Evidence for a sexual dimorphism in gene expression noise in metazoan species

Carlos Diaz-Castillo

Many biological processes depend on very few copies of intervening elements, which makes such processes particularly susceptible to stochastic fluctuations of these elements. The intrinsic stochasticity of certain processes is propagated across biological levels causing genotype- and environment-independent biological variation, which might permit populations better coping with variable environments. Biological variation of stochastic nature might also allow the accumulation of variation at the genetic level hidden from natural selection, which might have a great potential for population diversification. The study of any mechanism that resulted in the modulation of stochastic variation is, therefore, of potential wide interest. Here, I propose that sex might be an important modulator of the stochastic variation in gene expression, i.e., gene expression noise. Based on known associations between different patterns of gene expression variation, I hypothesize that in metazoans gene expression noise might be generally larger in heterogametic than in homogametic individuals. I directly test this hypothesis by comparing putative genotype- and environment-independent variation in gene expression between females and males of Drosophila melanogaster strains. Also, considering the potential effect of the propagation of gene expression noise across biological levels, I test indirectly the existence of a metazoan sexual dimorphism in gene expression noise by analyzing putative genotype- and environment-independent variation in phenotypes related with the interaction with the environment in D. melanogaster strains and metazoan species. As hypothesized, the results of these analyses are consistent with gene expression being generally noisier in heterogametic than in homogametic individuals. Further analyses and discussion of existing literature permits speculating the sexual dimorphism in gene expression noise is ultimately based on the nuclear dynamics in gametogenesis and very early embryogenesis of sex-specific chromosomes, i.e., Y and W chromosomes.
Evidence for a sexual dimorphism in gene expression noise in metazoan species

Carlos Díaz-Castillo
Department of Developmental & Cell Biology
University of California, Irvine
Irvine, CA 92697-2305
USA

Corresponding Author:
Carlos Díaz-Castillo,
Department of Developmental & Cell Biology
4311 Natural Sciences II
University of California, Irvine
Irvine, CA 92697-2305, USA
Phone number: (949) 892-9438
crlsdiazcastillo@gmail.com; cdiazcas@uci.edu
www.crlsdiazcastillo.wix.com/molcellbiodiversity
INTRODUCTION

Biological systems are prone to vary even in the absence of genetic modifications or environmental changes (Burga & Lehner 2012; Feinberg & Irizarry 2010; Kaern et al. 2005; Kilfoil et al. 2009; Lehner 2013; Raj & van Oudenaarden 2008; Raser & O'Shea 2005). Although not yet completely understood, biological variation in the absence of genetic or environmental cues ultimately depends on stochastic transitions and interactions of the elements that contribute to biological processes (Burga & Lehner 2012; Feinberg & Irizarry 2010; Kaern et al. 2005; Kilfoil et al. 2009; Lehner 2013; Raj & van Oudenaarden 2008; Raser & O'Shea 2005). One of the biological processes known to vary even in the absence of genetic or environmental cues is gene expression (Burga & Lehner 2012; Feinberg & Irizarry 2010; Kaern et al. 2005; Kilfoil et al. 2009; Lehner 2013; Raj & van Oudenaarden 2008; Raser & O'Shea 2005). Many of the multiple steps encompassed by gene expression depend on very small numbers for some of the intervening elements, making gene expression particularly susceptible to the intrinsic stochasticity of these elements transitions and interactions (Burga & Lehner 2012; Feinberg & Irizarry 2010; Kaern et al. 2005; Kilfoil et al. 2009; Lehner 2013; Raj & van Oudenaarden 2008; Raser & O'Shea 2005).

Biological variation ultimately caused by stochastic events at the molecular level can be an important evolutionary driving force (Burga & Lehner 2012; Feinberg & Irizarry 2010; Kaern et al. 2005; Kilfoil et al. 2009; Lehner 2013; Raj & van Oudenaarden 2008; Raser & O'Shea 2005). On one side, stochastic biological variation represents on its own a cheap phenotypic diversification that might permit populations coping with variable environments (Burga & Lehner 2012; Feinberg & Irizarry 2010; Kaern et al. 2005; Kilfoil et al. 2009; Lehner 2013; Raj & van Oudenaarden 2008; Raser & O'Shea 2005). On the other side, stochastic biological variation might act as genetic capacitor (Chalancon et al. 2012). Genetic capacitance refers to the accumulation of genetic variation with no
phenotypic effect, i.e., cryptic genetic variation, and its release upon capacitance attenuation (Masel 2013; Masel & Trotter 2010; Paaby & Rockman 2014). Genetic variation with phenotypes indistinguishable from the spectrum of stochastic phenotypes will be allowed to accumulate in a cryptic state until stochastic capacitance is somehow attenuated, and cryptic genetic variation becomes phenotypically relevant (Chalancon et al. 2012; Paaby & Rockman 2014). Thus, stochastic biological variation could serve natural populations short-term endurance to variable environments, while permitting the accumulation of cryptic genetic variation with great potential for their diversification. Although stochastic variation is an intrinsic property of biological systems, any factor that modulated stochastic biological variation would have direct and capacitance-driven indirect effects on short-term response to environmental changes and long-term diversification of natural populations. The identification of factors that resulted in the modulation of stochastic variation is, therefore, of broad interest.

The integration of literature on gene expression variation suggests sex can be an important modulator of stochastic biological variation in metazoan species. On one hand, as it would be expected from the contribution of any source of stochastic biological variation to genetic capacitance, stochastic variation in gene expression, or gene expression noise, has been shown to positively correlate with gene expression variation in response to conditional changes and divergence (Dong et al. 2011). Genetic variation for loci with noisier gene expression would be prone to accumulate in a cryptic state until capacitance is attenuated, whereas genetic variation for loci with less noisy gene expression would be more often phenotypically noticeable, and removed from populations if detrimental. On the other hand, in dipterans and mammals, male-biased expression responds to conditional changes and diverges faster than homogametic female-biased or unbiased gene expression, whereas in birds, it is the female-biased expression the one with faster conditional response and divergence (Assis et al. 2012; Ellegren &
Parsch 2007; Gallach et al. 2011; Jiang et al. 2010; Mank 2009; Mank et al. 2007; Meisel 2011; Parsch & Ellegren 2013; Singh & Artieri 2011; Wyman et al. 2010; Wyman et al. 2011). In dipterans and mammals, females are homogametic and males are heterogametic, and, in birds, females are heterogametic and males homogametic. Thus, it could be generalized that, in metazoan species, heterogametic sex-biased gene expression responds to conditional changes and diverges faster than homogametic sex-biased or unbiased gene expression (Assis et al. 2012; Ellegren & Parsch 2007; Gallach et al. 2011; Jiang et al. 2010; Mank 2009; Mank et al. 2007; Meisel 2011; Parsch & Ellegren 2013; Singh & Artieri 2011; Wyman et al. 2010; Wyman et al. 2011).

Considering the contribution of gene expression noise to genetic capacitance, the differences in sex-biased gene expression variation noticed in metazoan species could be explained if gene expression was generally noisier in the heterogametic sex (Figure 1). Phenotypes due to genetic variation in loci that are expressed in heterogametic individuals would often be indistinguishable from the broad spectrum of phenotypes dependent on the generally noisier gene expression of the heterogametic sex. Phenotypes due to genetic variation in loci that are expressed in homogametic individuals would be often distinguishable from the narrower spectrum of phenotypes dependent on the less noisy gene expression of the homogametic sex, and purged from populations if detrimental. Therefore, genetic variation for loci that are specifically expressed or overexpressed in the heterogametic sex would be prone to accumulate in a cryptic state, and become phenotypically relevant when capacitance is somehow attenuated by conditional changes in single and/or divergent populations.

Here, I present evidence for the existence of a sexual dimorphism in gene expression noise in metazoan species by comparing female and male putative genotype- and environment-independent variation for transcript abundance and other phenotypic traits in Drosophila melanogaster, and analyzing sex-biased
dispersal in metazoan species. These analyses are consistent with gene expression being generally noisier in the heterogametic sex, and point to the possibility that this sexual dimorphism in gene expression noise might be ultimately dependent on sex-specific chromosomes, i.e., $Y$ and $W$.

Mechanistic details on the effect of sex-specific chromosomes on gene expression noise are briefly discussed.

**MATERIALS AND METHODS**

The original source of datasets used in this article is: *D. melanogaster* transcript abundance in adult females and males (Diaz-Castillo et al. 2012), DGRP strains phenotypic data (Huang et al. 2014; Mackay et al. 2012), coordinates of regions with distinctive chromatin features in *D. melanogaster* somatic cells (Filion et al. 2010), and, sex-biased dispersal in metazoan species (Petit & Excoffier 2009). Coordinates of genes coding for transcripts in Diaz-Castillo and coworkers dataset were obtained from FlyBase (Diaz-Castillo et al. 2012; St Pierre et al. 2014). Microsoft® Excel® for Mac 2011 (Microsoft Corporation) was used to perform Monte Carlo-Wilcoxon matched-pairs signed-ranks tests, and Monte Carlo simulations. Prism 5 for Mac OS X (GraphPad Software, Inc) was used to analyze polynomial regressions between transcript abundance mean and CV for *D. melanogaster* strains and genotypes (Diaz-Castillo et al. 2012). Data processing and statistical analyses are described in detail on the Results and Discussion section, and Table footnotes.

**RESULTS AND DISCUSSION**

**Direct evidence for the existence of a sexual dimorphism in gene expression noise in *Drosophila***.

The existence of the hypothesized sexual dimorphism in gene expression noise should be easy to test. When directly comparing gene expression between sexes, putative genotype- and environment-independent variation should be generally larger in heterogametic than in homogametic individuals. In
2012, Diaz-Castillo and coworkers performed genome-wide microarray-based transcript abundance analyses of adult naive females and males in six *D. melanogaster* strains (Diaz-Castillo et al. 2012). These analyses resulted in abundance data of 16,637 transcripts for three biological replicates per strain and sex (36 samples: 6 strains x 2 sexes x 3 biological replicates). All analyzed samples were maintained under the same environmental conditions, and RNA was extracted simultaneously to minimize environmental and technical variation between biological replicates. With the only exception of the dosage of sex chromosomes, i.e., *X* and *Y*, genotype and environment were identical for biological replicates of each strain. Thus, any transcript abundance variation detected for biological replicates of a single strain might only be of sexual and/or stochastic nature, making it possible to test the potential existence of a generalized difference in gene expression noise between sexes.

Interestingly, five of the analyzed strains shared the same exact genotype with at least one another strain. INV1 and INV2 genotype is w<sup>1118</sup>/y<sup>+</sup> Y; In(2R) [P{FRT<sup>Cb-0236-3</sup>} P{FRT,w<sup>+</sup>}<sup>5-HA-1995</sup>]/SM6a, whereas SIM1, REV1 and REV2 genotype is w<sup>1118</sup>/y<sup>+</sup> Y; P{FRT,w<sup>-</sup>}<sup>Cb-0236-3</sup>, P{FRT, w<sup>-</sup>}<sup>5-HA-1995</sup>/SM6a. Since all samples were reared in the same conditions, transcript abundance data of strains with the same genotype could be pooled together permitting performing analyses with increasing analytical power (6 and 9 biological replicates per sex for INV and SIM/REV genotypes, respectively).

Since in *Drosophila*, females are homogametic and males are heterogametic, it would be expected, that in the dataset under study, measures for putative stochastic variation in transcript abundance were generally larger in males than in females. To test this prediction, I retrieved Diaz-Castillo and coworkers dataset from Gene Expression Omnibus database (GSE31120)(Diaz-Castillo et al. 2012). Transcript abundance data in this dataset had been already normalized to permit comparisons between samples (Diaz-Castillo et al. 2012). To compare transcript abundance variation putatively independent
of genetic and environmental cues, I calculated transcript abundance coefficients of variation for females and males in each strain and after pooling data of strains with the same genotype (CV\textsubscript{F} and CV\textsubscript{M}, respectively). The existence of a general trend for the difference of putative stochastic variation in transcript abundance between sexes was tested by performing Monte Carlo-Wilcoxon matched-pairs signed-ranks tests for each strain and genotype. For each transcript in the dataset, I subtracted CV\textsubscript{M} from CV\textsubscript{F}. CV differences were ranked upon the absolute value of their difference from lower to higher, and signs were assigned to each rank upon the sign of the difference between CV\textsubscript{F} and CV\textsubscript{M}. Wilcoxon sum of signed ranks (W) were calculated by adding signed ranks for all the transcripts in the dataset, for each strain and genotype. W is perfectly suited for the identification of general biases in collections of paired data, because it is sensitive both to the number or elements in the dataset with biased measures and the extent of such biases, eliminating the need to set arbitrary thresholds to infer trends of putative biologically significance. To estimate the significance of observed W, I recalculated W after randomly rearranging all CV data for each strain and genotype 10,000 times. Simulated W would represent the value W can adopt for a dataset with the same number of elements and the same value range without existing significant differences between sexes.

Considering W ultimately depends on the subtraction of CV\textsubscript{M} from CV\textsubscript{F} for each transcript, if gene expression was indeed generally noisier in males than in females, it would be expected that observed W were negative and commonly lower than simulated ones. In fact, except for strain INV2, all observed W were negative and lower than any simulated W (Table 1 and Figure 2). Strains INV1 and INV2 are genotypically identical and yet they differ with regard to the significance of the hypothesized sexual dimorphism in gene expression noise. It is worth noting that observed W are considerably more variable for single-strain analyses than for analyses of data from multiple strains sharing the same genotype. Since there is a difference in the extent of biological replication between single- and
multiple-strain analyses (Table 1), it is possible that non-significant observed $W$, as it is the case of
strain INV2, might be related with power limitations imposed by lower levels of biological replication.
In fact, non-significant trends exclusively associated to data encompassing lower levels of biological replication are also noticeable in forthcoming analyses, which stresses the importance of larger biological replication for the study of biological variation of putative stochastic nature.

The study of the association between transcript abundance mean and CV for each sex in the dataset under study also supports the existence of a sexual dimorphism in gene expression noise in *D. melanogaster*, and hints about its potential cause. It is commonly accepted that the variation in gene expression noise is mostly associated to changes in the mean level of gene expression (Bar-Even et al. 2006; Newman et al. 2006; Taniguchi et al. 2010). In those cases in which gene expression noise was modulated by factors other than the variation in the mean level of expression, it should be expected a weak association between measures for the mean level and stochastic variation in gene expression. In Diaz-Castillo and coworkers dataset, with the exception of INV1 and REV2, the coefficient of determination for the quadratic regression between transcript abundance mean and CV is considerably larger in females than in males (Table 2). These results suggest that the association between the mean level and stochastic variation in gene expression is weaker in males than in females. The male reduction in gene expression noise variation that could be explained solely through changes in the mean level of gene expression is consistent with the possibility that factors specific to males, heterogametic in *Drosophila*, promoted gene expression noise.

**Indirect evidence for the existence of a sexual dimorphism in gene expression noise in *Drosophila*.** It could be argued that the genome-wide difference in gene expression noise between *D. melanogaster* females and males here reported is an aberration of a particular genomic background, and/or the use of
a particular gene expression quantification methodology. For example, the experimental setup that resulted in Diaz-Castillo and coworkers dataset did not include technical replicates (Diaz-Castillo et al. 2012). Although it seems unlikely that consistent differences in the processing of female and male samples were to fully account for the generalized difference in transcript abundance variation detected in the dataset in question, it is currently not possible to remove from the measures of putative gene expression noise the variation due exclusively to technical issues. Other analyses using different D. melanogaster strains, other metazoan species, and/or other quantitative methods to measure gene expression are needed to confirm the existence and characterize further the hypothesized sexual dimorphism in gene expression noise in metazoan species.

Notwithstanding these limitations, the existence of a sexual dimorphism in gene expression noise in metazoan species should also be indirectly tested by studying sexual differences for the intrinsic variation of other phenotypic traits. Since the propagation of gene expression noise across biological levels would result in genotype- and environment-independent variation of other phenotypic traits that might permit populations coping with environmental changes (Burga & Lehner 2012; Feinberg & Irizarry 2010; Kaern et al. 2005; Kilfoil et al. 2009; Lehner 2013; Raj & van Oudenaarden 2008; Raser & O’Shea 2005), the existence of a sexual dimorphism in gene expression noise could be indirectly supported by clear sexual differences for the variation of phenotypic traits related with the interaction with the environment.

The D. melanogaster Genetic Reference Panel (DGRP) is a collection of over 200 inbred lines derived from a single D. melanogaster population in Raleigh (North Carolina, USA)(Huang et al. 2014; Mackay et al. 2012). The generation of these lines aimed to encompass standing genetic variation present in the founder population, so their intensive genotypic and phenotypic characterization helped
in addressing the relationship between the variation in genes and phenotypes. Among other assays, females and males from DGRP lines were used to measure their response to three environmental stressors, namely startle, starvation, and cold (Mackay et al. 2012). If the sexual dimorphism in gene expression noise detected in Diaz-Castillo and coworkers dataset was extensive to all *D. melanogaster* populations, and a generalized difference in gene expression noise between sexes resulted in a sexual difference in phenotypic variation, it would be expected that, regardless of genotypic and environmental differences, the response to these stressors were more variable in males than in females for DGRP lines.

To test this prediction, I performed Monte Carlo-Wilcoxon matched-pairs signed-ranks tests with DGRP startle response, starvation resistance and chill comma recovery data (http://dgrp2.gnets.ncsu.edu)(Mackay et al. 2012). For groups of females and males from the same strain assayed in the same conditions, I calculated $CV_F$ and $CV_M$. Although these three traits were originally measured individually, the assortment of assayed individuals was different for each trait. Data arrangement to best accommodate it in pairs of same genotype females and males assayed in the same conditions resulted in differences in the level of biological replication and the number of elements under study for each trait (Table 3). Startle response was assayed using groups of individuals of the same sex per vial in two different dates (Mackay et al. 2012). Since environmental conditions between dates might be slightly different, I considered assays in different days separately. I considered individuals within each group independent biological replicates, and calculated startle response $CV_F$ and $CV_M$ using measures for individuals of the same strain and sex assayed simultaneously. Starvation resistance was assayed using groups of individuals of the same sex per vial simultaneously (Mackay et al. 2012). I considered groups of individuals of the same strain and sex independent biological replicates, and averaged the starvation response measures for individuals in the same group. I
calculated starvation resistance $CV_F$ and $CV_M$ using averaged measures for each biological replicate of the same strain and sex. Chill comma recovery was assayed using same strain and sex groups of single individual vials assayed simultaneously (Mackay et al. 2012). I considered each group of individuals of the same strain and sex independent biological replicates, and averaged the chill comma recovery measures for individuals in the same group. I calculated chill comma recovery $CV_F$ and $CV_M$ using averaged measures for each biological replicate of the same strain and sex.

For each pair of $CV_F$ and $CV_M$, I subtracted the latter from the former. CV differences were ranked according to their absolute value from lower to higher, and signs were assigned to each rank upon the sign of the difference between $CV_F$ and $CV_M$. $W$ was calculated for each trait using the observed assortment of CV, and after randomly rearranging all CV data for each trait 10,000 times. As expected if gene expression noise-dependent phenotypic variation was generally larger in males than in females, observed $W$ were negative for all three traits (Table 3). Moreover, for startle response and starvation resistance, the fraction of simulated $W$ lower than observed $W$ was below the common threshold of significance ($P<0.05$)(Table 3). Although for chill comma recovery observed $W$ was not significantly different from simulated $W$ ($P>0.05$), it is worth noting that this is the trait here considered with the lowest level of biological replication, and, still, observed $W$ is lower than the majority of simulated $W$ (8,527 out of 10,000 simulated $W$ are larger than the observed one)(Table 3). Further analyses of sexual differences in the intrinsic variation of chill comma recovery with larger biological replication would be needed to decide if this trait shows a similar trend to the other two.

**Indirect evidence for the existence of a sexual dimorphism in gene expression noise in Metazoa.**

Sexual differences in gene expression noise with effect on phenotypic traits related with the interaction with the environment could influence each sex capacity to endure environmental changes, and,
therefore, their spatiotemporal dynamics. For instance, if larger gene expression noise resulted in a phenotypic variation that facilitated enduring environmental changes, it would be expected that noisier heterogametic individuals were able to better survive drastic environmental changes or occupy broader and/or more diverse geographies than less noisy homogametic individuals.

The movement of individuals away from their birth place or dispersal is constrained both by environmental conditions and factors intrinsic to dispersing individuals (Clobert et al. 2012). It is known that dispersal is sexually dimorphic in many species (Clobert et al. 2012; Dobson 2013; Greenwood 1980). For example, in mammals males tend to disperse more than females, whereas in birds females are the ones dispersing more (Clobert et al. 2012; Dobson 2013; Greenwood 1980). Since mammal males and bird females are heterogametic, their ability to disperse more than their homogametic relatives would fit with an enhanced ability to endure environmental changes possibly granted by a generalized larger gene expression noise in heterogametic individuals. The study of karyotypic constraints to sex-biased dispersal lends indirect support to the existence of the sexual dimorphism in gene expression noise in metazoan species promoted by heterogametic-specific factors.

In 2010, Petit and Excoffier studied the contribution of different factors to interspecific gene flow, using a dataset that included 36 widely distributed metazoan species for which sex-biased dispersal trends and chromosome systems were known (Petit & Excoffier 2009). In the dataset in question, birds, insects, and mammals were represented by 11, 11, and 14 cases, respectively (Supplemental Table S1). To confirm sex-biased dispersal could indeed reflect a generalized larger gene expression noise in heterogametic metazoans, I calculated the fraction of cases within Petit and Excoffier dataset where the heterogametic sex was the one dispersing the most, i.e., the fraction of heterogamety-biased dispersal. To test the significance of observed heterogamety-biased dispersal, I also calculated the fraction of
heterogamety-biased dispersal after randomly rearranging chromosome system tags 10,000 times. As
expected if metazoan heterogametic individuals tended to disperse more than homogametic
individuals, the observed fraction of heterogamety-biased dispersal is significantly larger than expected
by chance (Table 4).

Interestingly, Petit and Excoffier dataset includes five cases in which species commonly or completely
lack sex-specific chromosomes, $W/Y$ chromosomes hereinafter (Supplemental Table S1)(Ardila-Garcia
& Gregory 2009; Kandul et al. 2007; Morgan-Richards 1997; Narita et al. 2011; Petit & Excoffier
2009; Spence 1990). When the fraction of heterogamety-biased dispersal was calculated separately for
species where heterogametic individuals carried or lack $W/Y$ chromosomes, heterogametic individuals
significantly disperse the most only if they carried $W/Y$ chromosomes (Table 4). These results suggest
the presence of $W/Y$ chromosomes is an even better predictor for dispersal than heterogamety is.

Considering I argued differences in sex-biased dispersal could be consequence of sexual differences for
the variation in phenotypic traits that influence the interaction with the environment, ultimately caused
by a generalized difference in gene expression noise between sexes, $W/Y$ chromosomes-driven
dispersal could hint about a potential connection between $W/Y$ chromosomes and an increase in
putative gene expression noise-based phenotypic variation. This inference agrees with the possibility
that in $D. melanogaster$ gene expression noise is promoted by male-specific factors suggested by the
study of Diaz-Castillo and coworkers dataset. The connection between $W/Y$ chromosomes and gene
expression noise requires further consideration.

The sexual dimorphism in gene expression noise might depend on $W/Y$ chromosomes acting as
genomic tuning knobs. In $D. melanogaster$, the $Y$ chromosome has been shown to have a modulator
effect on phenotypic variation of stochastic nature, and gene expression variation across the genome.
On one side, *D. melanogaster* Y chromosome acts as a suppressor of the phenomenon known as position effect variegation (PEV), namely the stochastic inactivation of genes when they relocate into or juxtaposed to regions with highly compacted chromatin, i.e., heterochromatin (Elgin & Reuter 2013; Gowen & Gay 1934). On the other side, variation in *D. melanogaster* Y chromosome has been shown to have an indirect effect on the expression of multiple genes across the genome, a phenomenon referred to as Y-linked regulatory variation (YRV) (Lemos et al. 2010; Paredes et al. 2011; Sackton & Hartl 2013; Zhou et al. 2012). Although YRV affects genes widespread in the genome, it tends to be more accentuated for loci located in environments where gene expression is actively repressed, i.e., heterochromatin and nuclear periphery (Sackton & Hartl 2013). Since *D. melanogaster* Y chromosome indirect effects seem to be heterochromatin-centric, and *D. melanogaster* Y chromosome encodes for very few genes with functionalities not obviously connected with the regulation of gene expression, it has been argued that *D. melanogaster* Y chromosome indirect effects can depend on it acting as a sink for heterochromatin-forming elements (Berloco et al. 2014; Sackton & Hartl 2013; Zuckerkandl 1974). The highly heterochromatic *D. melanogaster* Y chromosome requires such a large amount of elements needed for heterochromatin formation that these would become depleted in heterochromatic loci located in other chromosomes, i.e., non-Y heterochromatic loci. Y chromosome-mediated depletion in heterochromatin-forming elements in non-Y loci will affect their level of chromatin compaction, and, therefore, their accessibility for the transcription machinery. Since some of the elements required for heterochromatin formation play also roles in non-heterochromatic loci (Cryderman et al. 2005; Fanti et al. 2008), *D. melanogaster* Y chromosome sink effect would be heterochromatin-centric, but not heterochromatin-exclusive.

*W/Y* chromosomes have originated from autosomes several independent times, and yet, they all seem to proceed through a progressive loss of coding elements and enrichment in repetitive DNA (Bachtrog
Except for chromosomes in early stages of this process, \( W/Y \) chromosomes tend to be the largest repositories in repetitive DNA in the genome, making of them the largest heterochromatic bodies in the nucleus (Ellegren 2011; Mank 2012; O'Meally et al. 2010). To test if the putative dependence on \( W/Y \) chromosomes of the sexual dimorphism in gene expression noise here documented was based on these chromosomes acting as a sink for heterochromatic-forming elements, I used again Diaz-Castillo and coworkers dataset (Diaz-Castillo et al. 2012).

In 2010, Filion and coworkers used genome-wide binding patterns of 53 chromatin elements to define five components in the \( D. \ melanogaster \) genome, which they symbolized with different colors (Filion et al. 2010). Four of these components are related with stable or transitory gene repression. BLUE and GREEN represent known heterochromatin repositories. BLACK and RED are enriched in loci with tissue-restricted gene expression and lamin-binding targets, suggesting they might be located towards the repressive environment of the nuclear periphery. YELLOW is the only component associated with broadly expressed loci. Transcripts in the Diaz-Castillo and coworkers dataset were assigned colors if the genomic region where they are encoded according to \( D. \ melanogaster \) genome annotation were spanned in their entirety by a single-color tract according to Filion and coworkers (Filion et al. 2010; St Pierre et al. 2014). As a measure of sexual dimorphism in gene expression noise, I calculated \( W \) for transcripts in each color per strain and genotype, following the afore explained process. The significance of the sexual dimorphism in gene expression noise for loci in each color was evaluated by recalculating \( W \) after randomly rearranging 10,000 times all CV for each strain and genotype.

A common pattern to strains and genotypes in the dataset is that observed \( W \) are negative and considerably smaller for YELLOW transcripts than for transcripts of any other color (Figure 3, and
Table 5). Such pattern suggests male-biased gene expression noise is less extreme for loci within repressive areas of the genome, which is consistent with what it would be expected if *D. melanogaster* Y chromosome sink effect altered chromatin compaction across the genome and, consequently, gene expression noise. Since in males much of the material required for heterochromatin formation would be diverted towards the Y chromosome, non-Y heterochromatic loci would be expected to be more compacted in females than in males (Figure 4). Higher chromatin compaction structures have been associated with an increase in gene expression noise due to slower gene expression dynamics (Kaern et al. 2005; Raj & van Oudenaarden 2008; Raser & O'Shea 2005). Thus, non-Y heterochromatic loci would be expected to be considerably noisy in females, and, therefore, show less extreme sexual dimorphism in gene expression noise.

The analysis of the sexual dimorphism in gene expression noise for genes in different compartments also showed that, with the exception of INV2 and REV2, observed W were negative and smaller than any simulated W for each chromatin color (Table 5). These results suggest gene expression is generally noisier in males than in females regardless of chromatin compaction and subnuclear localization. Thus, although the deployment of heterochromatin-forming elements in the presence/absence of *W/Y* chromosomes might cause differences in chromatin compaction and gene expression noise across genomes, on its own, a heterochromatin sink effect common to all *W/Y* chromosomes could not fully explain the existence of the putative sexual dimorphism in gene expression noise in metazoan species.

In 1997, David King and coworkers proposed the concept “genetic tuning knobs” to refer to the indirect effect the variation in repetitive DNA has on genetic elements they belong to or are intimately associated with (King et al. 1997). Small, reversible, and very frequent variation in repetitive DNA copy number would have a fine-tuning modifier effect for coding units and/or non-coding motives with
regulatory attributes (Gemayel et al. 2010; Kashi & King 2006; King et al. 1997). If repetitive DNA motives in $W/Y$ chromosome were susceptible to frequent variation in copy number, it could be expected that the fraction of repetitive DNA in $W/Y$ chromosomes was a very variable trait. Since the heterochromatic nature of $W/Y$ chromosomes depends on their content in repetitive DNA, it could be inferred that $W/Y$ chromosomes with slightly different fractions of repetitive DNA would withdraw slightly different amounts from the limiting pool of heterochromatin-forming elements (Figure 4). The variation in the amount of heterochromatin-forming elements deployed in $W/Y$ chromosomes with slightly different fractions of repetitive DNA would be translated in a variation in chromatin compaction in genes across the genome, and, therefore in their accessibility to the transcription machinery (Figure 4). Thus, a small variation in the fraction of repetitive DNA in $W/Y$ chromosomes could ultimately cause an increase in gene expression noise across the genome in heterogametic individuals. In other words, the sexual dimorphism in gene expression noise in metazoan species could be consequence of $W/Y$ chromosomes acting as tuning knobs at a genomic scale. Following with the musical theme, $W/Y$ chromosomes could be compared to the pulley-and-lever system of single-stringed whamolas. Handling of a whamola lever modifies the tension of the string causing a variation in noise when played. Similarly, variation in $W/Y$ chromosomes repetitive DNA would modify chromatin compaction across the genome, and, with it, the accessibility for the transcription machinery.

The characterization of genome regions enriched in repetitive DNA is one of the unresolved challenges in genomics (Chain et al. 2009; Treangen & Salzberg 2012). In fact, genomic regions enriched in repetitive DNA, such as $W/Y$ chromosomes, are remarkably misrepresented in genome assemblies. The confirmation of $W/Y$ chromosomes role as genomic tuning knobs, and its importance for the sexual dimorphism in gene expression noise here documented are currently hard to address empirically. Despite current technical limitations, in *D. melanogaster* indirect evidence exists consistent with the
possibility that the \textit{Y} chromosome might act as a genomic tuning knob. On one side, large and subtler induced changes in the fraction of repetitive DNA in \textit{D. melanogaster Y} chromosome have effects similar to PEV modulation and YRV (Berloco et al. 2014; Dimitri & Pisano 1989; Paredes et al. 2011).

Also, the fraction of repetitive DNA in \textit{W/Y} chromosomes has been shown to be very variable between closely related species and natural populations of a single species (Halfer 1981; Hughes & Rozen 2012; Lyckegaard & Clark 1989; Lyckegaard & Clark 1991; Nova et al. 2002; Paredes et al. 2011; Repping et al. 2006; Sahara et al. 2012; Singh et al. 1980). In \textit{D. melanogaster}, repetitive DNA derivatives suggested to mediate its copy number plasticity have been detected for the \textit{Y} chromosome in individuals of a single strain (Cohen et al. 2005; Cohen & Segal 2009), consistent with the possibility that \textit{D. melanogaster Y} chromosome repetitive DNA variation can be spontaneous and frequent. Thus, at least for \textit{D. melanogaster Y} chromosome, independent lines of evidence exist consistent with a frequent variation in its repetitive DNA and for this type of variation to have an effect on gene expression across the genome, as it would be expected if this chromosome acted as a genomic tuning knob. Further research is required to confirm \textit{W/Y} chromosomes acting as genomic tuning knobs caused a sexual dimorphism in gene expression noise in \textit{D. melanogaster} and other metazoan species.

The sexual dimorphism in gene expression noise might depend on nuclear dynamics in gametogenesis and early embryogenesis. The \textit{W/Y} chromosomes genomic tuning knob-sink effect model to explain sexual differences in gene expression noise in metazoan species is contingent not only upon the spontaneous and frequent variation in \textit{W/Y} chromosomes repetitive DNA, but also upon heterochromatin-forming elements being in roughly the same limiting amounts in nuclei of both sexes (Figure 4). There is only one moment common to all metazoans in which the pool of heterochromatin-forming elements might be in comparable limiting amounts regardless of chromosome composition. From oocyte fertilization until the zygotic genome gets fully activated, all new chromatin is formed at
the expense of maternally-deposited material (Banaszynski et al. 2010; Baroux et al. 2008; Tadros & Lipshitz 2009). Furthermore, sperm chromosomes enter the oocyte in a state of extreme compaction thanks to the almost complete substitution of histones by protamine-like proteins along spermatogenesis (Banaszynski et al. 2010; Tadros & Lipshitz 2009). During the transformation of sperm nuclei into paternal pronuclei, protamine-like proteins are substituted by histones, a process so intensive that will take a very large amount of the already limiting maternally-deposited material (Banaszynski et al. 2010; Tadros & Lipshitz 2009). In this context, the deployment of heterochromatin-forming elements during the first zygotic nuclear divisions might not run the same course in the presence or absence of \(W/Y\) chromosomes, or when the amount of repetitive DNA in these chromosomes is different (Figure 4). It could be speculated that the sexual dimorphism in gene expression noise in metazoan species ultimately depended on \(W/Y\) chromosomes acting as genomic tuning knobs-sinks very early in embryogenesis.

Interestingly, in *Drosophila* and Lepidoptera, evidence exists for the preferential use of DNA break repair strategies that can cause repetitive DNA variation in the germ line of heterogametic individuals (Cohen & Segal 2009; Díaz-Castillo 2013; Diaz-Castillo & Ranz 2012; Lieber 2010; Peng & Karpen 2007; Preston et al. 2006; Suzuki et al. 2009). Also, in *Drosophila*, silencing patterns of PEV reporters when maternally- or paternally-inherited are consistent with heterochromatin in later stages of development being dependent on the assortment of heterochromatic-forming elements very early in embryogenesis (Golic et al. 1998; Maggert & Golic 2002). Both the intrinsic spontaneous variation in repetitive DNA for \(W/Y\) chromosomes occurred along gametogenesis, and the preservation along development of early embryogenesis chromatin compaction genomic patterns are required for \(W/Y\) chromosomes acting as genomic tuning knobs-sinks in very early embryogenesis to be the base for the sexual dimorphism in gene expression noise in metazoan species. Further research is required to
confirm that nuclear dynamics along gametogenesis and early embryogenesis can indeed have such effect on the intrinsic variability in gene expression.

CONCLUSIONS

In the present article, I hypothesized gene expression might be generally noisier for heterogametic than for homogametic individuals in metazoan species. I present direct evidence of the hypothesized sexual dimorphism by studying a *D. melanogaster* dataset where transcript variation might be mostly of stochastic and sexual nature. Also, taking into consideration the direct and capacitance-mediated indirect contribution of gene expression noise to phenotypic variation, I predicted phenotypic traits related with the interaction with the environment should be more variable for heterogametic than homogametic individuals. Evidence of the putative gene expression noise-mediated dimorphism in phenotypic traits were found studying genotype- and environment-independent variation of the response to three stressors in *D. melanogaster*, and sex-biased dispersal in metazoan species. Further analyses permit speculating the sexual dimorphism in gene expression noise might be dependent on sex-specific chromosomes acting as genomic tuning knobs very early in embryogenesis. The intrinsically frequent variation in repetitive DNA for *W/Y* chromosomes might cause variation in chromatin compaction across the genome, which ultimately is transformed in an increase in gene expression noise. The genomic tuning knob-sink model for the origin of the sexual dimorphism in gene expression noise, which subsequently contributes to sexual differences for the variation of other phenotypic traits, illustrates the difficulty for the understanding of the connection between genotypes and phenotypes. *W/Y* chromosomes genomic tuning knob-sink effect could be described as epigenetic, i.e., it depends on chromatin compaction variation, but also as stochastic and genetic, i.e., it depends on spontaneous variation in repetitive DNA. Notwithstanding semantic tribulations, the possibility that chromosomes with very limited coding potential, i.e., *W/Y* chromosomes, can have genome-wide
effects such as PEV modifier, YRV or gene expression noise modulation talks about the potential importance nuclear dynamics might have for the phenotypic expression of genotypes.

ACKNOWLEDGEMENTS

The author wants to express his deepest gratitude to Raquel Chamorro-García for valuable comments during the preparation of this article. Also, the author thanks Richard Jorgensen, Alexander de Luna, and one more anonymous reviewer for their constructive criticism.

REFERENCES

Ardila-Garcia AM, and Gregory TR. 2009. An exploration of genome size diversity in dragonflies and damselflies (Insecta: Odonata). *Journal of Zoology* 278:163-173.

Assis R, Zhou Q, and Bachtro D. 2012. Sex-biased transcriptome evolution in Drosophila. *Genome Biol Evol* 4:1189-1200.

Bachtrog D. 2013. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nat Rev Genet* 14:113-124.

Banaszynski LA, Allis CD, and Lewis PW. 2010. Histone variants in metazoan development. *Dev Cell* 19:662-674.

Bar-Even A, Paulsson J, Maheshri N, Carmi M, O'Shea E, Pilpel Y, and Barkai N. 2006. Noise in protein expression scales with natural protein abundance. *Nat Genet* 38:636-643.

Baroux C, Autran D, Gillmor CS, Grimanelli D, and Grossniklaus U. 2008. The maternal to zygotic transition in animals and plants. *Cold Spring Harb Symp Quant Biol* 73:89-100.

Berloco M, Palumbo G, Piacentini L, Pimpinelli S, and Fanti L. 2014. Position effect variegation and viability are both sensitive to dosage of constitutive heterochromatin in Drosophila. *G3 (Bethesda)* 4:1709-1716.

Burga A, and Lehner B. 2012. Beyond genotype to phenotype: why the phenotype of an individual cannot always be predicted from their genome sequence and the environment that they experience. *FEBS J* 279:3765-3775.

Chain PS, Graffham DV, Fulton RS, Fitzgerald MG, Hostetler J, Muzny D, Ali J, Birren B, Bruce DC, Buhay C, Cole JR, Ding Y, Dugan S, Field D, Garrity GM, Gibbs R, Graves T, Han CS, Harrison SH, Highlander S, Hugenholtz P, Khouri HM, Kodira CD, Kolker E, Kyripides NC, Lang D, Lapidus A, Malfatti SA, Markowitz V, Metha T, Nelson KE, Parkhill J, Pitluck S, Qin X, Read TD, Schmutz J, Sozhamannan S, Sterk P, Strausberg RL, Sutton G, Thomson NR, Tiedje JM, Weinstock G, Wollam A, Genomic Standards Consortium Human Microbiome Project Jumpstart C, and Detter JC. 2009. Genomics. Genome project standards in a new era of sequencing. *Science* 326:236-237.
Chalancon G, Ravarani CN, Balaji S, Martinez-Arias A, Aravind L, Jothi R, and Babu MM. 2012. Interplay between gene expression noise and regulatory network architecture. *Trends Genet* 28:221-232.

Clobert J, Baguette M, Benton TG, and Bullock JM. 2012. *Dispersal ecology and evolution*: Oxford University Press.

Cohen S, Agmon N, Yacobi K, Mislovati M, and Segal D. 2005. Evidence for rolling circle replication of tandem genes in Drosophila. *Nucleic Acids Res* 33:4519-4526.

Cohen S, and Segal D. 2009. Extrachromosomal circular DNA in eukaryotes: possible involvement in the plasticity of tandem repeats. *Cytogenet Genome Res* 124:327-338.

Cryderman DE, Grade SK, Li Y, Fanti L, Pimpinelli S, and Wallrath LL. 2005. Role of Drosophila HP1 in euchromatic gene expression. *Dev Dyn* 232:767-774.

Díaz-Castillo C. 2013. Females and Males Contribute in Opposite Ways to the Evolution of Gene Order in Drosophila. *PLoS One* 8:e64491.

Díaz-Castillo C, and Ranz JM. 2012. Nuclear chromosome dynamics in the Drosophila male germ line contribute to the nonrandom genomic distribution of retrogenes. *Mol Biol Evol* 29:2105-2108.

Díaz-Castillo C, Xia XQ, and Ranz JM. 2012. Evaluation of the role of functional constraints on the integrity of an ultraconserved region in the genus Drosophila. *PLoS Genet* 8:e1002475.

Dimitri P, and Pisano C. 1989. Position effect variegation in Drosophila melanogaster: relationship between suppression effect and the amount of Y chromosome. *Genetics* 122:793-800.

Dobson FS. 2013. The enduring question of sex-biased dispersal: Paul J. Greenwood’s seminal contribution. *Animal Behaviour* 85:299-304.

Dong D, Shao X, Deng N, and Zhang Z. 2011. Gene expression variations are predictive for stochastic noise. *Nucleic Acids Res* 39:403-413.

Ellegren H. 2011. Sex-chromosome evolution: recent progress and the influence of male and female heterogamety. *Nat Rev Genet* 12:157-166.

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

Fanti L, Perrini B, Piacentini L, Berloco M, Marchetti E, Palumbo G, and Pimpinelli S. 2008. The trithorax group and Pc group proteins are differentially involved in heterochromatin formation in Drosophila. *Chromosoma* 117:25-39.

Feinberg AP, and Irizarry RA. 2010. Evolution in health and medicine Sackler colloquium: Stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. *Proc Natl Acad Sci U S A* 107 Suppl 1:1757-1764.

Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ, and van Steensel B. 2010. Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell* 143:212-224.

Gallach M, Domingues S, and Betran E. 2011. Gene duplication and the genome distribution of sex-biased genes. *Int J Evol Biol* 2011:989438.

Gelgel KG, Golic MM, and Pimpinelli S. 1998. Imprinted control of gene activity in Drosophila. *Curr Biol* 8:1273-1276.

Gowen JW, and Gay EH. 1934. Chromosome Constitution and Behavior in Eversporting and Mottling in Drosophila Melanogaster. *Genetics* 19:189-208.

Greenwood PJ. 1980. Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour* 28:1140-1162 %@ 0003-3472.

Halfer C. 1981. Interstrain heterochromatin polymorphisms in Drosophila melanogaster. *Chromosoma* 84:195-206.
Huang W, Massouras A, Inoue Y, Peiffer J, Ramia M, Tarone AM, Turlapati L, Zichner T, Zhu D, Lyman RF, Magwire MM, Blankenburg K, Carbone MA, Chang K, Ellis LL, Fernandez S, Han Y, Highnam G, Hjelmen CE, Jack JR, Javadi M, Jayaseelan J, Kalra D, Lee S, Lewis L, Munidasa M, Ongeri F, Patel S, Perales L, Perez A, Pu L, Rollmann SM, Ruth R, Saada N, Warner C, Williams A, Wu YQ, Yamamoto A, Zhang Y, Zhu Y, Anholt RR, Korbel JO, Mittelman D, Muzny DM, Gibbs RA, Barbadilla A, Johnston JS, Stone EA, Richards S, Deplancke B, and Mackay TF. 2014. Natural variation in genome architecture among 205 Drosophila melanogaster Genetic Reference Panel lines. Genome Res 24:1193-1208.

Hughes JF, and Rozen S. 2012. Genomics and genetics of human and primate y chromosomes. Annu Rev Genomics Hum Genet 13:83-108.

Jiang ZF, Croshaw DA, Wang Y, Hey J, and Machado CA. 2010. Enrichment of mRNA-like noncoding RNAs in the divergence of Drosophila males. Mol Biol Evol 28:1339-1348.

Kaern M, Elston TC, Blake WJ, and Collins JJ. 2005. Stochasticity in gene expression: from theories to phenotypes. Nat Rev Genet 6:451-464.

Kandul NP, Lukhtanov VA, and Pierce NE. 2007. Karyotypic diversity and speciation in Agrodiaetus butterflies. Evolution 61:546-559.

Kashi Y, and King DG. 2006. Simple sequence repeats as advantageous mutators in evolution. Trends Genet 22:253-259.

Kilfoil ML, Lasko P, and Abouheif E. 2009. Stochastic variation: from single cells to superorganisms. HFSP J 3:379-385.

King DG, Soller M, and Kashi Y. 1997. Evolutionary tuning knobs. Endeavour 21:36-40.

Lehner B. 2013. Genotype to phenotype: lessons from model organisms for human genetics. Nat Rev Genet 14:168-178.

Lemos B, Branco AT, and Hartl DL. 2010. Epigenetic effects of polymorphic Y chromosomes modulate chromatin components, immune response, and sexual conflict. Proc Natl Acad Sci USA 107:15826-15831.

Lieber MR. 2010. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem 79:181-211.

Lyckegaard EM, and Clark AG. 1989. Ribosomal DNA and Stellate gene copy number variation on the Y chromosome of Drosophila melanogaster. Proc Natl Acad Sci USA 86:1944-1948.

Lyckegaard EM, and Clark AG. 1991. Evolution of ribosomal RNA gene copy number on the sex chromosomes of Drosophila melanogaster. Mol Biol Evol 8:458-474.

Mackay TF, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu D, Casillas S, Han Y, Magwire MM, Cridland JM, Richardson MF, Anholt RR, Barron M, Bess C, Blankenburg KP, Carbone MA, Castellano D, Chaboub L, Duncan L, Harris Z, Javadi M, Jayaseelan JC, Jhangiani SN, Jordan KW, Lara F, Lawrence F, Lee SL, Librado P, Linheiro RS, Lyman RF, Mackey AJ, Munidasa M, Muzny DM, Nazareth L, Newsham I, Perales L, Pu LL, Qu C, Ramia M, Reid JG, Rollmann SM, Rozas J, Saada N, Turlapati L, Worley KC, Wu YQ, Yamamoto A, Zhu Y, Bergman CM, Thornton KR, Mittelman D, and Gibbs RA. 2012. The Drosophila melanogaster Genetic Reference Panel. Nature 482:173-178.

Maggert KA, and Golic KG. 2002. The Y chromosome of Drosophila melanogaster exhibits chromosome-wide imprinting. Genetics 162:1245-1258.

Mank JE. 2009. Sex chromosomes and the evolution of sexual dimorphism: lessons from the genome. Am Nat 173:141-150.

Mank JE. 2012. Small but mighty: the evolutionary dynamics of W and Y sex chromosomes. Chromosome Res 20:21-33.

Mank JE, Hultin-Rosenberg L, Axelsson E, and Ellegren H. 2007. Rapid evolution of female-biased, but not male-biased, genes expressed in the avian brain. Mol Biol Evol 24:2698-2706.

Masel J. 2013. Q&A: Evolutionary capacitance. BMC Biol 11:103.

Masel J, and Trotter MV. 2010. Robustness and evolvability. Trends Genet 26:406-414.
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.

Morgan-Richards M. 1997. Intraspecific karyotype variation is not concordant with allozyme variation in the Auckland tree weta of New Zealand, Hemideina thoracica (Orthoptera: Stenopelmatidae). *Biological Journal of the Linnean Society* 60:423-442.

Narita S, Kageyama D, Hiroki M, Sanpei T, Hashimoto S, Kamitoh T, and Kato Y. 2011. Wolbachia-induced feminisation newly found in Eurema hecabe, a sibling species of Eurema mandarina (Lepidoptera: Pieridae). *Ecological Entomology* 36:309-317.

Newman JR, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, and Weissman JS. 2006. Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* 441:840-846.

Nova P, Reutter BA, Rabova M, and Zima J. 2002. Sex-chromosome heterochromatin variation in the wood mouse, Apodemus sylvaticus. *Cytogenet Genome Res* 96:186-190.

O'Meally D, Patel HR, Stiglec R, Sarre SD, Georges A, Marshall Graves JA, and Ezaz T. 2010. Non-homologous sex chromosomes of birds and snakes share repetitive sequences. *Chromosome Res* 18:787-800.

Paaby AB, and Rockman MV. 2014. Cryptic genetic variation: evolution's hidden substrate. *Nat Rev Genet* 15:247-258.

Paredes S, Branco AT, Hartl DL, Maggert KA, and Lemos B. 2011. Ribosomal DNA deletions modulate genome-wide gene expression: "rDNA-sensitive" genes and natural variation. *PLoS Genet* 7:e1001376.

Parsch J, and Ellegren H. 2013. The evolutionary causes and consequences of sex-biased gene expression. *Nat Rev Genet* 14:83-87.

Peng JC, and Karpen GH. 2007. H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. *Nat Cell Biol* 9:25-35.

Petit RJ, and Excoffier L. 2009. Gene flow and species delimitation. *Trends Ecol Evol* 24:386-393.

Preston CR, Flores CC, and Engels WR. 2006. Differential usage of alternative pathways of double-strand break repair in Drosophila. *Genetics* 172:1055-1068.

Raj A, and van Oudenaarden A. 2008. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* 135:216-226.

Raser JM, and O'Shea EK. 2005. Noise in gene expression: origins, consequences, and control. *Science* 309:2010-2013.

Repping S, van Daalen SK, Brown LG, Korver CM, Lange J, Marszalek JD, Pyntikova T, van der Veen F, Skalsktsky H, Page DC, and Rozen S. 2006. High mutation rates have driven extensive structural polymorphism among human Y chromosomes. *Nat Genet* 38:463-467.

Sackton TB, and Hartl DL. 2013. Meta-analysis reveals that genes regulated by the Y chromosome in Drosophila melanogaster are preferentially localized to repressive chromatin. *Genome Biol Evol* 5:255-266.

Sahara K, Yoshido A, and Traut W. 2012. Sex chromosome evolution in moths and butterflies. *Chromosome Res* 20:83-94.

Singh L, Purdom IF, and Jones KW. 1980. Sex chromosome associated satellite DNA: evolution and conservation. *Chromosoma* 79:137-157.

Singh RS, and Artieri CG. 2011. Male sex drive and the maintenance of sex: evidence from Drosophila. *J Hered* 101 Suppl 1:S100-106.

Spence JR. 1990. Introgressive hybridization in Heteroptera: the example of Limnoporus Stal (Gerridae) species in western Canada. *Canadian Journal of Zoology* 68:1770-1782.

St Pierre SE, Ponting L, Stefancsik R, McQuilton P, and FlyBase C. 2014. FlyBase 102--advanced approaches to interrogating FlyBase. *Nucleic Acids Res* 42:D780-788.

Suzuki J, Yamaguchi K, Kajikawa M, Ichiyanagi K, Adachi N, Koyama H, Takeda S, and Okada N. 2009. Genetic evidence that the non-homologous end-joining repair pathway is involved in LINE retrotransposition. *PLoS Genet* 5:e1000461.
Tadros W, and Lipshitz HD. 2009. The maternal-to-zygotic transition: a play in two acts. *Development* 136:3033-3042.

Taniguchi Y, Choi PJ, Li GW, Chen H, Babu M, Hearn J, Emili A, and Xie XS. 2010. Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science* 329:533-538.

Treangen TJ, and Salzberg SL. 2012. Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat Rev Genet* 13:36-46.

Wyman MJ, Agrawal AF, and Rowe L. 2010. Condition-dependence of the sexually dimorphic transcriptome in Drosophila melanogaster. *Evolution* 64:1836-1848.

Wyman MJ, Cutter AD, and Rowe L. 2011. Gene duplication in the evolution of sexual dimorphism. *Evolution* 66:1556-1566.

Zhou J, Sackton TB, Martinsen L, Lemos B, Eickbush TH, and Hartl DL. 2012. Y chromosome mediates ribosomal DNA silencing and modulates the chromatin state in Drosophila. *Proc Natl Acad Sci U S A* 109:9941-9946.

Zuckerkandl E. 1974. A possible role of "inert" heterochromatin in cell differentiation. Action of and competition for "locking" molecules. *Biochimie* 56:937-954.
A sexual dimorphism in gene expression noise could explain differences in conditional response and divergence for sex-biased gene expression.

Charts symbolize gene expression dynamics for three transcripts in a population under different conditions or along time, under the assumption that gene expression is generally noisier in the heterogametic sex than in the homogametic sex. (A) Dynamics for a transcript that is overexpressed in the generally noisier heterogametic sex (Hm < Ht). (B) Dynamics for a transcript that is overexpressed in the generally less noisy homogametic sex (Hm > Ht). (C) Dynamics for a transcript equally expressed in both sexes (Hm ≈ Ht). Each line represents transcript abundance variation for a single individual in the population. Black and red lines represent variation in gene expression of stochastic or genetic nature, respectively. Grey areas represent gene expression levels with detrimental effects. In this case, both overexpression and underexpression beyond certain levels are detrimental. Individuals with detrimental expression stop contributing to the population. Noisier gene expression permits a better endurance for environmental changes, symbolized with narrower ranges for detrimental expression. Genetic variation in gene expression is prone to accumulate for transcripts that are overexpressed in the noisiest sex, because their phenotypes are often indistinguishable from the noise-driven phenotypic spectrum (A). Genetic variation in gene expression is less prone to accumulate for transcripts that are overexpressed in the less noisy sex or equally expressed in both sexes, because their phenotypes are often distinguishable from the noise-driven phenotypic spectrum, and removed from the population if detrimental (B and C).
A
Gene expression
Hm < Ht
Gene expression
Condition/Time

B
Gene expression
Hm > Ht
Gene expression
Condition/Time

C
Gene expression
Hm = Ht
Gene expression
Condition/Time

Stochastic variation in gene expression
Genetic variation in gene expression
Gene expression with detrimental effect
Direct evidence for the existence of a sexual dimorphism in gene expression noise in *D. melanogaster*.

Distribution of measures for per transcript gene expression noise sex bias before and after randomly rearranging observed gene expression noise measures for *D. melanogaster strains* with SIM/REV genotype according to Diaz-Castillo and coworkers dataset. Gene expression noise is measured as transcript abundance coefficient of variation in females and males (CV\(_F\) and CV\(_M\)). Sexual dimorphism in gene expression noise for the whole transcriptome before and after randomly rearranging the observed data was measured using W (see main text for further details). The negative skew of the distribution of observed data is consistent with gene expression being generally noisier in males than in females.
Sexual dimorphism in gene expression noise for genes in five different compartments of *D. melanogaster* genome.

Graph representing measures of sexual dimorphism in gene expression noise for loci with different chromatin structure and/or subnuclear localization according to Filion and coworkers (Filion et al. 2010), in *D. melanogaster* strains and genotypes represented in Diaz-Castillo and coworkers dataset (Diaz-Castillo et al. 2012). BLUE and GREEN components represent known heterochromatin repositories (Filion et al. 2010). BLACK and RED are enriched in loci with tissue-restricted gene expression and lamin-binding targets, suggesting they might be located towards the repressive environment of the nuclear periphery (Filion et al. 2010). YELLOW is the only component associated with broadly expressed loci (Filion et al. 2010). Sexual dimorphism in gene expression noise was measured using W (see main text for further details). The area of the graph representing W for BLUE, RED, and GREEN loci is amplified in the inner box.
Genomic tuning knob-sink effect of Y chromosomes.

Model for the assortment of heterochromatin-forming elements in homogametic and heterogametic nuclei, under the assumption that heterochromatic-forming elements are found in similar and limiting amounts in all nuclei. Font size is used to symbolize differences in repetitive DNA content of sexual chromosomes, i.e., X and Y, and autosomes (A). The fraction of heterochromatin-forming elements deployed in Y chromosomes can vary depending on their content in repetitive DNA. The assortment of heterochromatin-forming elements in non-Y loci would be very different if nuclei carry or lack Y chromosomes, i.e., XYAA or XXAA, and slightly different if Y chromosomes have different amounts of repetitive DNA. Slight differences in chromatin compaction across the genome ultimately based in small differences in Y chromosome repetitive DNA will cause an increase in gene expression noise across the genome, whereas higher level of chromatin compaction in non-Y heterochromatic loci in XXAA nuclei will be translated in less extreme male-biased gene expression noise in these loci (Figure 3, and main text).
Table 1 (on next page)

Monte Carlo-Wilcoxon matched-pairs signed-ranks tests for putative stochastic variation in transcript abundance in *D. melanogaster* females and males.
Table 1. Monte Carlo-Wilcoxon matched-pairs signed-ranks tests for putative stochastic variation in transcript abundance in *D. melanogaster* females and males.

| Strains | N replicates | N data pairs | Observed W | Simulated W [5th/95th percentiles] | $P_{upper}$/P$_{lower}$ |
|---------|--------------|--------------|------------|-------------------------------------|------------------------|
| REC     | 3            | 16,637       | -1.21 10^8 | -2.06 10^6/2.07 10^6               | 1.0000/0.0001          |
| INV1    | 3            | 16,637       | -1.24 10^8 | -2.04 10^6/2.03 10^6               | 1.0000/0.0001          |
| INV2    | 3            | 16,637       | -1.88 10^8 | -2.03 10^6/2.04 10^6               | 0.9354/0.0646          |
| SIM1    | 3            | 16,637       | -1.09 10^8 | -1.97 10^6/2.05 10^6               | 1.0000/0.0001          |
| REV1    | 3            | 16,637       | -5.96 10^7 | -2.05 10^6/2.08 10^6               | 1.0000/0.0001          |
| REV2    | 3            | 16,637       | -1.49 10^7 | -2.03 10^6/2.02 10^6               | 1.0000/0.0001          |
| INV     | 6            | 16,637       | -1.18 10^8 | -2.00 10^6/2.03 10^6               | 1.0000/0.0001          |
| SIM/REV | 9            | 16,637       | -1.05 10^8 | -2.00 10^6/2.04 10^6               | 1.0000/0.0001          |

Transcript abundance data was obtained from (Diaz-Castillo, et al. 2012). Observed Wilcoxon $W$ were obtained by subtracting $CV_M$ from $CV_F$ for each element in the dataset for each strain or genotype. CV differences were ranked from lower to higher according to their absolute value. Signs were assigned to ranks according to the sign of CV differences. Finally, $W$ were obtained by adding signed ranks for all the elements in the dataset for each strain or genotype. Simulated $W$ were obtained by repeating the same process after randomly rearranging all CV values for each strain or genotype 10,000 times. $P_{upper}$ and $P_{lower}$ values represent the fraction of random simulations with measures larger or equal, and lower or equal than the observed ones, respectively.
Table 2 (on next page)

Coefficients of determination ($R^2$) for quadratic regressions between transcript abundance mean and CV for *D. melanogaster* females and males.
Table 2. Coefficients of determination ($R^2$) for quadratic regressions between transcript abundance mean and CV for *D. melanogaster* females and males.

| Strains | Females | Males | Females / Males |
|---------|---------|-------|-----------------|
| REC     | 0.0496  | 0.0294| 1.6834          |
| INV1    | 0.0444  | 0.0750| 0.5923          |
| INV2    | 0.1218  | 0.0619| 1.9677          |
| SIM1    | 0.0918  | 0.0475| 1.9333          |
| REV1    | 0.0467  | 0.0123| 3.7853          |
| REV2    | 0.0639  | 0.0719| 0.8885          |
| INV     | 0.1563  | 0.0608| 2.5707          |
| SIM/REV | 0.1617  | 0.0281| 5.7585          |

Transcript abundance data was obtained from (Diaz-Castillo, et al. 2012).
Table 3 (on next page)

Monte Carlo-Wilcoxon matched-pairs signed-ranks tests for putative stochastic variation in phenotypic traits for *D. melanogaster* females and males.
Table 3. Monte Carlo-Wilcoxon matched-pairs signed-ranks tests for putative stochastic variation in phenotypic traits for *D. melanogaster* females and males.

| Trait                | N replicates | N data pairs | Observed $W$ | Simulated $W$ [5th/95th percentiles] | $P_{upper}$/$P_{lower}$ |
|----------------------|--------------|--------------|--------------|-------------------------------------|-------------------------|
| Startle response     | 18-40        | 405          | -13,773      | -7,788/7,639                        | 0.9986/0.0014           |
| Starvation resistance| 2-11         | 203          | -3,176       | -2,724/2,780                        | 0.9722/0.0280           |
| Chill coma recovery  | 2-4          | 174          | -1,395       | -2,201/2,151                        | 0.8527/0.1476           |

Phenotypic trait data was obtained from [http://dgrp2.gnets.ncsu.edu](http://dgrp2.gnets.ncsu.edu) (Mackay, et al. 2012). Observed Wilcoxon $W$ were obtained by subtracting $CV_M$ from $CV_F$ for each element in the dataset for each strain or genotype. CV differences were ranked from lower to higher according to their absolute value. Signs were assigned to ranks according to the sign of CV differences. Finally, $W$ were obtained by adding signed ranks for all the elements in the dataset for each strain or genotype. Simulated $W$ were obtained by repeating the same process after randomly rearranging all CV values for each strain or genotype 10,000 times. $P_{upper}$ and $P_{lower}$ values represent the fraction of random simulations with measures larger or equal, and lower or equal than the observed ones, respectively.
Table 4 (on next page)

Monte Carlo simulations for heterogamety-biased dispersal in metazoan species with and without sex-specific heterochromatic chromosomes.
Table 4. Monte Carlo simulations for heterogamety-biased dispersal in metazoan species with and without sex-specific heterochromatic chromosomes.

| Species | Observed fraction of heterogamety-biased dispersal | Simulated fraction of heterogamety-biased dispersal [5th-95th percentiles] | $P_{upper}/P_{lower}$ |
|---------|---------------------------------------------------|---------------------------------------------------------------------------|------------------------|
| All     | 0.81                                              | 0.36/0.64                                                                 | 0.0005/1.0000          |
| $XY+ZW$ | 0.90                                              | 0.36/0.65                                                                 | 0.0001/1.0000          |
| $XO+ZO$ | 0.20                                              | 0.20/0.80                                                                 | 0.9731/0.1788          |

Sex-biased dispersal and chromosome system data were obtained from (Spence 1990; Morgan-Richards 1997; Kandul, et al. 2007; Ardila-Garcia and Gregory 2009; Petit and Excoffier 2009; Narita, et al. 2011). $XO$ and $ZO$ represent species in which heterogametic individuals commonly or completely lack $W$ and $Y$. The fraction of heterogamety-biased dispersal represents the number of cases in the dataset in which heterogametic individuals tend to disperse more than homogametic individuals. Simulated fractions of heterogamety-biased dispersal were calculated after randomly rearranging chromosome system tags 10,000 times. $P_{upper}$ and $P_{lower}$ values represent the fraction of random simulations with measures larger or equal, and lower or equal than the observed ones, respectively.
Table 5 (on next page)

Monte Carlo-Wilcoxon matched-pairs signed-ranks tests for putative stochastic variation in transcript abundance in five different chromatin compartments in *D. melanogaster* females and males.
Table 5. Monte Carlo-Wilcoxon matched-pairs signed-ranks tests for putative stochastic variation in transcript abundance in five different chromatin compartments in *D. melanogaster* females and males.

| Strains | Chromatin color | N replicates | N data pairs | Observed $W$ | Simulated $W$ [5th/95th percentiles] | $P_{upper}$/P$_{lower}$ |
|---------|----------------|--------------|--------------|--------------|----------------------------------------|------------------------|
| REC     | YELLOW         | 3            | 4,907        | 1.03 10$^3$  | -3.30 10$^3$ /3.20 10$^3$              | 1.0000/0.0001         |
| REC     | BLACK          | 3            | 2,721        | -2.60 10$^3$ | 1.36 10$^3$ /1.34 10$^3$              | 1.0000/0.0001         |
| REC     | BLUE           | 3            | 2,721        | -2.60 10$^3$ | 2.06 10$^3$ /2.01 10$^3$              | 1.0000/0.0001         |
| REC     | RED            | 3            | 266          | 2.03 10$^4$  | -4.19 10$^4$ /4.06 10$^4$             | 1.0000/0.0001         |
| REC     | GREEN          | 3            | 147          | 1.00 10$^4$  | 1.10 10$^4$ /1.67 10$^3$              | 1.0000/0.0001         |
| INV1    | YELLOW         | 3            | 4,907        | 1.08 10$^3$  | -3.32 10$^4$ /3.22 10$^4$             | 1.0000/0.0001         |
| INV1    | BLACK          | 3            | 2,721        | -2.70 10$^3$ | 1.36 10$^4$ /1.36 10$^3$              | 1.0000/0.0001         |
| INV1    | BLUE           | 3            | 2,721        | -3.11 10$^4$ | 1.41 10$^4$ /4.17 10$^4$              | 1.0000/0.0001         |
| INV1    | RED            | 3            | 266          | 2.56 10$^3$  | -1.71 10$^3$ /1.68 10$^3$             | 1.0000/0.0001         |
| INV1    | GREEN          | 3            | 147          | -9.28 10$^3$ | -3.29 10$^3$ /3.27 10$^3$             | <0.0001/1.000         |
| INV2    | YELLOW         | 3            | 4,907        | 2.11 10$^4$  | -3.29 10$^3$ /3.27 10$^3$             | <0.0001/1.000         |
| INV2    | BLACK          | 3            | 2,721        | -9.20 10$^3$ | 1.37 10$^3$ /1.36 10$^3$              | 1.0000/0.0001         |
| INV2    | BLUE           | 3            | 2,721        | -4.60 10$^3$ | -2.11 10$^3$ /2.03 10$^3$             | 0.9999/0.0001         |
| INV2    | RED            | 3            | 266          | 4.19 10$^3$  | 0.1497/0.8505                         |
| INV2    | GREEN          | 3            | 147          | 2.17 10$^3$  | 0.0149/0.9851                         |
| SIM1    | YELLOW         | 3            | 4,907        | 9.21 10$^3$  | 3.33 10$^3$ /3.28 10$^3$              | 1.0000/0.0001         |
| SIM1    | BLACK          | 3            | 2,721        | 3.12 10$^4$  | 1.34 10$^3$ /1.37 10$^3$              | 1.0000/0.0001         |
| SIM1    | BLUE           | 3            | 2,721        | 2.48 10$^3$  | -2.02 10$^3$ /2.07 10$^3$             | 1.0000/0.0001         |
| SIM1    | RED            | 3            | 266          | 3.11 10$^4$  | 4.17 10$^3$ /4.12 10$^3$              | 1.0000/0.0001         |
| SIM1    | GREEN          | 3            | 147          | -9.26 10$^3$ | 2.10 10$^3$ /1.66 10$^3$              | 1.0000/0.0001         |
| REV1    | YELLOW         | 3            | 4,907        | -5.77 10$^3$ | 3.25 10$^3$ /3.21 10$^3$              | 1.0000/0.0001         |
| REV1    | BLACK          | 3            | 2,721        | -1.42 10$^4$ | 1.34 10$^3$ /1.37 10$^3$              | 1.0000/0.0001         |
| REV1    | BLUE           | 3            | 2,721        | -1.04 10$^3$ | -2.05 10$^3$ /2.03 10$^3$             | 1.0000/0.0001         |
| REV1    | RED            | 3            | 266          | 1.17 10$^4$  | 0.1497/0.8505                         |
| REV1    | GREEN          | 3            | 147          | 5.61 10$^3$  | -1.69 10$^3$ /1.71 10$^3$             | 1.0000/0.0001         |
| REV2    | YELLOW         | 3            | 4,907        | -4.94 10$^3$ | 3.27 10$^3$ /3.29 10$^3$              | 0.9930/0.0070         |
| REV2    | BLACK          | 3            | 2,721        | -8.79 10$^3$ | -1.34 10$^3$ /1.33 10$^3$             | 1.0000/0.0001         |
| REV2    | BLUE           | 3            | 2,721        | -8.35 10$^3$ | -2.06 10$^3$ /2.02 10$^3$             | 1.0000/0.0001         |
| REV2    | RED            | 3            | 266          | -9.38 10$^3$ | 0.9998/0.0002                         |
| REV2    | GREEN          | 3            | 147          | 3.15 10$^3$  | 0.0010/0.9990                         |
| INV     | YELLOW         | 6            | 4,907        | 1.00 10$^3$  | 3.29 10$^3$ /3.22 10$^3$              | 1.0000/0.0001         |
| INV     | BLACK          | 6            | 2,721        | 3.28 10$^3$  | 1.35 10$^3$ /1.38 10$^3$              | 1.0000/0.0001         |
| INV     | BLUE           | 6            | 2,721        | 2.63 10$^3$  | 2.06 10$^3$ /2.07 10$^3$              | 1.0000/0.0001         |
| INV     | RED            | 6            | 266          | 3.35 10$^3$  | 4.11 10$^3$ /4.09 10$^3$              | 1.0000/0.0001         |
|    |    |    |    |    |
|----|----|----|----|----|
| INV| GREEN | 6  | 147 | -8.39 $10^3$ | -1.68 $10^3$/1.70 $10^3$ | 1.0000/$<0.0001$ |
| SIM/REV| YELLOW | 9  | 4,907 | -9.13 $10^6$ | -3.21 $10^6$/3.30 $10^6$ | 1.0000/$<0.0001$ |
| SIM/REV| BLACK | 9  | 2,721 | -3.04 $10^8$ | -1.33 $10^8$/1.33 $10^8$ | 1.0000/$<0.0001$ |
| SIM/REV| BLUE | 9  | 775  | -2.52 $10^4$ | -2.07 $10^4$/2.07 $10^4$ | 1.0000/$<0.0001$ |
| SIM/REV| RED | 9  | 266  | -2.94 $10^4$ | -4.16 $10^4$/4.15 $10^4$ | 1.0000/$<0.0001$ |
| SIM/REV| GREEN | 9  | 147  | -6.51 $10^3$ | -1.69 $10^3$/1.73 $10^3$ | 1.0000/$<0.0001$ |

Transcript abundance data was obtained from (Diaz-Castillo, et al. 2012). Coordinates of chromatin components symbolized with colors were obtained from (Filion, et al. 2010). Observed Wilcoxon $W$ were obtained by subtracting $CV_M$ from $CV_F$ for each element in every chromatin compartment for each strain or genotype. CV differences were ranked from lower to higher according to their absolute value. Signs were assigned to ranks according to the sign of CV differences. Finally, $W$ were obtained by adding signed ranks for all the elements with the same color for each strain or genotype. Simulated $W$ were obtained by repeating the same process after randomly rearranging all CV values for each strain or genotype 10,000 times. $P_{upper}$ and $P_{lower}$ values represent the fraction of random simulations with measures larger or equal, and lower or equal than the observed ones, respectively.