

# Conserved Patterns of Sex Chromosome Dosage Compensation in the Lepidoptera (WZ/ZZ): Insights from a Moth Neo-Z Chromosome

Liuqi Gu<sup>1,\*</sup>, James R. Walters<sup>2</sup>, and Douglas C. Kipple<sup>1</sup>

<sup>1</sup>Department of Entomology, Cornell University, Geneva, NY

<sup>2</sup>Department of Ecology & Evolutionary Biology, The University of Kansas, Lawrence, KS

\*Corresponding author: E-mail: lg356@cornell.edu.

Accepted: March 1, 2017

Data deposition: This project has been deposited at GenBank under the accession BioProject PRJNA283415.

## Abstract

Where previously described, patterns of sex chromosome dosage compensation in the Lepidoptera (moths and butterflies) have several unusual characteristics. Other female-heterogametic (ZW/ZZ) species exhibit female Z-linked expression that is reduced compared with autosomal expression and male Z expression. In the Lepidoptera, however, Z expression typically appears balanced between sexes but overall reduced relative to autosomal expression, that is  $Z \approx ZZ < AA$ . This pattern is not easily reconciled with theoretical expectations for the evolution of sex chromosome dosage compensation. Moreover, conflicting results linger due to discrepancies in data analyses and tissues sampled among lepidopterans. To address these issues, we performed RNA-seq to analyze sex chromosome dosage compensation in the codling moth, *Cydia pomonella*, which is a species from the earliest diverging lepidopteran lineage yet examined for dosage compensation and has a neo-Z chromosome resulting from an ancient Z:autosome fusion. While supported by intraspecific analyses, the  $Z \approx ZZ < AA$  pattern was further evidenced by comparative study using autosomal orthologs of *C. pomonella* neo-Z genes in outgroup species. In contrast, dosage compensation appears to be absent in reproductive tissues. We thus argue that inclusion of reproductive tissues may explain the incongruence from a prior study on another moth species and that patterns of dosage compensation are likely conserved in the Lepidoptera. Notably, this pattern appears convergent with patterns in eutherian mammals ( $X \approx XX < AA$ ). Overall, our results contribute to the notion that the Lepidoptera present challenges both to classical theories regarding the evolution of sex chromosome dosage compensation and the emerging view of the association of dosage compensation with sexual heterogamety.

**Key words:** *Cydia pomonella*, Lepidoptera, sex chromosome dosage compensation, neo-Z chromosome, sex-biased gene expression.

## Introduction

In the animal kingdom, there are two predominant sex chromosome systems. In XX/XY species, females are the homogametic sex (XX) and males are the heterogametic sex (XY). This pattern is reversed in WZ/ZZ species, where females are heterogametic (WZ) and males are homogametic (ZZ). In both cases, however, sex chromosome dosage compensation is expected to evolve concomitantly with the sex chromosomes in order to offset a loss of gene dosage that accompanies the heterochromatic degradation of the Y or W chromosomes (Ohno et al. 1959; Ohno 1967; Charlesworth 1978; Rice 1984). As such, sex chromosome dosage compensation is

expected to yield two distinct but related patterns: 1) balanced sex-linked gene expression between the male and female ( $X \sim XX \mid Z \sim ZZ$  balance) and 2) equalized expression between the sex chromosome and the autosomes in the heterogametic sex ( $X \sim AA \mid Z \sim AA$  parity), on the assumption that autosomal expression is representative of the ancestral sex-linked expression.

Different organisms deploy distinct strategies to achieve these two aspects of dosage compensation. In XX/XY systems, the  $X \sim XX$  balance involves orchestrated chromosome-wide regulation (reviewed in Ferrari et al. 2014). In *Drosophila*, transcriptional up-regulation of the single X chromosome of XY

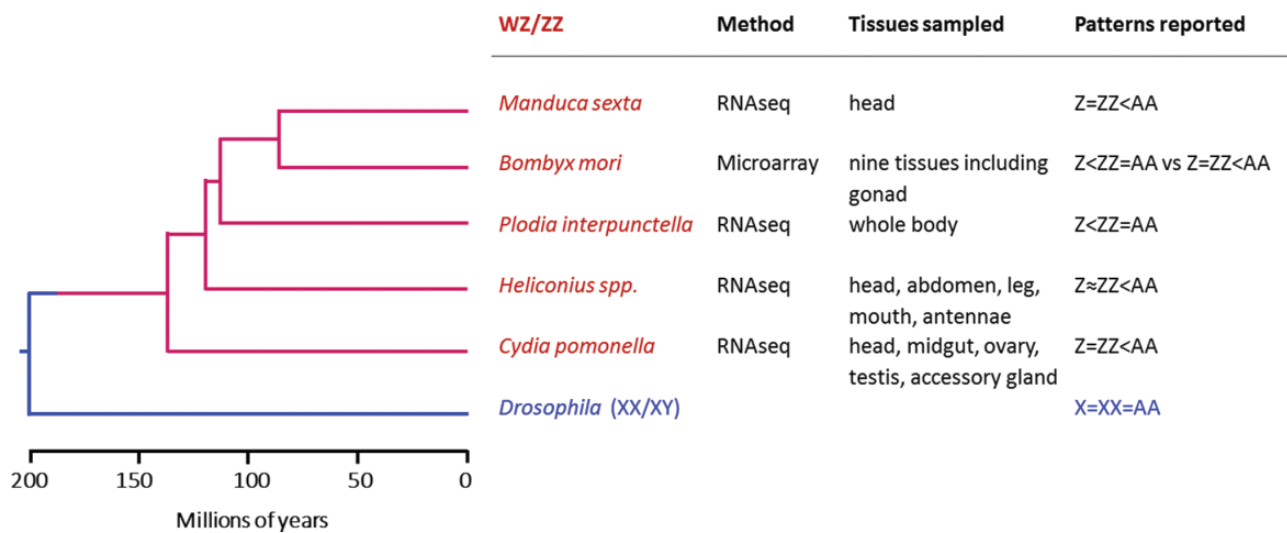
males equalizes its output to that from two X copies of XX females. This doubling of expression of X-linked genes in males also equalizes average expression with the autosomes (AA) (i.e.  $X \approx XX \approx AA$ , reviewed in Gelbart and Kuroda 2009; Laverty et al. 2010; Conrad and Akhtar 2012). In contrast, in both placental mammals and *Caenorhabditis* worms, X~XX balance is achieved by reducing X chromosome expression in the homogametic sex. Mammals inactivate one X chromosome (reviewed in Schulz and Heard 2013), while worms repress both X chromosomes by half (reviewed in Meyer 2010). Additionally, general 2-fold up-regulation of X-linked gene expression (global X~AA compensation), as first proposed by Ohno (Ohno 1967), is predicted to equalize the level of X-linked gene expression in both sexes to the level of diploid autosomal expression (i.e.  $X \approx XX \approx AA$ ). However, the latter notion is widely challenged by a growing body of evidence that favors the reduction in X chromosome gene expression in both mammals (reviewed in Pessia et al. 2014) and *C. elegans* (Albritton et al. 2014) (i.e.  $X \approx XX < AA$ ).

Patterns of dosage compensation in WZ/ZZ species generally have shown a pattern distinct from XX/XY systems. Female heterogametic species typically lack global Z chromosome compensation and therefore exhibit a gene dosage effect producing male bias in Z-linked gene expression. This  $Z < ZZ \approx AA$  pattern has been reported in several different taxa with independently evolved female heterogamety, including birds (Ellegren et al. 2007; Itoh et al. 2007, 2010; Wolf and Bryk 2011; Adolfsson and Ellegren 2013; Uebbing et al. 2013), snakes (Vicoso et al. 2013), *Cynoglossus semilaevis* (a flatfish) (Chen et al. 2014) and *Schistosoma mansoni* (a trematode parasite) (Vicoso and Bachtrog 2011). To date, the only exceptions to this pattern among female-heterogametic species come from the insect order of Lepidoptera (moths and butterflies), although results from the limited number of lepidopteran species examined thus far are not entirely consistent. As in other WZ/ZZ taxa, a pattern of  $Z < ZZ \approx AA$  was initially reported in the silk moth *Bombyx mori* (Zha et al. 2009) using microarrays, and subsequently in the Indian meal moth *Plodia interpunctella* (Harrison et al. 2012) using RNA-seq. These observations are consistent with an emerging view that female-heterogametic taxa lack global mechanisms for sex chromosome dosage compensation and thus exhibit a strong male bias in sex-linked gene expression (Graves and Disteche 2007; Mank 2009, 2013; Vicoso and Bachtrog 2009; Naurin et al. 2010; Wolf and Bryk 2011; Harrison et al. 2012). However, the analyses in *B. mori* was later challenged due to artifacts of microarray normalization and a failure to examine the female to male (F:M) autosomal gene expression ratios concurrently with the level of Z-linked expression (Walters and Hardcastle 2011). Re-analyses of the same data set revealed a  $Z \approx ZZ < AA$  pattern that mirrors the recent results in eutherian mammals (Julien et al. 2012; Lin et al. 2012). More recent RNA-seq studies in two additional lepidopterans, *Manduca sexta* (Smith et al. 2014) and

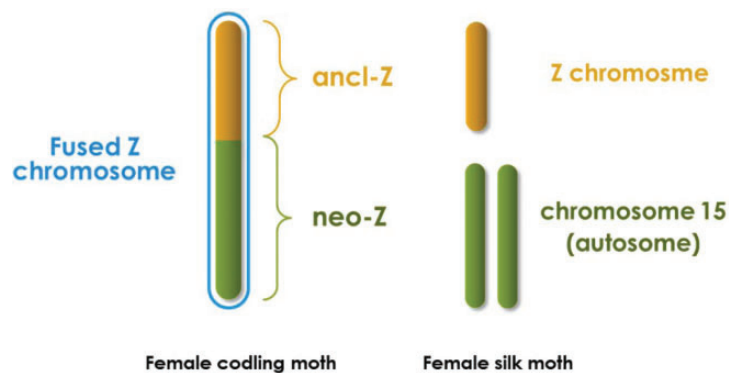
*Heliconius* butterflies (Walters et al. 2015), also revealed a pattern of both sexes having equally reduced Z-linked expression relative to autosomal expression (i.e.  $Z \approx ZZ < AA$ ).

Unfortunately, these lepidopteran studies are not easily compared because they differed in methodology (microarray vs. RNA-seq) and also in which tissues were sampled (fig. 1). Notably, in only one case were gonads analyzed separately from somatic tissues (*B. mori* in Walters and Hardcastle 2011), while two other cases used entire bodies or abdomen, potentially confounding patterns of dosage compensation in gonads and soma (Harrison et al. 2012; Walters et al. 2015). This is problematic because gonads should ideally be isolated in dosage compensation studies due to the widely observed germline-specific regulation of the sex chromosomes, such as the absence of dosage compensation (Kelly et al. 2002; Sugimoto and Abe 2007; Meiklejohn et al. 2011; Vicoso and Bachtrog 2015) and meiotic sex chromosome inactivation (Monesi 1965; Bean et al. 2004; Turner 2007; Schoenmakers et al. 2009; Guioli et al. 2012; Vibranovski 2014). In addition, genes with predominant expression in gonads are often disproportionately located on the sex chromosomes compared with the autosomes (Saifi and Chandra 1999; Lercher et al. 2003; Parisi et al. 2003; Khil et al. 2004; Reinke et al. 2004; Arunkumar et al. 2009). All these factors add layers of complexity to the assessment of dosage compensation and potentially bias the results when gonads and soma are analyzed together.

Beyond the technical artifacts and potential biases of including gonads, two more fundamental questions linger concerning dosage compensation in the Lepidoptera. First, to what extent do lepidopterans represent an exception to the otherwise broad observations that female-heterogametic taxa exhibit a strong dosage effect on the Z chromosome? Second, if the absence of a Z-chromosome dosage effect proves common among lepidopterans, does the mechanism of balancing expression between sexes involve reduced expression in the male or increased expression in the female (the latter is predicted by Ohno's model)? To refine this second question, consider that in the Lepidoptera, balanced expression between sexes seems to be associated with reduced Z-linked expression relative to autosomes. One explanation for this reduced Z expression could be that males are down-regulating gene expression from their diploid Z chromosomes. An alternative explanation could be that the proto-Z chromosome just happened to have unusually low average expression relative to other autosomes before degradation and divergence of the W chromosome. In this latter scenario, current observation of Z:A expression ratios distinctly  $< 1$  simply reflects the (unchanged) ancestral expression patterns and it is the females that up-regulate Z expression to fully mitigate gene dosage effects, a situation consistent with the canonical theory of dosage compensation. Discerning between these two scenarios is crucial for assessing whether dosage compensation in the



**FIG. 1.**—Summary of dosage compensation studies in the Lepidoptera. Phylogenetic relationships among insect species mentioned in this study is based on Marec et al. (2010) and Wahlberg et al. (2013).



**FIG. 2.**—Sex chromosomes comparison between *Cydia pomonella* and *Bombyx mori*. The “ancl-Z” and “neo-Z” chromosomal segments of *C. pomonella* are homologous to the Z chromosome and autosome 15 of *B. mori*, respectively. Z chromosomes are effectively hemizygous in lepidopteran females; W chromosomes are heterochromatic and are presumed to contain few or no functional genes in this order and are therefore not drawn.

Lepidoptera represents a distinct departure from theoretical expectations.

In order to address these questions, we generated RNA-seq data from several different tissues of adult female and male codling moth (*Cydia pomonella*). Importantly, we took care to isolate somatic from reproductive tissues, analyzing both, in order to directly address any inherent differences. *Cydia pomonella* represents a more basal phylogenetic group (Tortricidae) compared with the other lepidopterans previously examined for dosage compensation (fig. 1). Thus dosage analysis in this species distinctly extends the taxonomic sampling of dosage compensation performed in the Lepidoptera. Furthermore, *C. pomonella* has a large neo-Z chromosome that reflects a fusion between an autosome (homologous to *B. mori* chromosome 15) and the ancestral Z chromosome (fig. 2) (Nguyen et al. 2013). Both molecular and cytogenetic

evidence showed that the *C. pomonella* W chromosome is extensively heterochromatinized, a common characteristic of the lepidopteran W (Fukova et al. 2007; Nguyen et al. 2013; Sichova et al. 2013). Therefore, the autosomal progenitors of *C. pomonella* neo-Z genes were brought under the same sex-linked inheritance and accompanying sex-specific gene dosage as the ancestral-Z (ancl-Z) genes. This evolutionary partitioning of the *C. pomonella* Z chromosome provides a novel opportunity to examine the effect of sex-linkage on dosage compensation. In particular, genes on the neo-Z portion of *C. pomonella* can be compared directly to autosomal orthologs in other species. This yields a relatively direct assessment of the specific effects of Z-linkage on levels of gene expression and avoids assumptions about equal average expression across chromosomes or between sexes that are inherent in most assessments of dosage compensation (Julien

et al. 2012; Lin et al. 2012; Albritton et al. 2014; Nozawa et al. 2014). This neo-Z arrangement can thus help to discern between male down-regulation versus female up-regulation of the Z chromosome in the case that Z-linked expression is balanced between sexes.

## Materials and Methods

### Samples, RNASeq, and De Novo Assembly

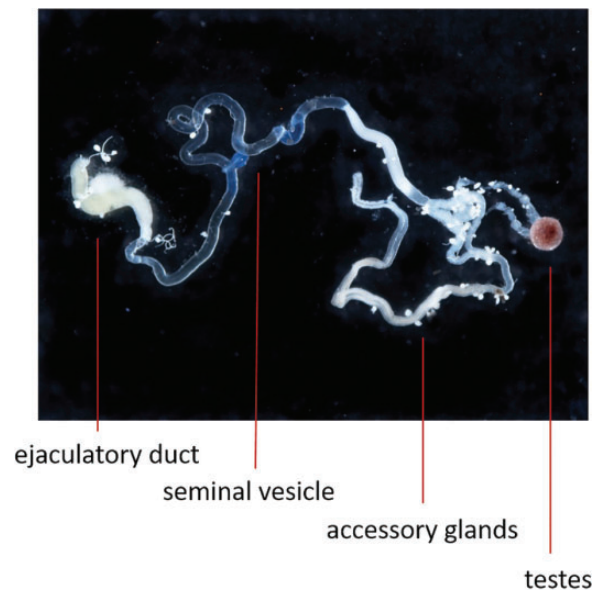
*Cydia pomonella* pupae were obtained from a commercial breeder (Benzon Research Inc., Carlisle, PA). The following seven tissue samples were dissected from young adult moths within 24 hours after emergence: head (female), head (male), midgut (female), midgut (male), ovary (female), testis (male) and accessory gland (male). The testis was isolated from the male reproductive tract, which is primarily comprised of accessory gland, along with seminal vesicles, ejaculatory duct and ejaculatory bulb (fig. 3). In this study, these somatic tissues of the male reproductive tract were dissected together, analyzed as a single tissue sample, and referred to as “accessory gland” for the purpose of simplicity.

Tissues dissected from five individuals were pooled into one sequencing sample, and two sequencing samples (replicates) were used for each tissue sample. RNA was separately extracted from each sequencing sample using the QIAGEN RNeasy Kit (QIAGEN, Foster City, CA), following manufacturer’s instruction and subsequently sent to Boyce Thompson Institute (Ithaca, NY) for the construction of bar-coded libraries using standard Illumina library preparation protocols (Meyer and Kircher 2010). The resulting 14 libraries were then pooled into three lanes and sequenced on the Illumina HiSeq2000 platform (150 bp paired-end) at the Cornell University Biotechnology Resource Center (Ithaca, NY).

The resulting 150-bp paired-end reads were first filtered and trimmed using Trimmomatic (V0.32) with default parameters (Bolger et al. 2014). The remaining reads from all 14 sequencing samples were pooled for de novo assembly on the Trinity platform with default parameters (Grabherr et al. 2011; Haas et al. 2013). The finished assembly contained all “Trinity components” (or contigs)  $\geq 200$  bp in length, which represented the *C. pomonella* transcriptomes from the seven adult tissues.

### Estimation of Expressional Intensity

Fragments per kilobase per million mappable reads (FPKM) were used to estimate gene expression intensity. Filtered read sets from each of the 14 sequencing samples were individually mapped back to the transcriptome assembly for FPKM calculation using RSEM (Li and Dewey 2011) contained in the Trinity package with default parameters. The resulting 14 FPKM sets represented the gene expression in the 7 tissues, each with two replicates. In order to facilitate cross-tissue comparison, FPKM values were further normalized by the Trimmed Mean of *M* values (TMM) method (Robinson and



**Fig. 3.**—Male reproductive tissues of *Cydia pomonella* dissected for RNAseq analyses. Testis were separated from the male reproductive tract. The latter was collected as a single tissue comprising the accessory gland (the predominant component), along with seminal vesicles, and the ejaculatory duct (excluding the ejaculatory bulb).

Oshlack 2010). Because the correlation of FPKM values between the two replicates was high for all the tissue samples (Pearson’s  $\rho$  ranging 0.84 ~ 0.99, mean = 0.93), average FPKMs were calculated and used in all the subsequent analyses.

### Chromosomal Assignment

*Cydia pomonella* 1:1 orthologous contigs to *B. mori* genes were identified using reciprocal best-hit approach with a BLAST algorithm e-value cut-off of  $1E-5$ . *Bombyx mori* gene sequences and chromosome map with scaffold information were downloaded from SilkDB v2.0 (<http://silkworm.genomics.org.cn>; last accessed March 8, 2017). The contigs mapped to *B. mori* chromosome 1 (Z) and chromosome 15 were identified as linked to ancl-Z and neo-Z segments, respectively, while those mapping to other chromosome locations were considered to be autosomal genes. Contigs that could not be mapped to a *B. mori* chromosome were excluded from subsequent analyses.

### Analysis of Dosage Compensation within Species

We assessed patterns of dosage compensation by contrasting the neo-Z or ancl-Z genes with autosomal genes in two different ways. First, to assess Z~ZZ balance, we calculated F:M ratios of FPKM values for all genes with FPKM > 0 in both sexes for the given tissues. Differences in average F:M ratios between chromosomal subsets were assessed via Mann–



Whitney *U* test, with a Bonferroni correction. Second, to assess Z~A parity, we examined the ratio of median expression for ancl-Z or neo-Z versus autosomal loci, with confidence intervals of these ratios generated via 1,000 bootstrapped replicates. All statistical analyses were performed using the R statistical analysis software (R Core Team 2014), with bootstrapping via the boot package (Davison and Hinkley 1997).

### Interspecies Comparative Transcriptome Analyses

RNAseq data from *M. sexta* and *H. melpomene* head samples was accessed from published studies (Smith et al. 2014) and (Walters et al. 2015). Both FPKM data sets were TMM-normalized within the species and averaged among the replicates. The orthologous pairs between *C. pomonella* and *M. sexta* and their chromosomal locations were identified by mutually mapping to the *B. mori* reference genome. For the *C. pomonella*~*H. melpomene* comparison, the *H. melpomene* reference genome (<http://www.butterflygenome.org/>; last accessed March 8, 2017) was used to identify 1:1 orthologs in *C. pomonella* and infer their chromosomal locations. *Heliconius melpomene* chromosome 11 corresponds to *B. mori* chromosome 15 and the two chromosomes share a high level of synteny (*Heliconius* Genome Consortium 2012).

We adopted a scaling procedure (Lin et al. 2012) to enable comparison of FPKM values between species. Specifically, all FPKM values from one species were linearly adjusted by a common factor in each sex so that the median FPKM values for autosomal genes were the same between the two species to be compared. To avoid infinite values when calculating ratios without changing the medians, all FPKM=0 values were converted to FPKM=0.01.

### Differential Expression Analysis

Differentially expressed genes were identified using EdgeR (Bioconductor) imbedded in the Trinity package (Robinson et al. 2010). EdgeR uses false discovery rate (FDR) to determine differentially expressed loci, which adjusts gene-specific *P* values for multiple tests (Storey and Tibshirani 2003). In order to reduce complexity, for head and midgut, reads from female and male samples were pooled to recalculate for nonsex-specific FPKM.

## Results

### Samples, RNA-Seq, De Novo Assembly, and Chromosomal Assignment of Genes

We sampled seven *C. pomonella* adult female (F) and male (M) tissues that represent different levels of sexual divergence. The digestive tissue of the adult midgut in the abdomen should have minimal sexual dimorphism since feeding activity is limited during lepidopteran adulthood. Insect adult head is the olfactory center for courtship and mating and is thus presumed to be one of the more sexually dimorphic tissues in

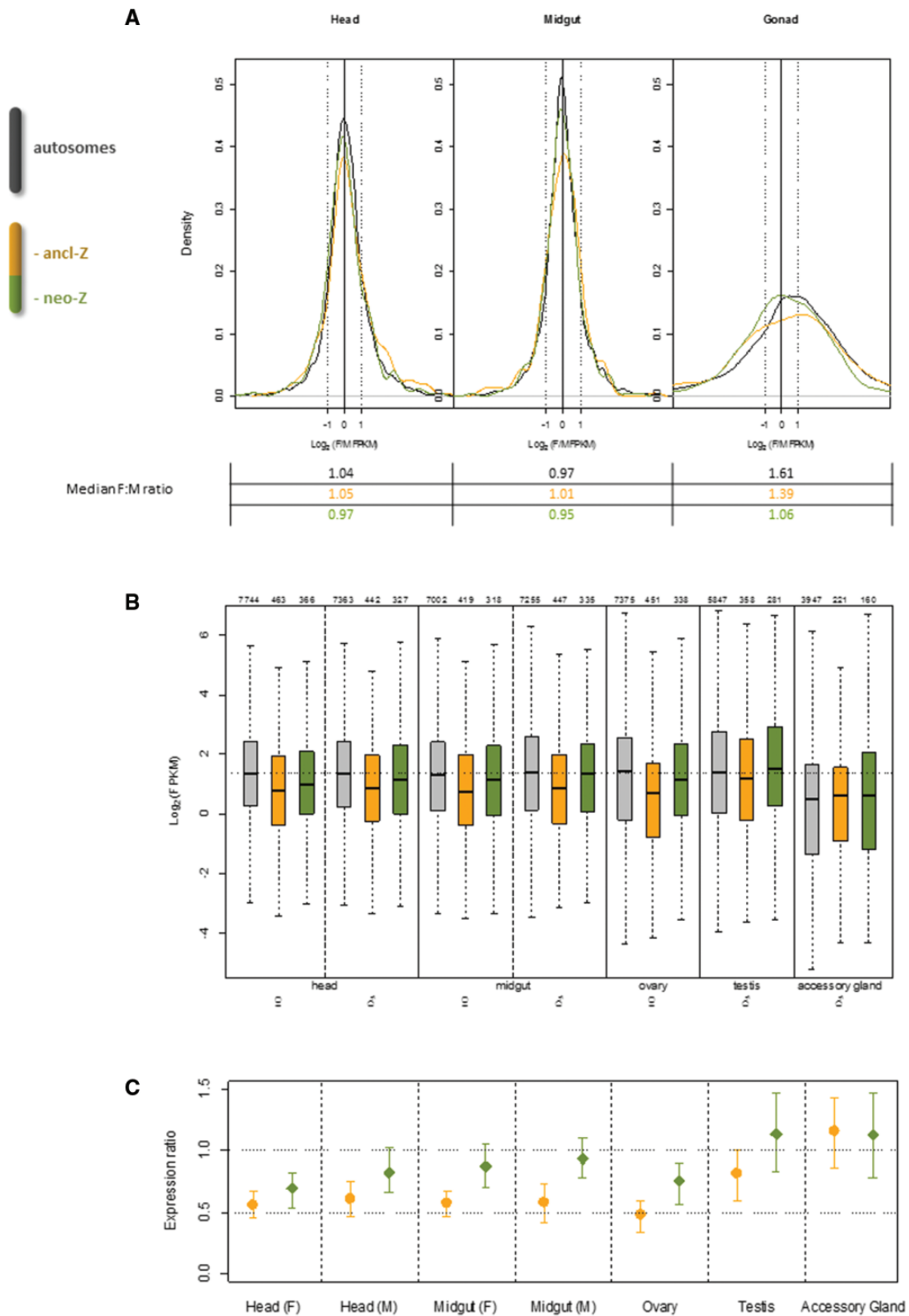
the insect soma. Reproductive tissues are sex-specific, representing a maximum of sexual dimorphism. Among the male reproductive tissues (fig. 3), testis is the gonadal counterpart of the ovary in females, whereas the remaining components are of somatic origin (sex-specific somatic tissues). The male reproductive tract was dissected separately from the testis and treated as a single tissue, referred to simply as “accessory gland” in this analysis. Pooling RNA-sequencing data from all tissues and samples yielded 238 million pairs of post-filtering reads, from which we constructed a de novo assembly from the combined transcriptome of these *C. pomonella* tissues. All sequencing data has been deposited into GenBank’s Sequence Read Archive (accession number SRP083782), associated BioProject PRJNA341278, and the de novo transcriptome assembly was deposited in Figshare (DOI: 10.6084/m9.figshare.4668223).

Synteny in the Lepidoptera is unusually well-conserved across very substantial evolutionary divergences (Labandeira et al. 1994; Pringle et al. 2007; Singh et al. 2009; *Heliconius* Genome Consortium 2012; Van’t Hof et al. 2013; Ahola et al. 2014; Kanost et al. 2016). Thus the conserved synteny in the Lepidoptera allowed us to assign putative chromosomal locations of *C. pomonella* genes based on homology to the *B. mori* reference genome. Among assembled contigs, we identified 447 *C. pomonella* genes putatively located on the ancestral Z segment (ancl-Z), 535 on the neo-Z segment (homologous to *B. mori* chromosome 15), and 9,053 on the autosomes. Only these 10,616 contigs were used in all subsequent analyses.

### Different Patterns of Dosage Compensation in Somatic Tissues and Gonads

Using the de novo transcriptome contigs assigned to neo-Z, ancl-Z, or autosomes, we first assessed dosage compensation in head and midgut. To assess Z~ZZ balance, we evaluated F:M expression ratios using all expressed genes (FPKM > 0) in both sexes in these tissues. We observed balanced expression (median F:M ratios close to 1) for both segments of the Z chromosome and autosomes with no significant difference detected between the autosomes and either Z segment (fig. 4A, left and center panels and supplementary table S1, Supplementary Material online). Also, no difference in average F:M expression ratio was detected between neo-Z versus ancl-Z in these somatic tissues.

We similarly assessed Z~ZZ balance in ovaries and testis using all expressed genes (FPKM > 0) in both sexes (fig. 4A, right panel). In these highly sexually dimorphic germinal tissues, the distribution of F:M expression ratios were quite distinct from those observed in head and midgut. Notably, the variance in ratios was much more extreme than in somatic tissues. This coincides with an extremely poor correlation in expression levels between sexes in gonad (Pearson’s  $\rho = 0.12$ ), compared with the head (Pearson’s  $\rho = 0.92$ ) or the midgut (Pearson’s  $\rho = 0.73$ ). The autosomes and Z segments all



**FIG. 4.**—Dosage compensation analysis in *Cydia pomonella* using tissue-specific RNAseq. (A) Z-Z balance in *C. pomonella* head, midgut and gonads assessed by the  $\log_2(\text{female}:\text{male})$  expression ratio. Key reference  $\log_2$  ratios are marked:  $-1$  (2-fold male over female expression);  $0$  (equal expression

showed a prominent female bias in median gene expression, though the magnitude of this bias was stronger for autosomes than either Z segment. In particular, the neoZ showed significantly higher male expression, on an average, than the autosomes ( $P < 3 \times 10^{-3}$ , Bonferroni-corrected Mann–Whitney  $U$  test; supplementary table S1, Supplementary Material online).

To assess Z~A parity, we compared average gene expression of the ancl-Z or neo-Z segments with autosomes, using all expressed genes (FPKM > 0) in the focal tissue and sex (fig. 4B). Ratios of median expression for neo-Z or ancl-Z versus autosomes, with bootstrapped 95% confidence intervals, are plotted in figure 4C. In head and midgut for both sexes, the ancl-Z:AA or ancl-ZZ:AA ratios (collectively symbolized as ancl-Z(Z):AA) ranged from 0.52 to 0.61, which are not significantly different than 0.5. In contrast, the neo-Z(Z):AA ratios were all significantly >0.5 and, except in female head, not significantly different from one. These results broadly indicate that in *C. pomonella* head and midgut, the extent of Z~A parity is greater for neo-Z genes than for ancl-Z genes.

Gonads showed a much greater discrepancy between sexes in Z~A parity than head or midgut (fig. 4C). In ovary, ancl-Z:AA and neo-Z:AA ratios were significantly <1, in contrast to testis where confidence intervals for both ratios overlapped one. Values for ancl-ZZ:AA and neo-ZZ:AA ratios are notably greater in the male reproductive tissues than other tissues.

### Interspecies Analyses to Assess Dosage Compensation

*C. pomonella* orthologs of neo-Z genes are autosomal in most lepidopteran species outside of the Tortricidae. Thus the existence of analogous RNA-seq data sets from adult head of *M. sexta* (Smith et al. 2014) and *Heliconius melpomene* (Walters et al. 2015) permits a novel assessment of the apparently reduced expression among neo-Z genes in this tissue in *C. pomonella*. Specifically, we can compare expression ratios for the same genes between species using a single sex. Orthologs of *C. pomonella* ancl-Z, neo-Z and autosomal genes in the reference species are denoted as ancl-Z(Z), neo-ZZ and AA, respectively. Crucially, the neo-Z(Z): neo-ZZ comparison contrasts Z-linked genes in *C. pomonella* with their autosomal orthologs in the other species, with neo-ZZ expression levels representing the ancestral state. The robustness of this comparative transcriptome method was supported by the substantial correlation in expression levels between interspecific

orthologous pairs (Pearson's  $\rho = 0.55$  for the *C. pomonella*~*M. sexta* pair and  $\rho = 0.4$  for the *C. pomonella*~*H. melpomene* pair).

Using either *M. sexta* or *H. melpomene* as the reference, the neo-Z(Z): neo-ZZ median ratios were ~0.7 in both sexes, significantly lower than autosomes (AA: AA) ( $P < 5e-05$ , Bonferroni-corrected Mann–Whitney  $U$  test, fig. 5A). This indicates that overall gene expression on neo-Z is equally reduced in both sexes relative to the inferred ancestral expression. This expression reduction does not appear to be universal among neo-Z genes, however, with strongly expressed genes appearing to be better compensated than more weakly expressed genes. For example, using *M. sexta* as the reference, the distributions of neo-Z(Z) and neo-ZZ expression are concordant for genes with  $\log_2(\text{FPKM})$  values >5 (an arbitrary threshold chosen post hoc to illustrate this point; fig. 5B). In contrast, below this threshold, neo-Z(Z) expression is shifted significantly towards lower values compared with those of neo-ZZ ( $P < 10^{-15}$ , Bonferroni-corrected Komolgorov–Smirnov tests) (fig. 5B).

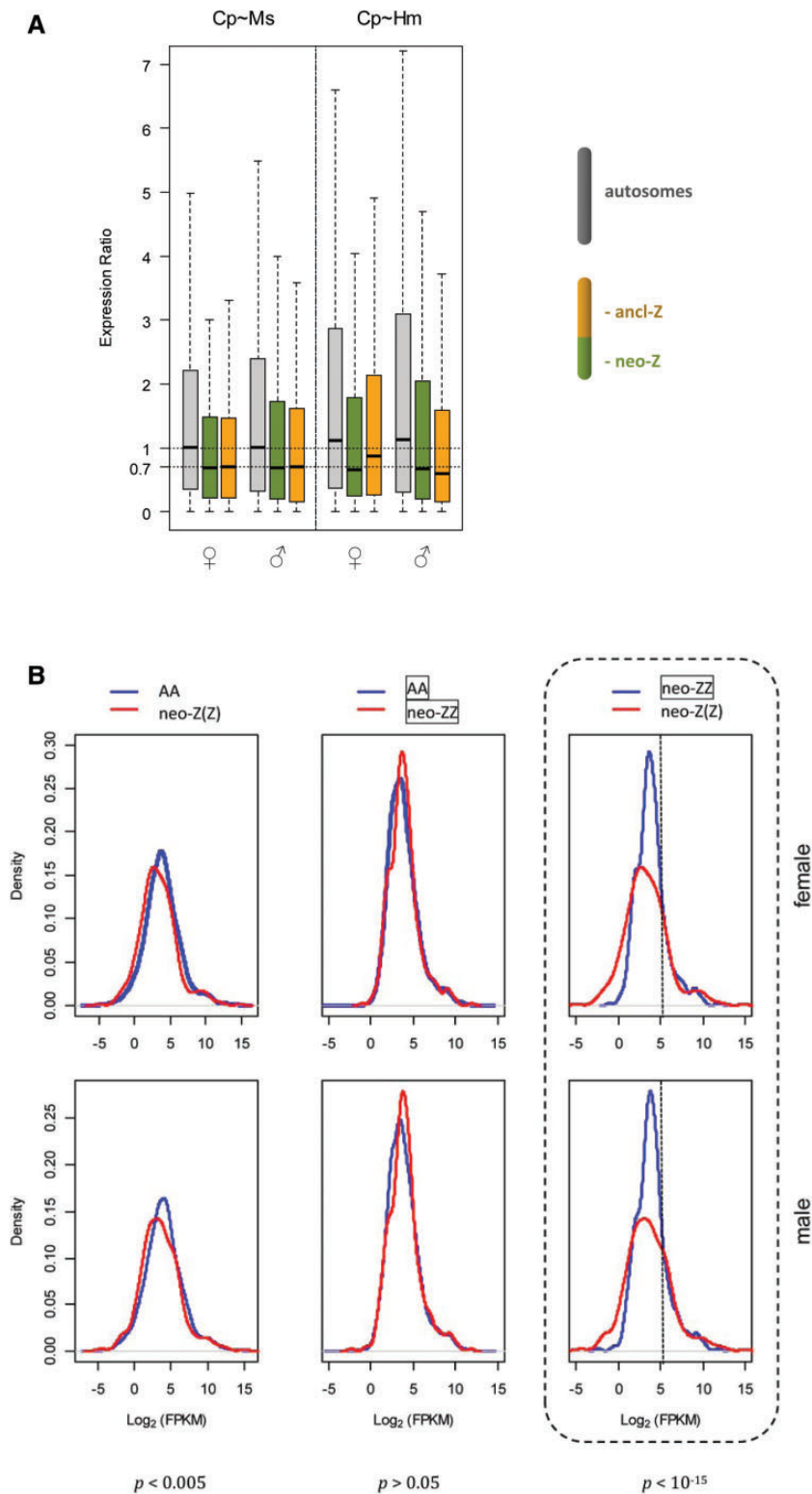
Because ancl-Z genes in *C. pomonella* are also Z-linked in *M. sexta* and *H. melpomene*, this comparative analysis is not informative concerning the ancestral diploid expression level of ancl-Z genes. Nonetheless, a strikingly similar reduction (~30%) in expression was observed for *C. pomonella* ancl-Z(Z) loci relative to *M. sexta* (fig. 5A). In comparison with *H. melpomene*, the reduction of ancl-Z expression was unequal between sexes, which corroboratively reflects the minor Z dosage-effect in *H. melpomene* reported by (Walters et al. 2015).

### Nonrandom Genomic Distribution of Tissue-Specific Genes in *C. pomonella*

We used EdgeR to identify genes differentially expressed between tissues, employing a significance threshold of FDR-adjusted  $P$  value < 0.05. Using these pairwise comparisons between tissues, we identified tissue-specific genes and normalized the numbers of tissue-specific genes to the total numbers of expressed genes found on autosomes, ancl-Z, or neo-Z segments (fig. 6). Tissue-specific genes are defined via differential expression as having consensus up-regulation in a particular tissue as compared with any other tissue for a given fold change (FC) cutoff value. Few differentially expressed genes were identified between the sex-specific samples in head (20 genes) and midgut (8 genes); we therefore pooled

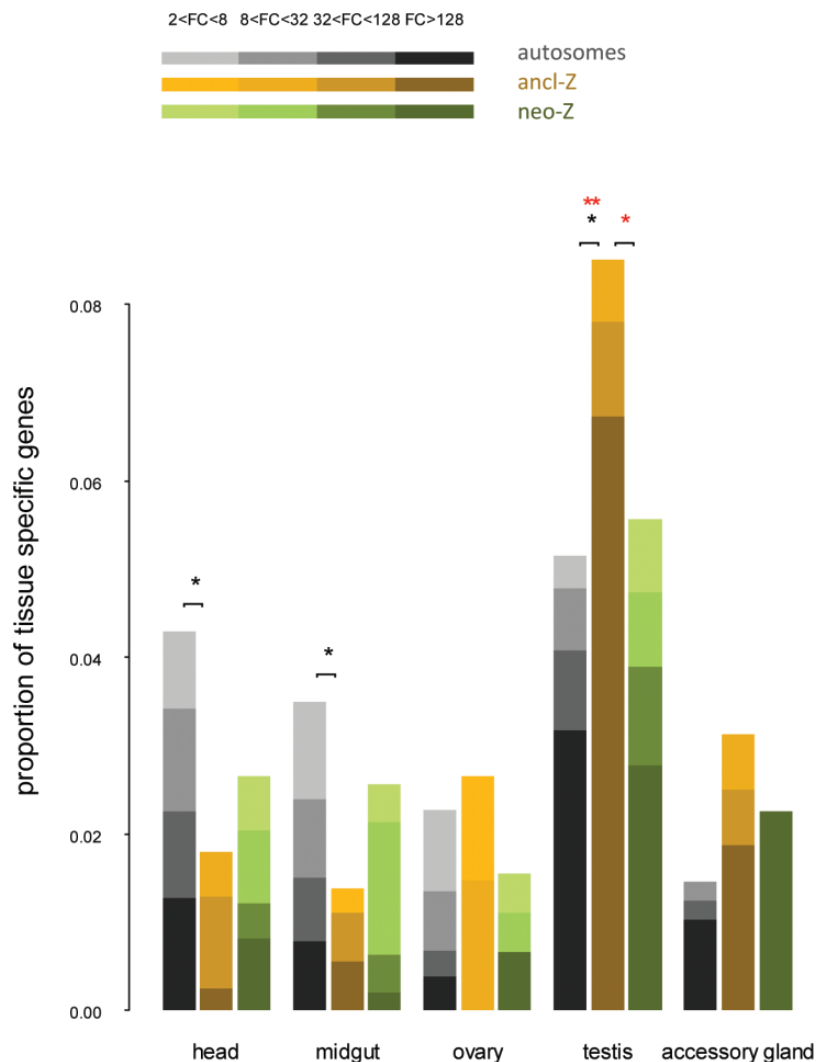
Fig. 4.—Continued

between sexes) and 1 (2-fold female over male expression). (B) Expression levels ( $\log_2$  transformed FPKM) of all expressed genes (FPKM > 0) in each tissue. The number of total expressed genes is indicated above each boxplot. Plot boxes represent the median and interquartile range of expression levels. The whiskers extend to the most extreme data point that is no more than 1.5 times the interquartile range. The dotted horizontal line across the whole plotting area denotes the mean autosomal expression in the head (F&M) and midgut (F&M). (C) Z(Z):AA ratios of median expression of all expressed genes (FPKM > 0) in each tissue (orange: ancl-Z(Z):AA; green: neo-Z(Z):AA). Error bars show 95% confidence intervals estimated by 1,000 bootstrap replicates. The horizontal dotted lines denote the reference ratios: 0.5 indicates Z(Z) expression equal to monosomic autosomal expression (A); 1 indicates Z(Z) expression equal to disomic autosomal expression (AA).



**FIG. 5.**—Dosage compensation analysis in *Cydia pomonella* based on interspecific adult head RNAseq data. (A) Expression ratios of orthologous genes expressed in adult head of *C. pomonella* (Cp) and reference species (Ms: *Manduca sexta*, Hm: *Heliconius melpomene*). Grey: AA: AA, green: neo-Z(Z): neo-ZZ, orange: ancl-Z(Z): ancl-ZZ, where AA represents all autosomal genes of *M. sexta* or *H. melpomene* except those orthologous to *C. pomonella* neo-Z(Z), and neo-ZZ represents autosomal genes of *M. sexta* or *H. melpomene* orthologous to *C. pomonella* neo-Z(Z). (B) Distributions of expression levels of orthologous genes expressed in adult head of *C. pomonella* and *M. sexta*. Bonferroni corrected *P* values from Komolgorov–Smirnov tests are shown below each category.





**Fig. 6.**—Genomic distribution of tissue-specific genes in *Cydia pomonella*. Tissue-specific genes were identified under fold-change (FC) cutoff values from 2 to 128. Statistically significant differences assessed by two tailed Fisher's exact test are indicated on top of bars: single asterisk (\*)  $P < 0.05$ ; double asterisk (\*\*)  $P < 0.01$  | black asterisk: FC cutoff=2; red asterisk: FC cutoff=128.

sexes for these somatic tissues when contrasted to reproductive tissues in order to reduce complexity of the analysis.

Overall, testis-specific genes were generally overrepresented in the genome compared with other tissue-specific genes. Notably, the ancl-Z segment harbored a much larger proportion of testis-specific genes than either the autosomes or the neo-Z segment (fig. 6). This pattern was significant at all FC cutoff values ( $P < 0.05$ , Fisher's exact test) and became more pronounced with increasing FC thresholds, such that at  $FC = 128$  the proportion of testis-specific genes was twice as large on the ancl-Z segment as on the autosomes. In contrast, head- and midgut-specific genes were significantly underrepresented on the ancl-Z segment ( $P < 0.05$ , Fisher's exact test). Tissue-specific genes from ovary and accessory glands showed no significant differences in chromosomal

distribution, though accessory gland genes did show a pattern of chromosomal distribution qualitatively similar to testes, with relatively more genes on the Z.

## Discussion

### Somatic Tissues: Reduced Z-Linked Expression in Males Balances Expression between the Sexes

We focus our discussion first on the results from nonreproductive somatic tissues. Comparison of gene expression patterns between sexes in *C. pomonella*, and also relative to other lepidopteran species, generally supports the idea that sex chromosome dosage compensation in the Lepidoptera involves males down-regulating their diploid complement of Z-chromosome genes to achieve balanced expression with

the monosomic Z in females. The balanced expression between males and females is most evident from the lack of any dosage effect involving Z chromosome; that is gene expression on both ancl-Z and neo-Z segments has distributions of F:M ratios comparable to that of the autosomes (fig. 4A). This lack of gene dosage effect on the Z chromosome indicates that *C. pomonella* completely compensates Z-linked expression for differences in gene dosage between the sexes in nonreproductive somatic tissues. These results are consistent with those obtained in previous investigations of dosage compensation in *B. mori* (Walters and Hardcastle 2011) and *M. sexta* (Smith et al. 2014), and similar to results from *Heliconius* butterflies (which did show a modest dosage effect) (Walters et al. 2015). In contrast to the absence of an obvious Z dosage effect in *C. pomonella*, data obtained from *P. interpunctella* suggest a substantial difference in Z:A ratios between the sexes of this species, although the distributions of F:M expression ratios were not directly reported for *P. interpunctella* (Harrison et al. 2012).

In these nonreproductive tissues, evidence that this balanced Z expression between sexes results from male-down regulation on the Z stems from two different comparisons to autosomal expression: 1) Z versus autosomal expression within *C. pomonella* and 2) comparisons of genes on the *C. pomonella* neo-Z segment with autosomal orthologs in other species. There is a long-established precedent of examining Z:A expression ratios in each sex as an assessment of sex chromosome dosage compensation, with the assumption that average autosomal expression serves as a proxy for ancestral expression levels of a pair of proto-Z chromosomes (Mank et al. 2011). As has been reported previously for a few lepidopteran species (*B. mori*, *M. sexta*, *Heliconius spp.*), expression of ancl-Z genes is greatly reduced relative to autosomal loci in both sexes of *C. pomonella*, suggesting that expression of ancl-Z genes is reduced in both sexes relative to expectations from an unmodified diploid state. In these other species, the Z:A expression ratios were reported to be  $\sim 0.7$ . Notably, in the case of *C. pomonella*, the ancl-Z(Z):AA ratio is quite close to 0.5, a value consistent with functionally monosomic expression. These results fit cleanly with the notion that males are down-regulating their diploid Z-chromosome gene expression. However, they also do not necessarily exclude the possibility of female up-regulation of proto-Z gene expression that was unusually low relative to that of the other autosomes.

Comparisons of the expression of *C. pomonella* neo-Z genes to autosomal gene expression yielded more equivocal evidence for reduced Z-linked expression than comparisons between the expression of ancl-Z genes and the autosomes. Point estimates of neo-Z(Z):AA ratios were all  $< 1$ , and often substantially so, which is consistent with generally reduced neo-Z expression relative to autosomes. However, only in one case (female head) did the confidence intervals not overlap one, so these data do not provide strong statistical support for reduced neo-Z expression relative to autosomes.

Unfortunately, these mixed results from the neo-Z segment compared with the other *C. pomonella* autosomes do not offer a robust answer to the fundamental question about the nature of dosage compensation in the Lepidoptera: does the reduced expression of Z-linked loci in both sexes reflect down-regulation in the male or up-regulation of the monosomic female? However, a comparison of the expression of *C. pomonella* neo-Z genes to their autosomal orthologs in other species strongly supports the down-regulation of Z chromosome genes in males.

Before further discussing the outcome of comparative analyses, it is worth re-emphasizing that evaluating dosage compensation by comparison of the Z chromosome to other autosomes in the same species is problematic because it typically involves contrasting the expression between sets of nonhomologous genes. The evolution of dosage compensation during sex chromosome divergence is thought to arise from a need to restore the expression of sex-linked genes to their ancestral levels on the proto-sex chromosomes (Ohno et al. 1959; Ohno 1967). This ancestral expression level cannot be measured directly, but in some cases it may be reasonably inferred from the expression of autosomal orthologs in related species. As such, interspecific comparisons to the expression of autosomal orthologs arguably provides a better reference point for inferring ancestral expression than conspecific comparisons to nonhomologous autosomal loci (He et al. 2011; Julien et al. 2012; Lin et al. 2012; Albritton et al. 2014; Nozawa et al. 2014).

In most other lepidopterans, *C. pomonella* orthologs of neo-Z genes are autosomal. This unique genetic circumstance enables us to evaluate expression reduction of *C. pomonella* neo-Z genes in the head using the existing data sets from *M. sexta* (Smith et al. 2014) and *H. melpomene* (Walters et al. 2015). This comparison between species revealed that the expression of neo-Z loci was substantially reduced in both sexes relative to the expression of their autosomal orthologs (neo-Z(Z): neo-ZZ  $\sim 0.7$ ) (fig. 5A), indicating that overall expression of genes on the *C. pomonella* neo-Z segment is equally reduced in both sexes relative to the inferred ancestral expression. This observation supports the notion of broadly down-regulated Z-linked gene expression in males to achieve parity with females. However, this pattern seems to depend on the absolute magnitude of expression, as illustrated by comparing genes with expression levels greater versus less than  $\log_2(\text{FPKM})=5$ . We found that strongly expressed genes were more similar to autosomal orthologs than weakly expressed genes, and therefore better compensated relative to ancestral expression (fig. 5B). Such genes are expected to be dosage-sensitive and hence under the strongest selective pressure for dosage compensation (Gout et al. 2010).

This comparative analysis indicating reduced expression in males as a mechanism for achieving balanced expression with the monosomic Z in females nicely corroborates recent genetic experiments in *B. mori* examining sex determination and

dosage compensation mechanisms (Kiuchi et al. 2014). In *B. mori*, the sex determination pathway involves interactions between a W-linked piRNA that targets a Z-linked gene, *Masc*, that masculinizes embryos. When siRNA is used to disrupt *Masc* function in males, there is chromosome-wide, Z-specific increase in gene expression. This indicates that indeed there is a male-specific mechanism, depending on the function of *Masc*, that specifically reduces expression across the Z chromosome. Our comparative analysis provides an important complement to this experimental result, which stand together as independent evidence for male-specific down-regulation of the Z chromosome. Thus both experimental and computational analyses increasingly support the idea that lepidopteran dosage compensation does not fit well with the precedent from other female-heterogametic taxa (Ellegren et al. 2007; Itoh et al. 2007, 2010; Vicoso and Bachtrog 2011; Wolf and Bryk 2011; Adolfsson and Ellegren 2013; Uebbing et al. 2013; Vicoso et al. 2013; Chen et al. 2014), or the theoretical expectations for up-regulation of the monosomic Z chromosome in females (Ohno 1967; Charlesworth 1978; Vicoso and Bachtrog 2009; Mank 2013).

One unexpected result arising from this comparative analysis was that, like the neo-Z, *C. pomonella* ancl-Z loci showed reduced expression relative to orthologs in either reference species. One possible scenario to explain this reduction in *C. pomonella* ancl-Z expression relative to Z-linked orthologs is that Z-linked expression may have increased among the more derived species. Indeed, an intriguing phylogenetic trend of increasing Z:AA ratios is observed among more derived species, at least in the head, which is the only tissue available for comparison. Specifically, ancl-(Z:AA|ZZ:AA) ratios were the lowest in *C. pomonella* (0.56|0.62), whereas Z:AA|ZZ:AA ratios are 0.63|0.69 in *H. melpomene* (Walters et al. 2015), 0.761|0.766 in *B. mori* (Walters and Hardcastle 2011) and 0.8|0.83 in *M. sexta* (Smith et al. 2014). We thus suggest that the extent of Z~A parity correlates with the phylogenetic hierarchy in the Lepidoptera such that more derived species are generally more completely compensated. Nevertheless, this taxonomic sampling is sparse and this pattern must be confirmed with data from additional lepidopteran species. An alternative explanation for this pattern is that ancl-Z expression has become reduced in *C. pomonella* (and related Tortricids) relative to other lineages, perhaps as a result of the Z-autosomal fusion. If there were selection to reduce expression of genes on the neo-Z segment, as appears to have occurred, then perhaps expression of the entire chromosome was reduced, with lowered ancl-Z expression as a by-product.

### Reproductive Tissues: Reduced Female Expression on the Z Chromosome

Analyses of dosage compensation in reproductive tissues (ovary, testis, and accessory gland) revealed patterns different

from those observed in head and midgut, with the overall pattern indicating dosage compensation is substantially reduced or absent. This is most apparent in the comparisons of absolute expression levels between the Z chromosome and the autosomes. Among all the tissues, testis and accessory gland have the highest ancl-ZZ:AA and neo-ZZ:AA ratios, with confidence intervals overlapping one (fig. 4C). This result is in good agreement with the *B. mori* data, which also show a median ZZ:AA ratio near one in the testis and no significant difference in Z versus autosomal average expression (Walters and Hardcastle 2011). In contrast, ovary had the lowest median ancl-Z:AA ratio of all tissues, with Z:Autosome median ratios falling significantly <1 (fig. 4C). It thus appears that in the reproductive tissues, gene expression of both ancl-Z and neo-Z genes is strongly influenced by the underlying chromosomal ploidy. This observation stands in stark contrast to the apparent absence of a gene dosage effect between sexes in both the head and midgut, and suggests the absence of any global form of dosage compensation in the reproductive system; in these tissues males apparently are not down-regulating their diploid Z chromosomal gene expression.

Additional evidence supporting the absence of dosage compensation in reproductive tissues comes from the comparative transcriptome analyses. The expression levels of both neo-ZZ and AA represent gene expression in the unmodified diploid state. The neo-ZZ: AA ratios of median expression estimated from head tissues in both reference species were close to 1.1, which were significantly (bootstrap tests) higher than the expected value of 1 (supplementary fig. S1, Supplementary Material online). Strikingly, this value of 1.1 coincides with the neo-ZZ:AA ratios of median expression in *C. pomonella* testis and accessory gland (1.13 and 1.14, respectively). This consistency between species in median expression of neo-Z relative to autosomes suggests that neo-Z gene expression in the male reproductive tissues reflects ancestral expression levels, a pattern that is consistent with the absence of any dosage compensating mechanism modifying Z-linked expression in males.

Since comparisons of absolute expression levels indicate increased male Z-linked expression in gonads, we might also expect F:M expression ratios to show a prominent male bias on the Z chromosome relative to autosomes. Unfortunately, robust interpretations of F:M expression ratios in the gonads is complicated by a strong overall female-bias, which is unexpected and cannot be easily explained. This bias may be an artifact of the extreme sexual dimorphism observed in gonadal expression levels. In comparison to somatic tissues, this strong gonadal dimorphism is evident in the very poor correlation in expression levels between sexes and the much reduced count of genes with expression in both sexes. Such dimorphism likely makes data based on ratios extremely “noisy,” and may account for the increased variance and the apparent female bias in point estimates of median expression. Nonetheless, there is some qualitative evidence in these results that point to a

distinct dosage effect that increases male expression on the Z in the gonads. Taking the autosomal distribution of F:M expression ratios as a reference point, both Z segments exhibit much reduced average values. This trend towards lower F:M ratios is also evident in the plots of distributions (fig. 4A, right panel), where the Z-segments have a notably greater density of observations among lower values as compared with the autosomal distribution. In short, strong sexual dimorphism in gonadal expression seemingly limits a confident assessment of dosage effects (i.e. reduced dosage compensation) based on expression ratios, but the qualitative pattern is consistent with results from absolute expression levels discussed earlier.

### Nonrandom Genomic Distribution of Tissue-Specific Genes in *C. pomonella*: Implications for Dosage Compensation

Studies on a variety of species, including lepidopterans, indicate that sex chromosomes typically have unusual complements of genes that are expressed predominantly in one sex (sex-biased genes), especially in the germline (Saifi and Chandra 1999; Lercher et al. 2003; Parisi et al. 2003; Khil et al. 2004; Reinke et al. 2004; Arunkumar et al. 2009). Walters and Hardcastle (2011) suggested that in *B. mori*, over-representation of strongly expressed testis-specific genes on the Z chromosome could account for the exceptional ZZ:AA ratio in the testis, yet this explanation was not directly tested. To further address this issue, we explored the interplay between genomic distribution of tissue-specific genes and dosage compensation patterns in *C. pomonella*. Indeed, testis-specific genes were generally over-represented in the genome compared with other tissue-specific genes, with the ancl-Z segment harboring a much larger proportion of testis-specific genes than either the autosomes or the neo-Z segment (fig. 6). Additionally, all the subsets of tissue-specific genes, particularly the testis- and accessory gland-specific genes, were among the most highly expressed genes in every tissue (supplementary fig. S2, Supplementary Material online). Nevertheless, removing all of the 1,247 tissue-specific genes (at FC = 2) from analyses of dosage compensation did not appreciably change the ancl- or neo-Z(Z):AA ratios in any tissue in comparison to calculations based on the complete data set (supplementary fig. S3, Supplementary Material online). Therefore, it appears that the nonrandom genomic distribution of tissue-specific genes did not confound the assessment of Z~A compensation in *C. pomonella*. In particular, this result lends confidence to the notion that the distinct Z:A expression ratios reflect largely unmitigated dosage effects in reproductive tissues and does not arise from a biased Z-linked distribution of highly expressed genes as suggested by Walters and Hardcastle (2011).

Sex-biased gene expression is typically associated with sexual dimorphism and, accordingly, germline-specific genes represent a substantial fraction of sex-biased genes (Ellegren and Parsch 2007). The over-representation of testis-specific

genes on the Z chromosome in *C. pomonella* provided further evidence supporting the role of sexually antagonistic selection in shaping the chromosomal distribution of genes (Rice 1984). However, this theory cannot explain the under-representation of head- and midgut-specific genes, since these genes do not involve sex-specific expression. We suggest that the lack of general Z~A compensation in the head and midgut would make the Z chromosome a suboptimal environment for genes that require high expression, hence driving selection for their localization on the autosomes. Consistent with such a scenario, the neo-Z and the ancl-Z segments both showed a deficit in head- and midgut-specific genes but to different extents, which may reflect the distinct level of Z~A compensation between the two chromosomal segments. This pattern seems to agree well with the findings from *D. melanogaster*, where constraints from the dosage compensation mechanism are suggested to account for the paucity both of male-biased genes with high expression on the X (Vicoso and Charlesworth 2009) and of tissue-biased genes that are not directly associated with reproduction (Mikhaylova and Nurminsky 2011).

### Inconsistencies in Dosage Compensation Patterns Reported for *P. interpunctella* Compared with Studies of Other Lepidopteran Species May Reflect Analytical Artifacts

Dosage compensation patterns are typically conserved in taxa with highly conserved sex chromosomes like the Lepidoptera (Julien et al. 2012; Lin et al. 2012; Vicoso and Bachtrog 2015). In this respect, the pattern ( $Z < ZZ \approx AA$ ) reported in *P. interpunctella* (Harrison et al. 2012) appears anomalous since analyses of a more basal species (*C. pomonella*) and three more derived species (*B. mori*, *M. sexta* and *Heliconius spp.*) suggest a conserved phylogenetic pattern ( $Z \approx ZZ < AA$ ). While it may be that *P. interpunctella* represents a true anomaly in the Lepidoptera, this inconsistent finding motivates a careful consideration of possible artifacts that may explain this inconsistency with the noted phylogenetic trend. Consider first that the analysis of dosage compensation in *P. interpunctella* sampled whole-body adult insects, which contain a substantial fraction of reproductive tissues. Evidence from both *B. mori* (Walters and Hardcastle 2011) and *C. pomonella* (this study) point to an absence of dosage compensation in the reproductive tissues. Thus analyzing reproductive tissues intermixed with soma is prone to bias the overall assessment of dosage compensation.

Second, in the *P. interpunctella* study the FPKM values were not TMM-normalized, which would contribute to the median autosomal expression being notably higher in females relative to males. This is demonstrated in the analysis of our *C. pomonella* data, where the median autosomal expression was higher in the ovary than in the testis before TMM normalization, whereas TMM normalization resulted in a distribution of FPKM values in the gonads similar to somatic tissues



(supplementary fig. S4, Supplementary Material online). By comparison, TMM normalization did not appreciably change the FPKM values for the head and midgut, which is consistent with the *M. sexta* study where only head was sampled (Smith et al. 2014). In accordance with the rationale for TMM normalization (Robinson and Oshlack 2010), this observation likely reflects the substantial differences in RNA populations between the testis and the ovary relative to the other tissues. Further clarification of dosage compensation in *P. interpunctella* will require sampling tissue(s) in isolation from reproductive tissues.

### Similarities in Dosage Compensation between Lepidopterans and Therian Mammals

Intriguingly, our results from *C. pomonella* add further evidence that lepidopterans and therian mammals have evolved convergent patterns of dosage compensation. First, both taxa show a general pattern of reduced sex-linked gene expression in the homogametic sex in the soma ( $Z \approx ZZ < AA$  vs.  $X \approx XX < AA$ ) (Julien et al. 2012; Lin et al. 2012). Also, both appear to lack dosage compensation in the germline (Sugimoto and Abe 2007). A more nuanced similarity exists in the patterns of X~A/Z~A dosage compensation. In mammals, after substantial rounds of analysis and debate, it has become recognized that X~A parity (i.e. “compensatory up-regulation” of the monosomic X) operates on a minority of X-linked genes that are mostly highly expressed, that is expression is closer to the inferred ancestral state (Julien et al. 2012; Lin et al. 2012; Pessia et al. 2012; Chen and Zhang 2015). Through comparative interspecies transcriptome analyses of genes expressed on the *C. pomonella* neo-Z, we have shown the same trend, that is that Z-linked genes retaining ancestral expression levels are mostly those with higher expression.

### Conclusion

In this study we address the issue of conservation of dosage compensation patterns in the Lepidoptera by analyzing tissue- and sex-specific gene expression data from *C. pomonella*, while scrutinizing previous studies of other species. We show equal reduction of Z-linked expression relative to autosomal expression in adult soma of females and males ( $Z \approx ZZ < AA$ ), with the most compelling evidence from direct comparisons of expression between *C. pomonella* neo-Z genes with their autosomal orthologs in *M. sexta* and *H. melpomene*. Our data also indicate the absence of dosage compensation and highly sex-specific gene expression in the reproductive tissues. We thus argue that inclusion of these germinal tissues coupled with artifacts in data analyses contributed to the incongruous results and conclusions obtained in the study of dosage compensation in *P. interpunctella* (Harrison et al. 2012). Indeed, with exception of the latter, the bulk of published evidence suggests that dosage compensation is likely conserved in the Lepidoptera. Assuming this

conclusion is borne out by additional studies of other lepidopteran species, this insect order would represent the only female heterogametic (WZ/ZZ) taxon known to date to equalize sex-linked gene expression between the sexes. The tantalizing similarities of dosage compensation patterns of lepidopterans and therian mammals, with opposing sex chromosome constitutions (WZ/ZZ vs. XX/XY), point to a compelling area for future studies addressing the extent to which these patterns share similar molecular mechanism. Studies of moths and butterflies challenge the traditional view of dosage compensation’s association with sexual heterogamety, while offering a unique opportunity to decipher the evolution of sex chromosome dosage compensation.

### Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

### Acknowledgments

We thank David Soderlund, Ping Wang and František Marec for their valuable comments and input on the manuscript. This work was supported in part by the Federal Formula Fund (6217449) granted to D.C.K. and in part by the Cornell University Griswold Award for graduate research (granted to L.G.) from the Department of Entomology, Cornell University (1398103-00514). This research was also supported in part by the National Science Foundation NSF-DEB 1457758 (granted to J.R.W.).

### Literature Cited

- Adolfsson S, Ellegren H. 2013. Lack of dosage compensation accompanies the arrested stage of sex chromosome evolution in ostriches. *Mol Biol Evol.* 30:806–810.
- Ahola V, et al. 2014. The *Glanville fritillaria* genome retains an ancient karyotype and reveals selective chromosomal fusions in Lepidoptera. *Nat Commun.* 5:4737.
- Albritton SE, et al. 2014. Sex-biased gene expression and evolution of the X chromosome in nematodes. *Genetics* 197:865–883.
- Arunkumar KP, Mita K, Nagaraju J. 2009. The silkworm Z chromosome is enriched in testis-specific genes. *Genetics* 182:493–501.
- Bean CJ, Schaner CE, Kelly WG. 2004. Meiotic pairing and imprinted X chromatin assembly in *Caenorhabditis elegans*. *Nat Genet.* 36:100–105.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120.
- Charlesworth B. 1978. Model for evolution of Y chromosomes and dosage compensation. *Proc Natl Acad Sci U S A.* 75:5618–5622.
- Chen S, et al. 2014. Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle. *Nat Genet.* 46:253–260.
- Chen X, Zhang J. 2015. No X-chromosome dosage compensation in human proteomes. *Mol Biol Evol.* 32:1456–1460.
- Conrad T, Akhtar A. 2012. Dosage compensation in *Drosophila melanogaster*: epigenetic fine-tuning of chromosome-wide transcription. *Nat Rev Genet.* 13:123–134.
- Davison AC, Hinkley DV. 1997. Bootstrap methods and their application. London: Cambridge University Press.



- Ellegren H, et al. 2007. Faced with inequality: chicken do not have a general dosage compensation of sex-linked genes. *BMC Biol.* 5:40.
- Ellegren H, Parsch J. 2007. The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev Genet.* 8:689–698.
- Ferrari F, Alekseyenko AA, Park PJ, Kuroda MI. 2014. Transcriptional control of a whole chromosome: emerging models for dosage compensation. *Nat Struct Mol Biol.* 21:118–125.
- Fukova I, et al. 2007. Probing the W chromosome of the codling moth, *Cydia pomonella*, with sequences from microdissected sex chromatin. *Chromosoma* 116:135–145.
- Gelbart ME, Kuroda MI. 2009. *Drosophila* dosage compensation: a complex voyage to the X chromosome. *Development* 136:1399–1410.
- Gout JF, Kahn D, Duret L, Consortium PP-G. 2010. The relationship among gene expression, the evolution of gene dosage, and the rate of protein evolution. *PLoS Genet.* 6:e1000944.
- Grabherr MG, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* 29:644–652.
- Graves JA, Distech CM. 2007. Does gene dosage really matter?. *J Biol.* 6:1.
- Guioli S, Lovell-Badge R, Turner JM. 2012. Error-prone ZW pairing and no evidence for meiotic sex chromosome inactivation in the chicken germ line. *PLoS Genet.* 8:e1002560.
- Haas BJ, et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc.* 8:1494–1512.
- Harrison PW, Mank JE, Wedell N. 2012. Incomplete sex chromosome dosage compensation in the Indian meal moth, *Plodia interpunctella*, based on de novo transcriptome assembly. *Genome Biol Evol.* 4:1118–1126.
- He X, et al. 2011. He et al. reply. *Nat Genet.* 43:1171–1172.
- Heliconius Genome Consortium 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* 487:94–98.
- Itoh Y, et al. 2007. Dosage compensation is less effective in birds than in mammals. *J Biol.* 6:2.
- Itoh Y, et al. 2010. Sex bias and dosage compensation in the zebra finch versus chicken genomes: general and specialized patterns among birds. *Genome Res.* 20:512–518.
- Julien P, et al. 2012. Mechanisms and evolutionary patterns of mammalian and avian dosage compensation. *PLoS Biol.* 10:e1001328.
- Kanost MR, et al. 2016. Multifaceted biological insights from a draft genome sequence of the tobacco hornworm moth, *Manduca sexta*. *Insect Biochem Mol Biol.* 76:118–147.
- Kelly WG, et al. 2002. X-chromosome silencing in the germline of *C. elegans*. *Development* 129:479–492.
- Khil PP, Smirnova NA, Romanienko PJ, Camerini-Otero RD. 2004. The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. *Nat Genet.* 36:642–646.
- Kiuchi T, et al. 2014. A single female-specific piRNA is the primary determinant of sex in the silkworm. *Nature* 509:633–636.
- Labandeira C, Dilcher D, Davis D, Wagner D. 1994. Ninety-seven million years of angiosperm-insect association: paleobiological insights into the meaning of coevolution. *Proc Natl Acad Sci U S A.* 91:12278–12282.
- Laverty C, Lucci J, Akhtar A. 2010. The MSL complex: X chromosome and beyond. *Curr Opin Genet Dev.* 20:171–178.
- Lercher MJ, Urrutia AO, Hurst LD. 2003. Evidence that the human X chromosome is enriched for male-specific but not female-specific genes. *Mol Biol Evol.* 20:1113–1116.
- Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323.
- Lin F, Xing K, Zhang J, He X. 2012. Expression reduction in mammalian X chromosome evolution refutes Ohno's hypothesis of dosage compensation. *Proc Natl Acad Sci U S A.* 109:11752–11757.
- Mank JE. 2009. The W, X, Y and Z of sex-chromosome dosage compensation. *Trends Genet.* 25:226–233.
- Mank JE. 2013. Sex chromosome dosage compensation: definitely not for everyone. *Trends Genet.* 29:677–683.
- Mank JE, Hosken DJ, Wedell N. 2011. Some inconvenient truths about sex chromosome dosage compensation and the potential role of sexual conflict. *Evolution* 65:2133–2144.
- Marec F, Sahara K, Traut W. 2010. Rise and fall of the W chromosome in Lepidoptera. In: Goldsmith M, Marec F, editors. *Molecular biology and genetics of the Lepidoptera*. CRC Press. p. 49–63.
- Meiklejohn CD, Landeen EL, Cook JM, Kingan SB, Presgraves DC. 2011. Sex chromosome-specific regulation in the *Drosophila* male germline but little evidence for chromosome dosage compensation or meiotic inactivation. *PLoS Biol.* 9:e1001126.
- Meyer BJ. 2010. Targeting X chromosomes for repression. *Curr Opin Genet Dev.* 20:179–189.
- Meyer M, Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc.* 2010:t5448.
- Mikhaylova LM, Nurminsky DI. 2011. Lack of global meiotic sex chromosome inactivation, and paucity of tissue-specific gene expression on the *Drosophila* X chromosome. *BMC Biol.* 9:29.
- Monesi V. 1965. Differential rate of ribonucleic acid synthesis in the autosomes and sex chromosomes during male meiosis in the mouse. *Chromosoma* 17:11–21.
- Naurin S, Hansson B, Bensch S, Hasselquist D. 2010. Why does dosage compensation differ between XY and ZW taxa?. *Trends Genet.* 26:15–20.
- Nguyen P, et al. 2013. Neo-sex chromosomes and adaptive potential in tortricid pests. *Proc Natl Acad Sci U S A.* 110:6931–6936.
- Nozawa M, Fukuda N, Ikee K, Gojobori T. 2014. Tissue- and stage-dependent dosage compensation on the neo-X chromosome in *Drosophila pseudoobscura*. *Mol Biol Evol.* 31:614–624.
- Ohno S. 1967. *Sex chromosomes and sex-linked genes*. Berlin: Springer-Verlag.
- Ohno S, Kaplan WD, Kinoshita R. 1959. Formation of the sex chromatin by a single X-chromosome in liver cells of *Rattus norvegicus*. *Exp Cell Res.* 18:415–418.
- Parisi M, et al. 2003. Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. *Science* 299:697–700.
- Pessia E, Engelstadter J, Marais GA. 2014. The evolution of X chromosome inactivation in mammals: the demise of Ohno's hypothesis?. *Cell Mol Life Sci.* 71:1383–1394.
- Pessia E, Makino T, Bailly-Bechet M, McLysaght A, Marais GA. 2012. Mammalian X chromosome inactivation evolved as a dosage-compensation mechanism for dosage-sensitive genes on the X chromosome. *Proc Natl Acad Sci U S A.* 109:5346–5351.
- Pringle EG, et al. 2007. Synteny and chromosome evolution in the Lepidoptera: evidence from mapping in *Heliconius melpomene*. *Genetics* 177:417–426.
- R Core Team 2014. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Available from: <http://www.R-project.org/>.
- Reinke V, Gil IS, Ward S, Kazmer K. 2004. Genome-wide germline-enriched and sex-biased expression profiles in *Caenorhabditis elegans*. *Development* 131:311–323.
- Rice WR. 1984. Sex chromosomes and the evolution of sexual dimorphism. *Evolution* 38:735–742.
- Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140.

- Robinson MD, Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11:1.
- Saifi GM, Chandra HS. 1999. An apparent excess of sex- and reproduction-related genes on the human X chromosome. *Proc Royal Soc B Biol Sci.* 266:203–209.
- Schoenmakers S, et al. 2009. Female meiotic sex chromosome inactivation in chicken. *PLoS Genet.* 5:e1000466.
- Schulz EG, Heard E. 2013. Role and control of X chromosome dosage in mammalian development. *Curr Opin Genet Dev.* 23:109–115.
- Sichova J, Nguyen P, Dalikova M, Marec F. 2013. Chromosomal evolution in tortricid moths: conserved karyotypes with diverged features. *PLoS One* 8:e64520.
- Singh ND, Larracunte AM, Sackton TB, Clark AG. 2009. Comparative genomics on the *Drosophila* phylogenetic tree. *Annu Rev Ecol Evol Syst.* 40:459–480.
- Smith G, Chen YR, Blissard GW, Briscoe AD. 2014. Complete dosage compensation and sex-biased gene expression in the moth *Manduca sexta*. *Genome Biol Evol.* 6:526–537.
- Storey JD, Tibshirani R. 2003. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A.* 100:9440–9445.
- Sugimoto M, Abe K. 2007. X chromosome reactivation initiates in nascent primordial germ cells in mice. *PLoS Genet.* 3:e116.
- Turner JM. 2007. Meiotic sex chromosome inactivation. *Development* 134:1823–1831.
- Uebbing S, Kunstner A, Makinen H, Ellegren H. 2013. Transcriptome sequencing reveals the character of incomplete dosage compensation across multiple tissues in flycatchers. *Genome Biol Evol.* 5:1555–1566.
- Van't Hof AE, et al. 2013. Linkage map of the peppered moth, *Biston betularia* (Lepidoptera, Geometridae): a model of industrial melanism. *Heredity* 110:283–295.
- Vibrantovski MD. 2014. Meiotic sex chromosome inactivation in *Drosophila*. *J Genomics* 2:104–117.
- Vicoso B, Bachtrog D. 2009. Progress and prospects toward our understanding of the evolution of dosage compensation. *Chromosome Res.* 17:585–602.
- Vicoso B, Bachtrog D. 2011. Lack of global dosage compensation in *Schistosoma mansoni*, a female-heterogametic parasite. *Genome Biol Evol.* 3:230–235.
- Vicoso B, Bachtrog D. 2015. Numerous transitions of sex chromosomes in Diptera. *PLoS Biol.* 13:e1002078.
- Vicoso B, Charlesworth B. 2009. The deficit of male-biased genes on the *D. melanogaster* X chromosome is expression-dependent: a consequence of dosage compensation?. *J Mol Evol.* 68:576–583.
- Vicoso B, Emerson JJ, Zektser Y, Mahajan S, Bachtrog D. 2013. Comparative sex chromosome genomics in snakes: differentiation, evolutionary strata, and lack of global dosage compensation. *PLoS Biol.* 11:e1001643.
- Wahlberg N, Wheat CW, Peña C. 2013. Timing and patterns in the taxonomic diversification of Lepidoptera (butterflies and moths). *PLoS One* 8:e80875.
- Walters JR, Hardcastle TJ. 2011. Getting a full dose? Reconsidering sex chromosome dosage compensation in the silkworm, *Bombyx mori*. *Genome Biol Evol.* 3:491–504.
- Walters JR, Hardcastle TJ, Jiggins CD. 2015. Sex chromosome dosage compensation in Heliconius butterflies: global yet still incomplete?. *Genome Biol Evol.* 7:2545–2559.
- Wolf JB, Bryk J. 2011. General lack of global dosage compensation in ZZ/ZW systems? Broadening the perspective with RNA-seq. *BMC Genomics* 12:91.
- Zha X, et al. 2009. Dosage analysis of Z chromosome genes using microarray in silkworm, *Bombyx mori*. *Insect Biochem Mol Biol.* 39:315–321.

Associate editor: Mar Alba