evaluate the chromosomal distribution of testis-specific genes in an unbiased way we should take into account only older genes. Moreover, neither study accounted for the relationship between gene age and expression breadth. Testis-specific genes are genes narrowly expressed in testis. The young genes tend to be more narrowly expressed than the older genes (Zhang et al. 2012). We tested the hypothesis that the random chromosomal distribution of testis-specific, but not of testis-biased, genes is caused by a large number of testis-specific genes being newly-evolved genes. We classified coding genes according to their expression breadth and separated older and young genes according to ages defined Zhang et al. (2010a), branch 0 and branches 1-6, respectively. First, we confirmed that narrowly expressed coding genes are enriched with new genes (Figure S6; Zhang et al. 2012). This pattern is also true for testis-specific coding genes (Figure S6). Second, young testis-specific coding genes are enriched on the X chromosome, whereas older testis-specific coding genes are deficient from the X chromosome (Figure 3B and 3C, respectively). We therefore conclude that the result of random distribution of testis-specific coding genes (Meiklejohn and Presgraves 2012; Meisel et al. 2012) is a consequence of the enrichment of testis-specific coding genes with recently-evolved coding genes in a short evolutionary period. Therefore, neither demasculization nor MSCI can be ruled out as important players in determining the chromosomal distribution of male-biased coding genes in *Drosophila* as older testis-specific coding genes are under-represented in the X chromosome (Figure 3C).

Although there is no argument over the presence of X chromosome reduced expression in male germline cells, there were different opinions for the period of this expression reduction. MSCI studies presented evidence of expression reduction in the meiotic stage (e.g. Vibranovski et al. 2009) whereas one study believed that the reduction is also extended to the mitotic stage of spermatogenesis (Miklejohn et al. 2011). However, recent analyses reported three independent lines of evidence in favor of MSCI analyzing the expression of testis-specific promoter reporters, testis from larval stages and from meiotic arrest mutants (Kemkemer et al. 2013; Vibranovski et al. 2012; Deng et al. 2011). Nevertheless, the work present here does not provide evidence in favor or against MSCI, but the patterns found are consistent with the phenomenon.

Male-biased incRNA gene linkage depends on gene age

Both male-biased incRNAs and male-biased coding transfrags are significantly deficient from the X chromosome, although this trend is stronger for coding than for noncoding transfrags (Figure 2). The comparison of testis-biased coding vs. non-coding transfrags shows OR = 1.21 vs. 1.49 (p = 6.63e-04) as well as the entire gene unit vs. non-coding transfrags shows OR = 1.21vs. 1.39 (p = 0.026, see Supplementary Table S4 for details). Do different evolutionary ages of male-biased coding and non-coding genes play any role in determining the evolutionary dynamics? We inferred the evolutionary age of incRNAs through comparative sequence analysis of the 12 sequenced *Drosophila* genomes (Clark et al. 2007; Stark et al. 2007). Given the relatively fast evolutionary rate of incRNAs (Pang et al. 2006), we took a conservative dating strategy (See *Material and Methods*). Following the parsimony principle, 23,165 incRNAs were assigned to a unique phylogenetic branch compared to out-group species. Among those, 2,660 (or 11.48 % out

of 23,165) were identified as "male-biased" in *D. melanogaster* during at least one comparison (Figure 4a). By comparing the chromosomal distribution of incRNAs across two age groups, *i.e.*, before and after the split of the *melanogaster* subgroup (3-6 million years ago, Figure 4), we found that the older male-biased incRNAs, those that originated before the split, are significantly enriched on autosomes (OR = 1.18, p = 0.037 for whole-body and OR = 1.38, p = 3.096e-4 for gonad). In contrast, young male-biased incRNAs show the opposite pattern, they are enriched on the X chromosome, in both whole-body (OR = 0.72, p = 0.05) and gonad (OR = 0.71, p = 0.03) comparisons. Furthermore, we found that a significantly larger percentage (10.86%, or 289 out of 2,660) of male-biased incRNAs emerged very recently (less than 3 million years, on branches 5 and 6 in Figure 4a) compared to male-biased coding genes that emerged during the same period (4.03%, or 400 out of 9,931) (Fisher Exact test, OR = 2.90, p < 2.2e-16). A significantly negative correlation (Spearman correlation rho= 0.95, p = 0.0008) between the age of the lineages and the proportion of X-linked male-biased incRNAs was observed (Figure 4b).

The age analyses implemented in this study work on the DNA sequence level. That means, for sex-biased and unbiased transfrags, we can infer the age of the corresponding DNA locus based on presence or absence information. However, although the DNA sequence can be old, the transcription pattern may have only recently evolved. In this sense, our strategy provides an upper age estimate for the expression pattern. Since male-biased expression has higher turnover rate, such an approximation maybe too conservative and the age of male-biased transfrags could have been overestimated. Therefore, there could be an even larger proportion of younger male-biased incRNAs, further strengthening our conclusions.

Zhang et al. (2010a) reported an excess of X-linked new protein coding genes in *Drosophila* that had been recently generated from DNA-level duplication or *de novo* gene origination, and that the proportion of male-biased genes among the X-linked new genes diminishes with gene age. The incRNA genes in this study show a similar pattern to the new protein-coding genes. However, it appears that turnover is more recent in the incRNA genes, consistent with the known rapid evolution of ncRNA genes (Pang et al. 2006) and higher turnover rate of microRNA genes (Lu et al. 2008). These data indicate that different evolutionary forces, e.g. MSCI and sexual antagonism, might play roles at different evolutionary timescales (Zhang et al. 2010a).

In summary, we experimentally identified male-biased non-coding RNAs in *D. melanogaster* and analyzed their chromosomal distribution. The identification of a large number of incRNAs that showed male-biased expression patterns may explain the signals of natural selection previously detected in the non-coding genomic regions (Andolfatto 2005). By systematically profiling the whole transcriptome of *D. melanogaster* male and female adult whole bodies as well as reproductive tract organs, we revealed a long-term removal of male-biased incRNA genes from the X chromosome resulting in, an uneven distribution of male-biased incRNAs between X and autosomes. This led to the long-term X chromosome demasculinization, probably through sexual antagonism and MSCI. Finally, we identified distinctive chromosomal preferences between young and old male-biased incRNAs. This pattern of male-biased incRNAs further generalizes the uneven pattern of male-biased gene content on *Drosophila* autosomes and X chromosomes and suggests that the pattern is shaped by natural selection acting on male functions. Our results contribute to a global picture of the sex chromosome evolution in the *Drosophila* genome.

METHODS

Data sources

Genome sequence and annotations were obtained from UCSC genome database (http://genome.ucsc.edu, dm2). Gene model annotations were obtained from FlyBase (http://www.flybase.org, v5.46, downloaded in March, 2013). The Affymetrix array BPMAP annotation was downloaded from the Affymetrix website (http://www.affymetrix.com/products/arrays/specific/drosophila_tiling1_0r.affx). The 12 sequenced *Drosophila* genome sequences (CAF1) were downloaded from http://rana.lbl.gov/drosophila/.

Sample preparation and microarray hybridization

We extracted total RNA from whole bodies, thoraxes, digestive tracts (guts), testes, accessory glands, and ovaries from virgin Oregon R adults using the Qiagen RNeasy Mini Kit (CA, USA) with on-column DNase digestion. All flies were virgin and aged for 1-6 days post-eclosion before being used in extractions or dissections. Whole body and thorax RNA was immediately extracted from 50 males or females. Gut tissue was completely removed from thoraxes. Testes, accessory glands, ovaries, and guts were dissected, placed in RNA*later* (Qiagen, CA, USA), and stored at -20ºC until RNA extraction. Roughly 200 testes, accessory glands or ovaries, or 100 guts were required for each replicate. Testes were separated from all other reproductive tract tissues (seminal vesicles, accessory glands, and ducts). For statistical independence, all tissues were harvested from independent sets of flies. Each of the biological replicates was labeled

with GeneChip® WT Sense Target Labeling and Control Reagents (Part#: 900652) and hybridized to an Affymetrix™ *D. melanogaster* genome tiling array as previously reported by (Manak et al. 2006). Labeling and hybridizations were performed at the University of Chicago Functional Genomics Facility. Three replicates of hybridization were performed for each tissue.

Tiling array data processing and analysis

Affymetrix Tiling Analysis Software (TAS v1.1.02) was used to process raw tiling array data (Cheng et al. 2005; Manak et al. 2006). Raw data were normalized by quantile normalization and the median of target intensities was scaled to 100. As suggested in Affymetrix TAS user manual (http://www.affymetrix.com/support/developer/downloads/TilingArrayTools/index.affx), each probe position was analyzed in a local smoothing window with bandwidth (BW) equal to 50 bp (resulting in a window width of 101 bp) for better statistical power. To assess the performance of replicates, standard Pearson correlation coefficients between replicates were calculated pair-wise. The significant correlation (spearman correlation rho > 0.98, p-value = 1e-5) indicated reasonable consistency between replicated samples.

A one-tail Wilcoxon signed-rank test on all the probes in the window was performed with the alternative hypothesis that the true intensity difference between the perfect match (PM) probe and mismatched probe is significantly greater than zero. Only probes with a p-value < 0.1 were called "positive". Neighboring positive probes with max-gap 50bp and a minimum run of 90bp were grouped as transfrags (*trans*cribed *frag*ments), then *Unified Transfrags*, or *UTS*, were further derived by assembling overlapped transfrags ("*supporting transfrags*") in different samples as suggested by previous literature (Cheng

et al. 2005; Manak et al. 2006). "Present" call was produced for each unified transfrag of UTS. A unified transfrag will be called "Present" in a given sample if and only if at least one of its supporting transfrags is identified in this sample. Moreover, the median intensity value of all comprising probe sets within each unified transfrag was calculated for each sample. Among UTS, ~48% unified transfrags are tissue-specifically transcribed, and 78% consist of multiple supporting transfrags with >=50% overlap. Further inspection indicated that 73.52% (68,310 out of 92,916) of coding exons on autosomes (chr2L, 2R, 3L, and 3R) and X chromosome annotated in FlyBase were detected as expressed in at least one of the five *D. melanogaster* samples (adult male/female whole body and three reproductive organ tissues (testis, ovary and accessory gland)) with >=70% coverage.

An integrated procedure was applied to assess sample-biased expression for each unified transfrag (Figure S5). To be conservative, a global Kruskal-Wallis test ("non-parametric one-way ANOVA") among the five samples was applied before the pair-wise Mann–Whitney U test. To assess the detection power of our procedure, we compared the identified sex-biased protein-coding transfrags to the Sebida database that integrates data derived from multiple previous high-throughput studies comparing male versus female protein-coding genes expression in *D. melanogaster* (Gnad and Parsch 2006). Up to 88% (613 out of 695) male-biased and 84% (663 out of 789) female-biased genes in the Sebida Two-fold high quality dataset were also identified by our procedure with >=70% length coverage, suggesting a high consistency of our procedure with previous studies.

Annotation of detected transfrags

Transfrags were classified as "coding" or "intergenic" based on their genomic

coordinates. Because our work focused on noncoding genes, as a conservative estimation, we only consider a transfrag as "intergenic" if it does not overlap with any annotated FlyBase protein-coding gene models on both the sense and antisense strands. We further assessed the coding potentials of these intergenic transfrags by a SVM-based classifier (Kong et al. 2007), the results indicated that >98% intergenic transfrags are truly non-coding. After excluding 945 transfrags showing putative coding potential at either sense or antisense strand, we classified the remaining transfrags as "intergenic noncoding transfrags". The intronic transfrags were excluded from follow-up analyses as current annotations might not distinguish true intronic noncoding RNAs from novel exons of protein-coding genes or non-degraded intronic transcripts. Moreover, in case of potential bias resulted by the relatively larger exon number in protein-coding genes (when comparing to non-coding RNAs), we re-run analysis in gene level by assigning transfrags to annotated FlyBase genes according to the coordination.

Assessing the relationship between expression breadth and gene age for protein-coding genes

According to Meisel et al. (2012), microarray signal intensities from 14 adult *D. melanogaster* tissues were obtained from FlyAtlas (Chintapalli et al. 2007). Expression breadth was calculated according to the tissue specificity index, τ (Yanai et al. 2005). Genes were considered as narrowly (tissue-specific) and broadly expressed depending of their τ value ($\tau > 0.7$ and $\tau \leq 0.4$, respectively). Testis-specific genes were considered to be encoded proteins in the sperm proteome if they were found at least in one of the two sperm proteomes (Dorus et al. 2006; Wasbrough et al. 2010). Gene age was obtained by crosslinking CG identifiers with information available from Zhang et al. (2010a).

Comparative genomics analysis to infer evolutionary ages along the *Drosophila* phylogeny

In silico comparative sequence analysis was performed with all 12 sequenced Drosophila genomes similar to the procedure reported by Sturgill *et al.* (Sturgill et al. 2007). We ran NCBI BLAST against genomic DNA of each species. To handle the relatively low sequence conservation of noncoding genes, optimized BLAST parameters were employed as suggested in previous literatures (Korf et al. 2003; Freyhult et al. 2007). A *D. melanogaster* incRNA is called "absent" in another species if there are no hits with E-value < 1e-4 and coverage > 80% found in that species. After making the "present"/"absent" call for each incRNA, we dated their origination along the *Drosophila* genus phylogenetic tree (Figure S7) following the parsimony principle described below.

incRNA *x* is assigned to branch *X* if and only if it is called "present" in all in-group species of branch *X* and "absent" in all out-group species of *X*. For example, branch 0 includes incRNAs that are "present" in all 12 sequenced *Drosophila* genomes. And branch 4 include incRNAs that are "present" in *D.mel, D.sim, D.sec, D.yac*, and *D.ere*, but "absent" in *D.ana, D.pse, D.per, D.wil, D.moj, D.vir,* and *D.gri*.

Assessment of the uneven distribution of the transfrags among X chromosome and autosomes

We applied the Fisher exact test to assess the uneven chromosomal distribution using the estimated Odds ratio (OR) as an intuitive measurement:

	# of transfrags on Autosomes	# of transfrags on X chromosome
# of male-biased transfrags	*a*	*b*
# of non-male-biased transfrags	*c*	*d*

$$OddRatio = \frac{\frac{a}{b}}{\frac{c}{d}}$$

Thus, Odds ratio greater than 1 indicates male-biased genes are enriched on autosomes, and less than 1 indicates X enrichment.

All statistically computations were performed by R (http://www.r-project.org) and in-house scripts available on request.

Independent RT-PCR assay

RNA extraction. Adults were collected within 10 days after eclosion. Tissues like heads, ovaries, testes were separately dissected into tubes with TRIzol (Invitrogen Company). RNAs were extracted following TRIzol reagent instructions. About 10ug of sample RNA was mixed with 20ul RQ1 RNase-free DNase (1unit/ul; Promega Company), 10ul 10× DNase buffer, 2ul RNase inhibitor HPR1 (Takara Company) and DEPC water (up to 100ul), and incubated for 3 hours at 37℃. Contamination of RNAs with DNA was ruled out by PCR amplification of 2 pairs of primers for Gapdh2 and II171a taking the extracted RNAs as template.

Reverse transcription. 0.5ug 6-mer random primer (Takara Company) and 5uL dNTPs (2.5mM, Takara Company) per microgram of RNA sample were mixed in total volume of ≤ 15uL in a tube; the tube was heated to 70℃ for 5 min to melt secondary structure; the

tube was cooled immediately on ice for at least 3min to prevent secondary structure from reforming; 5uL M-MLV 5×Reaction Buffer (Promega Company), 5uL dNTPs (2.5mM, Takara Company)(to be deleted), 0.5uL RNase Inhibitor (HPR I, Takara Company), and 1uL M-MLV RTase (Promega Company) were added to the annealed primer/template. Then add DEPC water was added up to the volume of 25uL. PCR of primer II171a was conducted to guarantee no genomic DNA contamination; and PCR of Gapdh2 was conducted to guarantee the quality of cDNA.

Data Access

The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (Edgar et al. 2002) and are accessible through GEO Series accession number GSE53421 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53421).

ACKNOWLEDGMENTS

LW was supported by National Outstanding Young Investigator award from Natural Science Foundation of China (31025014), 973 grant from China Ministry of Science and Technology (2012CB837600), and China Ministry of Education 111 Project (B06001). The research of GG was supported in part by the National Outstanding Youth Talent Initiative Program. ML is currently supported by NIH grants 1R01GM100768-01A1 and NSF1051826 for investigating genetics and evolution in *Drosophila*. MDV was supported by Pew Latin America Fellowship and Brazilian Counsel of Technological and Scientific Development (CNPq). NWV was partially supported by National Institutes of Health Grant T32 GM007197.

FIGURE LEGENDS

Figure 1. Genomic distribution of sex-biased transfrags. Expression profiling was done with Affymetrix whole-genome tiling arrays. Exon/intron/intergenic annotations were retrieved from FlyBase (version 5.46). Rows represent comparisons of male and female whole body RNA (a to c), testis versus ovary (d to f), accessory gland versus ovary (g to i), male versus female gut (j to l) and male versus female thorax (m to o). Columns represent male-biased (a, d, g, j, m), female-biased (b, e, h, k, n), and non-sex-biased expression (c, f, i, l, o).

Figure 2. X chromosome demasculinization was observed for incRNAs as well as coding transfrags, based on comparison of whole body of males versus females (a), testes versus ovaries (b) and accessory glands versus ovaries (c). Odds ratios greater than 1 indicate enrichment on autosomes, and less than 1 indicates enrichment on X chromosome. Significant deviations are indicated (*** $p < 0.001$, ** $p < 0.01$, Fisher exact test). Blue and red bars represent male-biased and female-biased transfrags, respectively.

Figure 3. Percentage of *D. melanogaster* X-linked genes that are broadly and narrowly expressed. Following methods in Meisel et al. (2012), genes narrowly expressed also called specific genes ($\tau > 0.7$) in each of four sex-limited tissues, with testis-specific 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 ($\tau \leq 0.4$). Percent of X-linked in the genome are shown in dashed lines. Significant deviations (Fisher exact test) are indicated (***$P < 0.001$; **$P <$

0.01, *P < 0.02). All X-linked genes (A) were separated according to their evolutionary age. New (B) and old (C) genes were defined according to Zhang et al (2010a) where old genes are at least as old as the split between Sophophora and Drosophila subgenera. At first (A), there is no paucity of testis-specific in the X chromosome. However, opposite patterns are found for old and young genes: while old testis-specific genes are under-represented (C), new testis-specific genes are found in excess in the X chromosome (B).

Figure 4. Analysis of incRNAs with different ages. (A) Numbers of newly originated incRNAs in each age branch inferred by comparative genomics analysis. For each branch, the count of male-biased and all incRNAs were given as the underlined numbers in parentheses, separated by a slash ("/"). (B) Proportions of male-biased incRNAs among all identified male-biased transfrags in each age branch. Significant correlation between the age of the lineages and the proportion of male-biased ncRNAs was observed (spearman correlation rho = 0.89, p = 0.01). The phylogeny and divergence times were from ref (Stark et al. 2007)

References

Andolfatto P. 2005. Adaptive evolution of non-coding DNA in Drosophila. *Nature* **437**(7062): 1149-1152.

Bachtrog D, Toda NR, Lockton S. 2010. Dosage compensation and demasculinization of X chromosomes in Drosophila. *Curr Biol* **20**(16): 1476-1481.

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

Charlesworth B, Coyne J, Barton N. 1987. The relative rates of evolution of sex chromosomes and autosomes. *American Naturalist* **130**(1): 113.

Cheng J, Kapranov P, Drenkow J, Dike S, Brubaker S, Patel S, Long J, Stern D, Tammana H, Helt G et al. 2005. Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science* **308**(5725): 1149-1154.

Chintapalli VR, Wang J, Dow JA. 2007. Using FlyAtlas to identify better Drosophila melanogaster models of human disease. *Nature genetics* **39**(6): 715-720.

Clark AG Eisen MB Smith DR Bergman CM Oliver B Markow TA Kaufman TC Kellis M Gelbart W Iyer VN et al. 2007. Evolution of genes and genomes on the Drosophila phylogeny. *Nature* **450**(7167): 203-218.

Deng X, Hiatt JB, Nguyen DK, Ercn 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*. *Nature 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. *Nature genetics* **38**(12): 1440-1445.

Edgar R, Domrachev M, Lash AE 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**(1): 207-210.

Ellegren H, Parsch J. 2007. The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev Genet* **8**(9): 689-698.

Emerson JJ, Kaessmann H, Betran E, Long M. 2004. Extensive gene traffic on the mammalian X chromosome. *Science* **303**(5657): 537-540.

Freyhult EK, Bollback JP, Gardner PP. 2007. Exploring genomic dark matter: a critical assessment of the performance of homology search methods on noncoding RNA. *Genome Res* **17**(1): 117-125.

Gnad F, Parsch J. 2006. Sebida: a database for the functional and evolutionary analysis of genes with sex-biased expression. *Bioinformatics* **22**(20): 2577-2579.

Hense W, Baines JF, Parsch J. 2007. X chromosome inactivation during Drosophila spermatogenesis. *PLoS Biol* **5**(10): e273.

Kemkemer C, Catalán A, Parsch J. 2013. 'Escaping' the X chromosome leads to increased gene expression in the male germline of *Drosophila melanogaster*. *J. Heredity* (Epub ahead of print).

Kemkemer C, Hense W, Parsch J. 2011. Fine-scale analysis of X chromosome inactivation in the male germ line of *Drosophila melanogaster*. *Mol Biol Evol.* **28**(5): 1561-1563.

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**(6): 642-646.

Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, Gao G. 2007. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res* **35**(Web Server issue): W345-349.

Korf I, Yandell M, Bedell J. 2003. *Blast*. O'Reilly Media, Inc.

Li Z, Liu M, Zhang L, Zhang W, Gao G, Zhu Z, Wei L, Fan Q, Long M. 2009. Detection of intergenic non-coding RNAs expressed in the main developmental stages in Drosophila melanogaster. *Nucleic Acids Res*.

Lifschytz E, Lindsley DL. 1972. The role of X-chromosome inactivation during spermatogenesis. *Proc Natl Acad Sci U S A* **69**(1): 182-186.

Liu J, Mercer JM, Stam LF, Gibson GC, Zeng ZB, Laurie CC. 1996. Genetic analysis of a morphological shape difference in the male genitalia of Drosophila simulans and D. mauritiana. *Genetics* **142**(4): 1129-1145.

Lu J, Shen Y, Wu Q, Kumar S, He B, Shi S, Carthew RW, Wang SM, Wu C-I. 2008. The birth and death of microRNA genes in *Drosophila*. *Nat. Genet.* **40**: 351–355.

Manak JR, Dike S, Sementchenko V, Kapranov P, Biemar F, Long J, Cheng J, Bell I, Ghosh S, Piccolboni A et al. 2006. Biological function of unannotated transcription during the early development of Drosophila melanogaster. *Nat Genet* **38**(10): 1151-1158.

Mank J. 2009. Sex Chromosomes and the Evolution of Sexual Dimorphism: Lessons from the Genome. *The American Naturalist* **173**(2): 141-150.

Mattick JS, Makunin IV. 2005. Small regulatory RNAs in mammals. *Hum Mol Genet* **14 Spec No 1**: R121-132..

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**(8): e1001126.

Meiklejohn CD, Presgraves DC. 2012. Little evidence for demasculinization of the Drosophila X chromosome among genes expressed in the male germline. *Genome biology and evolution* **4**(10): 1007-1016.

Meisel RP, Han MV, Hahn MW. 2009. A Complex Suite of Forces Drives Gene Traffic from Drosophila X Chromosomes. *Genome Biol Evol* **2009**(0): 176-188.

Meisel RP, Malone JH, Clark AG. 2012. Disentangling the relationship between sex-biased gene expression and X-linkage. *Genome research* **22**(7): 1255-1265.

Pang KC, Frith MC, Mattick JS. 2006. Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. *Trends Genet* **22**(1): 1-5.

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**(5607): 697-700.

Prasanth KV, Spector DL. 2007. Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum. *Genes Dev* **21**(1): 11-42.

Qvarnstrom A, Bailey RI. 2009. Speciation through evolution of sex-linked genes. *Heredity* **102**(1): 4-15.

Ranz JM, Castillo-Davis CI, Meiklejohn CD, Hartl DL. 2003. Sex-dependent gene expression and evolution of the Drosophila transcriptome. *Science* **300**(5626): 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.

Reinke V, Gil IS, Ward S, Kazmer K. 2004. Genome-wide germline-enriched and sex-biased expression profiles in Caenorhabditis elegans. *Development* **131**(2): 311-323.

Rice W. 1984. Sex chromosomes and the evolution of sexual dimorphism. *Evolution*: 735-742.

Russo CAM, Takezakib N, Nei M. 1995. Molecular Phylogeny and Divergence Times of Drosophilid Species. *Mol. Biol. Evol.* 12: 391-404.

Schrider DR, Stevens K, Cardeno CM, Langley CH, Hahn MW. 2011. Genome-wide analysis of retrogene polymorphisms in Drosophila melanogaster. *Genome research* **21**(12): 2087-2095.

Stark A, Lin MF, Kheradpour P, Pedersen JS, Parts L, Carlson JW, Crosby MA, Rasmussen MD, Roy S, Deoras AN et al. 2007. Discovery of functional elements in 12 Drosophila genomes using evolutionary signatures. *Nature* **450**(7167): 219-232.

Sturgill D, Zhang Y, Parisi M, Oliver B. 2007. Demasculinization of X chromosomes in the Drosophila genus. *Nature* **450**(7167): 238-241.

Vibranovski MD, Zhang YE, Kemkemer C, Lopes HF, Karr TL and Long M. 2012.

Re-analysis of the larval testis data on meiotic sex chromosome inactivation revealed evidence for tissue-specific gene expression related to the drosophila X chromosome. *BMC Biology* **10**:49.

Vibranovski MD, Lopes HF, Karr TL, Long M. 2009a. Stage-specific expression profiling of Drosophila spermatogenesis suggests that meiotic sex chromosome inactivation drives genomic relocation of testis-expressed genes. *PLoS Genet* **5**(11): e1000731.

Vibranovski MD, Zhang Y, Long M. 2009b. General gene movement off the X chromosome in the Drosophila genus. *Genome Res* **19**(5): 897-903.

Vicoso B, Charlesworth B. 2006. Evolution on the X chromosome: unusual patterns and processes. *Nat Rev Genet* **7**(8): 645-653.

-. 2009. Effective population size and the faster-X effect: an extended model. *Evolution* **63**(9): 2413-2426.

Wang PJ, Page DC, McCarrey JR. 2005. Differential expression of sex-linked and autosomal germ-cell-specific genes during spermatogenesis in the mouse. *Hum Mol Genet* **14**(19): 2911-2918.

Wasbrough ER, Dorus S, Hester S, 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**(11): 2171-2185.

Wu CI, Xu EY. 2003. Sexual antagonism and X inactivation--the SAXI hypothesis. *Trends Genet.* **19**(5): 243-247.

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**(5): 650-659.

Zhang YE, Landback P, Vibranovski M, Long M. 2012. New genes expressed in human brains: implications for annotating evolving genomes. *BioEssays : news and reviews in molecular, cellular and developmental biology* **34**(11): 982-991.

Zhang YE, Vibranovski MD, Krinsky BH, Long M. 2010a. Age-dependent chromosomal distribution of male-biased genes in Drosophila. *Genome research* **20**(11): 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**(10): e1000494.