PROCEEDINGS B

royalsocietypublishing.org/journal/rspb

Research

Cite this article: Barata C, Vicoso B. 2026 Single-nucleus resolution of sex-biased expression and dosage compensation in *Drosophila melanogaster*. *Proc. R. Soc. B* **293**: 20252471.
https://doi.org/10.1098/rspb.2025.2471

Received: 24 September 2025
Accepted: 5 December 2025

Subject Category:
Genetics and genomics

Subject Areas:
evolution, genomics, bioinformatics

Keywords:
sexual dimorphism, sex-biased expression, dosage compensation, fruit fly

Author for correspondence:
Carolina Barata
e-mail: carolina.barata@ist.ac.at

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.8208815.

THE ROYAL SOCIETY PUBLISHING

Single-nucleus resolution of sex-biased expression and dosage compensation in *Drosophila melanogaster*

Carolina Barata and Beatriz Vicoso

Institute of Science and Technology Austria, Klosterneuburg, Austria

CB, 0000-0003-1945-2245; BV, 0000-0002-4579-8306

In many species, sex-biased expression is widespread and thought to contribute to sexual dimorphism. While bulk RNA-sequencing has been instrumental in identifying strongly sex-biased genes, it lacks resolution to assess variation across cell-types and tissue compartments. Using single-nucleus expression data from the Fly Cell Atlas, we investigate sex differences in adult *Drosophila melanogaster*. We find that differences in cell-type composition between the sexes are not a major source of sex-bias, as for the vast majority of genes, the degree of sex-bias is similar regardless of whether sex differences in cell-type composition are controlled for or not. Our analysis confirms a deficit of X-linked male-biased genes in the body's somatic tissues that is widespread across cell-types. We also find the excess of X-linked female-biased genes to be associated with nervous system cells in the head but with epithelial cells in the body's somatic tissues, showing that single-nucleus data crucially resolves sex-bias at the cell-type level. We investigate dosage compensation (DC) across 15 tissues and 17 cell-types. We observe that it varies throughout the body. Surprisingly, we observe a lack of DC in a cluster of main cells within the male accessory glands. This result highlights the importance of understanding context-dependent DC.

1. Introduction

Sexual dimorphism is a striking form of phenotypic plasticity. The two sexes can differ in their physiology, their morphology—the lion's mane and the peacock's tail—and their behaviour—e.g. *Drosophila* courtship behaviour [1]. A large proportion of this plasticity between the sexes is thought to stem from sex differences in gene expression [2]. These sex differences in gene expression arise in spite of a shared genome and presumably because of sex-specific fitness optima [3]. The fitness of a genotype is a function of how it performs across diverse circumstances that include different tissues, cell-types, developmental stages, environments [4] and even sexes [5]. Therefore, studying gene expression variation between the sexes is key to describing plasticity in a genome that is shared between females and males.

One way to study sexual dimorphism in gene expression is through the characterization of sex-bias in bulk RNA sequencing (RNA-seq) data, where RNA is extracted from a pooled sample of cells and then sequenced. A large body of literature has described varying levels of sex-bias across taxa (reviewed in [6]), ranging from 1% in rabbit [7] to approximately 90% in *Drosophila melanogaster* inbred lines [8]. There is also variation between tissues or organs [7,9], and there is substantial sex-bias in adult gonadal samples [10] that is not observed in the soma [11]. Notably, sex-bias in *Drosophila* heads is weak compared to other somatic tissues [12]. Sexually dimorphic gene expression at the bulk level within a tissue can be generated as a result of three different conditions: (i) genes are sex-biased in the same cell-types

within a tissue; (ii) genes are sex-biased in one direction in some cell-types, but show sex-bias in the opposite direction in others; and (iii) there are differences in the proportion of different cell-types between females and males. Therefore, bulk RNA-seq limits the resolution at which we can study sex-biased expression.

Gene expression differences between the sexes can also arise as a byproduct of other mechanistic and evolutionary processes and, in that case, have deleterious effects. In an XY sex chromosome system, females have two copies of the X chromosome, whereas males have only one. This dosage imbalance is assumed to be detrimental to males and to select for dosage compensation (DC) mechanisms that regulate X-linked expression and re-establish optimal expression balance. Although there are reported cases of variable DC within an organism, namely in the gonads [13–16] and across somatic tissues [12,17–19], we lack a general understanding of why we observe such variation. Investigating DC at the cell-type level may clarify variation seen in bulk RNA-seq studies.

Expression and dosage differences between the sexes have been extensively studied using bulk RNA-seq data. However, we still lack a general understanding of whether sex-biased gene expression and DC vary at the cell-type level. Single-cell RNA-seq data can improve the resolution at which we are able to detect sex differences. Cell-type clusters within a tissue can be annotated, and sex differences can be estimated by comparing equivalent cell populations. Darolti & Mank [20] produced a single-cell RNA-seq dataset from guppy somatic and reproductive tissues. They identified more sex-biased genes at the cell-type level than when comparing female and male whole tissues. This points to the importance of considering differences in cell-type abundance between the sexes. It also shows that non-isometric scaling in cell populations within each tissue, where cell-type cluster sizes do not scale proportionally across female and male development, can inflate the amount of spurious results. Therefore, bulk RNA-seq might provide an incomplete picture of evolutionary patterns in studies that investigate rates of evolution. This variation in sex-bias suggests that a finer resolution is needed to understand the evolution of context-dependent gene expression.

The molecular mechanism that establishes DC in *Drosophila* male cells is a complex of proteins and long non-coding RNA(lncRNAs). The male-specific-lethal (MSL) complex ensures that the transcription of genes in the male hemizygous X matches the level produced in females. The MSL complex consists of five protein subunits (MSL-1, MSL-2, MSL-3, *males absent on the first*—MOF and *maleless*—MLE), and two noncoding RNAs (roX1 and roX2). The complex, also referred to as *Drosophila* DC complex (DCC), moves along the X chromosome and results in histone H4K16 acetylation in males (via MOF [21]; reviewed in [22]). Understanding the role of the different components of the DCC is important for investigating how variation in DC can be achieved. In male germline cells, lack of DC has been associated with absence of MSL expression: MLE is the only component of the MSL complex that is expressed in these cells where acetylation of H4K16 is absent [23]. One study [19] found a link between average X-to-autosome expression and MSL-2 expression across tissues, suggesting that MSL-mediated DC is involved in tissue-specific variation in the extent of DC.

Here, we use single-nucleus data from the Fly Cell Atlas [24], annotate over 250 cell-types in *D. melanogaster* flies, and identify and curate a list of marker genes for each cell-type. The dataset includes whole head and body with replicate samples, as well as 15 other individual tissue samples of both females and males. This makes the Fly Cell Atlas suitable for investigating differences in gene expression between the sexes at the tissue- and at the cell-type level. Li *et al*. [24] investigated sex-biased expression by comparing all female to all male cells within each cell-type. In particular, the authors performed a differential expression analysis for each gene within a cell-type cluster. They found that cell-types were typically either highly female- or highly male-biased. However, detailed analyses of the set of genes that are sex-biased in different cell-types and tissues, how much overlap there is between them and between results obtained from bulk RNA-seq, and the relationship between cell- and tissue-level sex-bias have yet to be performed.

We first focus on sex differences in cell-type composition. Because sex-bias may be confined to specific cell-type populations, it may be masked when analyses are conducted at the whole-tissue level. We therefore hypothesized that sex-bias could be affected by sex differences in cell-type composition. Then, we examined sex differences in X-linked gene expression. While distinct sex differences in the head and body have been described, less is known about differences at the cell-type level. We explore whether the direction of sex-bias of X-linked genes is consistent across body parts and cell-types. To address these questions, we implement complementary approaches that take advantage of the Fly Cell Atlas' replicated experimental design, such as pseudo-bulk analyses, where expression data from cell-type clusters is aggregated across biological replicates. These types of analyses can show greater power to infer differential expression [25]. Finally, we investigate X-linked-to-autosomal gene expression ratios in females and males to determine how DC varies throughout the body. We then focus on the male accessory glands, where the most extreme differences in DC were surprisingly detected.

2. Methods

(a) Data filtering and normalization

We used the 10x genomics single-nucleus RNA-seq datasets available on the Fly Cell Atlas website (https://flycellatlas.org/), which include single-nucleus gene expression profiles for several tissues and body parts. We analysed RNA-seq counts from 15 individual tissue samples (antenna, body wall, fat body, gut, haltere, heart, leg, male reproductive glands, Malpighian tubule, oenocyte, ovary, proboscis and maxillary palps, testis, trachea and wing), of which three were sex-specific (ovary, testis, male reproductive glands). Whole head and carcass samples were also analysed. The carcass includes the thorax and abdomen but excludes reproductive tissues, as we detail below. We downloaded the samples corresponding to the *relaxed dataset*, which includes raw gene expression counts that were produced with the Fly Cell Atlas data processing pipeline for

all annotated cell-types. Then, we performed a series of quality filtering and normalization steps which we now describe. We also implemented an alternative normalization step to the 'counts per million' normalization provided in the stringent Fly Cell Atlas files. We have chosen sctransform (R package), which normalizes counts using a negative binomial regression model and has been shown to be well suited for zero-inflated single-cell datasets, characterized by an excess of gene expression entries with zero counts [26]. We started by filtering out cells found in somatic tissues present in both sexes that were annotated as belonging to sex-specific cell-types, i.e. from the ovary, the testis or the male accessory gland. Moreover, we filtered out cells annotated as 'artefact' from all samples. We then removed contamination from ambient RNA with DecontX (R package) [27]. Corrected counts were rounded using the base R function *round*(). Then, we filtered genes and cells using criteria identical to the Fly Cell Atlas' *stringent dataset* filtering criteria. These include removing cells expressing fewer than 200 genes and 500 total DecontX corrected gene counts, cells that show more than 5% mitochondrial content and cells whose median total counts and number of expressed genes are three times greater than the deviation from the sample medians. We also removed genes that were expressed in fewer than three cells. Finally, count normalization was performed using sctransform [26]. Here, the total counts per cell and the percentage of mitochondrial content were regressed out since these potential confounding sources of variation were already accounted for in previous quality control steps.

After removing cells annotated to the reproductive system and those annotated as artefact cells, we retained 40 650 female and 27 954 male cells for the carcass. Carcass and head cells are, therefore, solely somatic to ensure sex-bias inference was unaffected by the strongly sex-biased nature of the gonad. Similarly, after filtering, we retained 38 387 female and 37 713 male cells for the head. We also analysed another 12 individual somatic tissue samples (antenna, body wall, fat body, gut, haltere, heart, leg, Malpighian tubule, oenocyte, proboscis and maxillary palps, trachea and wing), for which replicate information was not available and therefore differential expression analysis could not be performed, but which could still be used to infer mean levels of sex-bias. Samples ranged from 2800 to 12 476 female and 3440 to 13 438 male cells.

Our study involved: (i) a pseudo-bulk differential expression analysis to infer sex differences in whole head and carcass; and (ii) computing sex-bias with and without accounting for differences in cell-type composition to directly quantify the contribution of said differences to sex-bias (electronic supplementary material, figure S1). We describe the analyses in detail below.

(b) Differential expression analysis

For head and carcass data, we performed a pseudo-bulk analysis where one sums gene expression counts across cells within a tissue or cell-type cluster. With this approach, we do not treat each cell as an independent sample, as expression in cells from the same biological sample is probably correlated. Therefore, we aggregated counts across the whole head and carcass—tissue-level analysis—and across cell-type clusters within each body part—cluster-level analysis. For that, we used the *aggregateAcrossCells*() function (scuttle R package) [28] to sum over cells within a category (tissue or cell-type cluster) and sample. Any gonad-specific cell-types were removed from whole carcass data to avoid sex-biased expression patterns owing to the extremely sex-biased nature of the ovary and testis. For the cluster-level analysis, we used the broader annotation which includes nine cell-type categories: neurons, sensory neurons, glial cells, tracheolar cells, haemocytes, fat cells, epithelial cells, oenocytes and muscle cells. We then performed a differential expression analysis to compare gene expression in females and males at the tissue- and cluster-level, using DESeq2 (R package) [29], which fits a generalized linear model where counts are modelled with a negative binomial distribution. A gene was characterized as sex-biased based on two criteria: (i) the gene's absolute $\log_2$fold change value is greater than or equal to 1 (i.e. expression levels in one sex are twice those in the other); and (ii) the corrected *p*-value is less than or equal to 0.05.

We investigated the distribution of sex-biased genes throughout the genome by calculating the ratio of observed to expected female- and male-biased genes for each chromosome. The expected number of sex-biased genes was obtained by computing the product of the proportion of genes in a given chromosome found in the dataset and the total number of female- and male-biased genes inferred from our differential expression analysis. Then, we calculated odds ratios to compare X-linked and autosomal sex-biased genes within a body part or cell-type cluster. With these odds ratios, we performed Fisher's exact tests of significance and calculated Benjamini–Hochberg corrected *p*-values.

To compare cell-type clusters within the male accessory glands, we performed a differential expression analysis using a method that was designed to account for zero-inflated single-cell RNA-seq entries. Model-based analysis of single-cell transcriptomics fits a generalized linear model and performs a likelihood-ratio test to compare two different groups [30]. We performed *p*-value false discovery rate correction (same as Benjamini–Hochberg correction; *p.adjust*() function from *stats* package) and considered a gene to be differentially expressed between two groups if it has an adjusted *p*-value < 0.01. Here, we do not impose a $\log_2$fold change threshold, but instead use a stricter *p*-value threshold. This allows for detecting subtler expression differences between clusters of main cells. For further details see the electronic supplementary material, Supplementary Methods.

(c) Estimating sex-bias at the tissue level

We calculated sex-bias at the tissue level for each gene to either (i) account for differences in cell-type composition between the sexes or (ii) assume that females and males have the same proportion of a given cell-type in a tissue. Sex-bias was computed for each somatic tissue as well as whole head and carcass using the following formulae:

$$\text{sex-bias}_{\text{no cell-type differences}} = \frac{\sum_{i=1}^{n} E_{\text{f},i} \times p_i + 10^{-9}}{\sum_{i=1}^{n} E_{\text{m},i} \times p_i + 10^{-9}},$$

where n corresponds to the total number of cell-type clusters within a tissue, i is the cluster index, $E_{\text{f},i}$ and $E_{\text{m},i}$ are the mean expression across female and male cells of type i, respectively, and p_i is the proportion of cells of type i which is defined as:

$$p_i = \frac{1}{2}(p_{\text{f},i} + p_{\text{m},i}),$$

where $p_{\text{f},i}$ and $p_{\text{m},i}$ correspond to the proportion of female and male cells of type i within a tissue, respectively. Here, we assume that females and males have the same number of cells per cell-type cluster ($n_{\text{f cells}} = n_{\text{m cells}}$). Sex-bias accounting for differences in cell-type composition is calculated as below:

$$\text{sex-bias}_{\text{cell-type composition}} = \frac{\sum_{i=1}^{n} E_{\text{f},i} \times p_{\text{f},i} + 10^{-9}}{\sum_{i=1}^{n} E_{\text{m},i} \times p_{\text{m},i} + 10^{-9}} \times \frac{n_{\text{f cells}}}{n_{\text{m cells}}}.$$

Here, we grouped cells using the Fly Cell Atlas' broad annotation and computed sex-bias using mean expression (sctransform normalized) counts for each cell-type cluster. To each proportion, we added 10^{-9} to avoid infinitely high numbers when calculating $\log_2$(sex-bias). This quantity is biologically relevant when calculating proportional changes, i.e. fold changes between females and males.

To assess the degree to which sex-biased cell-type composition contributes to differences in sex-biased expression, we fitted a standard linear regression ($\text{sex-bias}_{\text{no cell-type differences}} \sim \text{sex-bias}_{\text{cell-type composition}}$). Outliers were calculated using an outlier test that reports Bonferroni-corrected p-values with a 0.05 threshold (car R package [31]).

(d) Dosage compensation analysis

To investigate DC in this single-nucleus dataset, we calculated mean X-to-autosome expression ratios in whole heads and carcass samples, as well as across all 15 individual tissues described above. Here, genes were assigned to the X chromosome or to an autosome based on the complete *D. melanogaster* genome annotation Dmel 6.31 (FlyBase, June 2019). Then, we computed mean X-to-autosome expression (X:A mean expression) for each cell. We investigated X:A ratio of densities in female and male tissues or cell-type clusters within each tissue. In summary, we investigated cell-type clusters that are a part of the circulatory-respiratory system (haemocyte, cardial cell, tracheolar cell and excretory system clusters), the nervous system (neuron and sensory neuron clusters), the endocrine and immune system (glial cell, salivary gland, fat cell and oenocyte clusters), the reproductive system (germline cell, reproductive system and male accessory gland) and other somatic cell clusters (muscle cells, epithelial cells, antimicrobial epithelial cells and somatic precursor cells). Female and male densities were compared using non-parametric Kolmogorov–Smirnov tests (*ks.test*() R function). p-values were corrected for multiple comparisons using a Bonferroni correction.

To further investigate DC in the male accessory gland, we performed immunofluorescence staining of a histone modification—H4K16ac—that has been shown to mark the X chromosome to be upregulated. We dissected and stained male accessory glands and testis. Sample acquisition and immunostaining are described in the electronic supplementary material, Supplementary Methods.

3. Results

(a) Cluster-specific expression contributes strongly to sex-bias, but cell composition does not

(i) More sex-biased genes at the cluster level when compared to bulk

Heterogeneity in sex-bias between different cell-types might influence differential expression analyses at the tissue level. We hypothesized that, should most sex-bias be driven by sex differences that are limited to some cell-type clusters, it should be more easily detected at the cluster than at the tissue level. To investigate sex-bias in the Fly Cell Atlas, we performed differential expression analyses using data derived from whole head and carcass, for which several replicates were available (six female and six male replicates). Our analysis found 77 and 1985 (0.6% and 13.4% of all genes) differentially expressed genes between the sexes at the tissue-level in head and carcass, respectively (electronic supplementary material, table S1(A)). By contrast, we found 207 and 2908 (1.7% and 19.6%, respectively) different sex-biased genes at the cluster-level (electronic supplementary material, table S1(B)). The results are in line with our hypothesis: 64% (71 out of 110) and 59% (792 out of 1334) of female-biased genes in the head and carcass, respectively, were restricted to the cluster-level analysis (figure 1). The remaining female-biased genes were shared between the tissue- and cluster-level analyses. For male-biased genes, the patterns differed between head and carcass: 84% (82 out of 97) of said genes in the head were only found to be male-biased at the cluster level. In carcass samples, 50% (790 out of 1575) of male-biased genes were limited to the cluster-level analysis. Overall, only 26 genes were shared between body parts in both the tissue- and the cluster-levels analyses (electronic supplementary material, table S2). Relaxing threshold stringency—specifically, removing the $\log_2$fold change value cut-off—reveals a meaningful increase in shared sex-biased genes across tissue- and cluster-level analyses (electronic supplementary material, figure S2). Nevertheless,

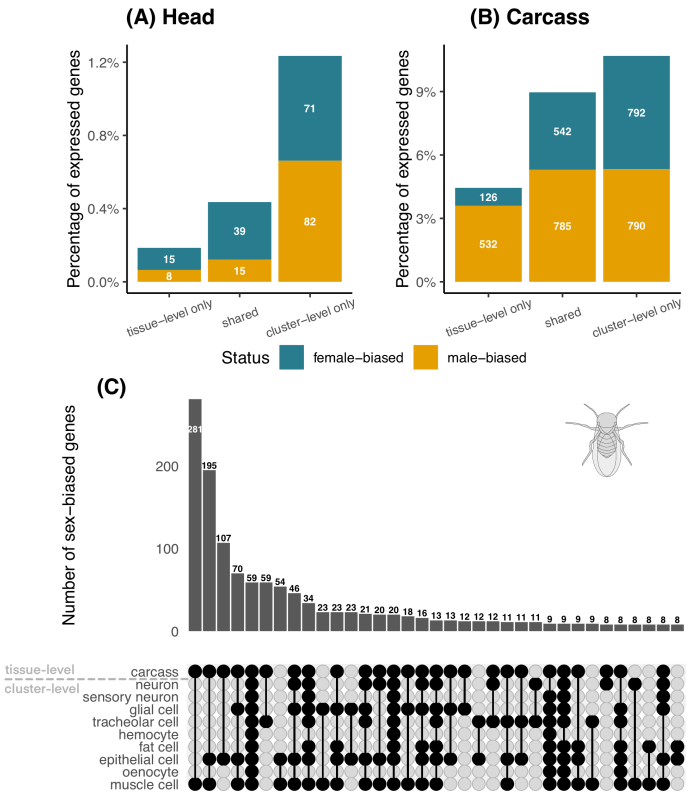

Figure 1. Sex-biased genes are largely shared between the whole head (or carcass) and cell-type clusters within each of the two body parts. (A) Percentage of female- (in blue) and male-biased (in yellow) genes for whole head (*tissue-level only*), all cell-type clusters combined (*cluster-level only*) and shared between whole head and at least one cell-type cluster (*shared*). Each bar plot also includes the total number of sex-biased genes in each category. (B) Same as (A) but for the carcass. In the carcass, 79 genes were included twice as they are both female- and male-biased in at least one cell-type cluster or whole tissue. (C) Upset plot that shows the total number of sex-biased genes shared across the carcass and cell-type clusters within the carcass. Intersections under eight genes were excluded for simplicity.

many sex-biased genes are still detected only at the cluster-level. This highlights the importance of choosing appropriate significance criteria to account for the inherent variability of RNA-seq data, especially with single-cell datasets. These patterns suggest that bulk analyses of whole body parts have the potential to overlook sex-bias that is only detectable at the cluster level.

Moreover, we observe a greater proportion of sex-biased genes in the head only at the cluster-level when compared to the carcass. This suggests that many sex differences in the head may only be detectable at the resolution of individual cell-type clusters. Accordingly, we emphasize the value of single-nucleus RNA-seq studies for investigating sex differences in the head in particular.

In light of these differences in power to detect sex-bias, we asked whether patterns observed in our differential expression analysis are consistent with those reported in studies of bulk RNA-seq. Previous studies have found evidence of greater sex-bias in the carcass than in the head [32] and an excess of male-biased compared to female-biased genes [33,34]. Our differential expression analysis supports both observations, but the excess of male-biased genes is only observed in the carcass (figure 1). In the head, there are more female-biased than male-biased genes (figure 1A) similar to [35]. These patterns are recovered both at the tissue- and at the cluster-level (electronic supplementary material, table S1). We observe that approximately 52% and 46% of sex-biased genes that are shared across tissue- and cluster-level analyses in carcass and head, respectively, have sex-biased status in only one or two cell-type clusters (figure 1C; electronic supplementary material, figure S3). These results suggest that, despite the different numbers of differentially expressed genes that are detected, broad patterns are consistent with studies that used bulk RNA-seq.

Finally, we observed that some genes are sex-biased at the tissue-level but not at the cluster-level. That is the case for 23 genes in the head (10.0% of sex-biased genes in the head; figure 1A) and 656 genes in the carcass (18.4% of sex-biased genes in the carcass; figure 1B). To understand why this many genes are not recovered by the cluster-level analysis, we examined their expression levels per cell: the majority of these genes are expressed at low levels and in few cells. In the carcass, 28 out of 656 genes were reported as unbiased with a $\log_2$fold change that fell below the significance threshold ($-1 < \log_2\text{fold change} < 1$) but a significant p-value (<0.05) in at least one cell-type cluster. Such subtle sex differences may be missed by the cluster- but not tissue-level analysis, as pooling all cells in a tissue increases the power to detect them. On the other hand, many of these tissue-only differences affect genes that are lowly and sparsely expressed and may reflect a high rate of false positives for such low-expression genes in bulk analyses.

(ii) Sex-bias is mostly unaffected by differences in cell-type composition

Thus far, we have shown that cluster-level analyses can detect sex-bias that is cell-type-specific, which is indiscernible at the tissue level. Next, we considered the impact of sex differences in cell-type composition on sex-bias estimates. Specifically, we compared the extent of sex-bias for each gene when accounting for or not accounting for cell-type composition differences. We

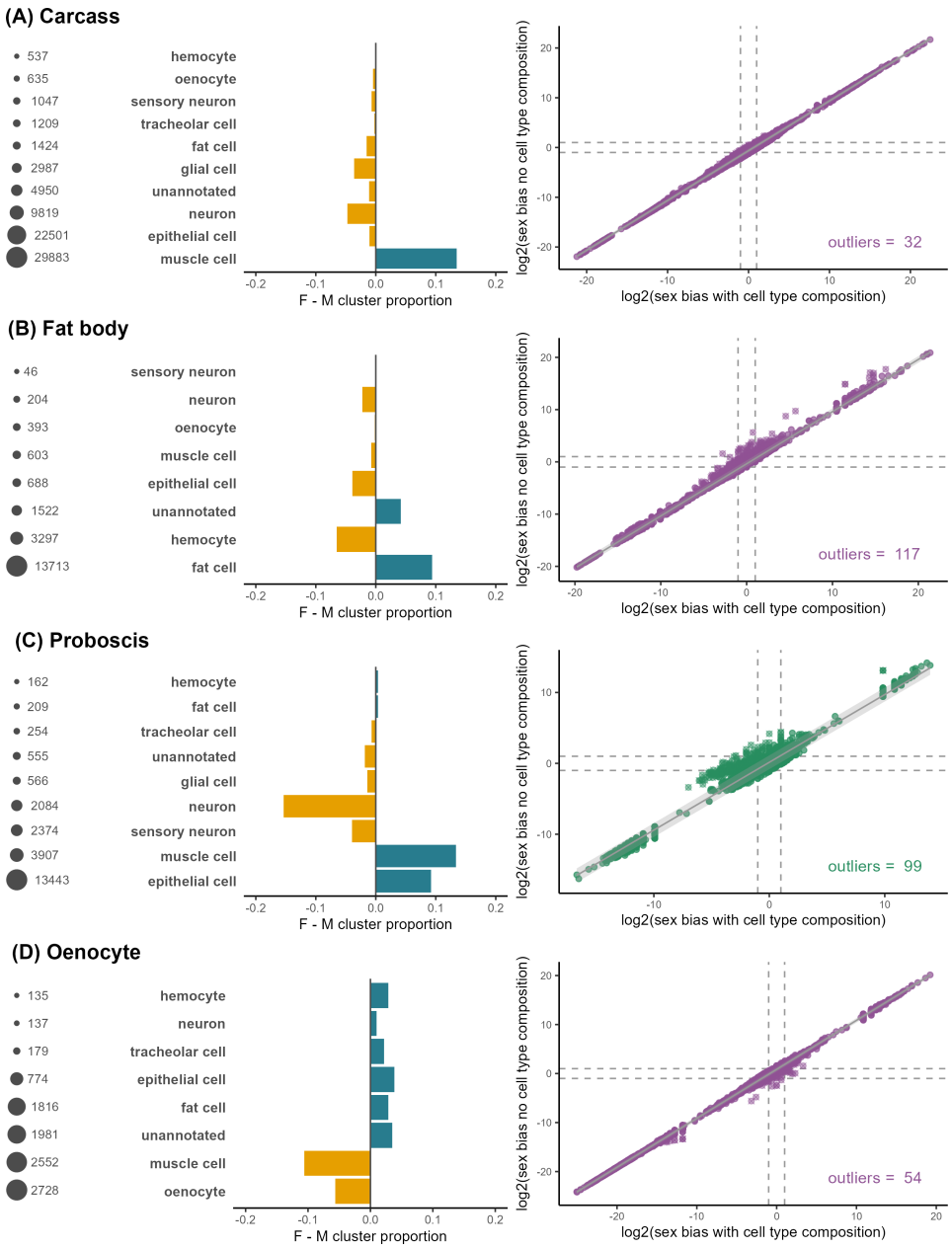

Figure 2. Sex-bias with and without differences in cell-type composition between females and males is highly correlated in the (A) carcass, (B) fat body, (C) proboscis and maxillary palps and (D) oenocyte, regardless of sex differences in cell-type proportions. Total cell-type cluster size (left-hand side), female-to-male difference in cluster proportion (centre) and sex-bias accounting for cell-type composition versus without cell-type composition (right-hand side) are shown for each tissue. Head and carcass tissues are coloured in green and purple, respectively.

found that the two are highly correlated both in whole body parts (head and carcass) and in individual tissues ($\underline{\beta} = 1$, $\underline{r^2} = 0.996$; figure 2; electronic supplementary material, figures S4–S6), suggesting at most a minor contribution of cell-type composition differences to sex-bias. This seems to be owing to the fact that sex differences in the proportion of each cell-type cluster are themselves modest (average absolute difference approx. 3.3%). The largest absolute sex difference, 21.5%, is found in excretory system cells within the Malpighian tubule, which show a male-biased proportion (electronic supplementary material, figure S5). Even for tissues where the largest clusters were also the most sex-biased in proportion, the two sex-bias measures were strongly correlated (figure 2). This indicates that sex differences in cell-type composition might generate inappreciable changes to sex-biased expression.

To further confirm that the sex-bias detected through differential expression was not influenced by heterogeneity in cell-type composition, we compared the ratio of female-to-male-biased gene ratios (no. female-biased/no. male-biased genes) with female-to-male cells ($\log_2$(female cells : male cells)) per tissue or cell-type cluster (electronic supplementary material, figure S7). We predict sex-biased gene expression not to scale with sex-biased cell count if composition is not a major driver of bias. We fitted a linear model to assess the relationship between no. female-biased/no. male-biased genes and sex-bias in cell-type abundance ($\log_2$(female : male cells)). Consistent with our prediction, there is no significant linear relationship between the two variables (head: adjusted r^2 approx. 0.05, p-value approx. 0.277, slope approx. −0.592; carcass: r^2 0.013, p-value 0.315, slope approx. −0.35), again suggesting that cell-type abundance differences between the sexes do not explain differences in the number of genes found to be sex-biased. If anything, some clusters with the highest number of sex-biased genes have more male bias. For instance, muscle cells in the carcass are more than two times more abundant in females but show the highest

number of male-biased genes among individual clusters of cells. These observations further confirm that differences in cell-type composition are unlikely to strongly affect differential expression analysis at the tissue level.

(iii) Sex bias of X-linked genes: female-biased genes always enriched, male-biased genes deficient in the carcass

Previous studies have identified an enrichment of X-linked sex-biased genes in the head [4], in particular of male-biased genes [12]. This enrichment was largest in the brain when compared to other body parts and gonads [12,35]. We expected to recover this pattern at the single-nucleus level, and our results align with earlier findings, showing an enrichment of X-linked sex-biased genes in the head (electronic supplementary material, figure S8(A) and table S3). Notably, X-linked male-biased genes are significantly enriched in neurons (Fisher's exact test p-value = 0.019). By contrast, we observed a deficit of X-linked male-biased genes in the carcass, spanning most cell-type clusters (8 out of 10; electronic supplementary material, table S4), suggesting that the lack of X-linked male bias is widespread among somatic cells, not limited to the male gonads. Additionally, the enrichment of X-linked female bias in the head is driven by specific cell-type clusters—neurons (Fisher's exact test p-value ≈ 0.01), sensory neurons (Fisher's exact test p-value ≈ 0.005) and glial cells (Fisher's exact test p-value ≈ 0.003)—whereas in the carcass, it is mainly observed in epithelial cells (Fisher's exact test p-value < 0.001).

(b) Dosage compensation varies across tissues and cell-types

(i) Tissue-dependent XA expression suggests that incomplete dosage compensation and overcompensation coexist

To assess whether DC varies across the body, we calculated the average X:A expression ratios for each cell. Then, we examined the distributions of X:A mean expression for females and males. There was substantial variation at the tissue level, suggesting that the overall expression of X-linked genes, as well as the extent of DC, is heterogeneous throughout the body. The female and male density curves of X:A expression in whole head and carcass were significantly different (Kolmogorov–Smirnov test p-value < 0.001), with female distributions centred around 1, the expectation under complete dosage balance, and male cells X:A distribution centred slightly above 1 (figure 3; 'head' and 'carcass'). This might indicate that the majority of male cells in somatic tissues show overexpression of X genes to the point where, on average, X-linked expression is higher than autosomal gene expression. This trend is however not consistent across individual tissues, where only male antenna and wing peak above 1 (figure 3). Half of male tissues (body wall, fat body, gut, heart, leg, oenocyte, male accessory gland and testis) show X:A ratio distributions centred below 1, with a subset also having significantly lower X:A than females, suggesting incomplete DC. In addition, almost half of female tissues show X:A densities that peak just below 1. This result might suggest that not all X-linked genes require high expression levels at all times in females. Alternatively, expression balancing in these tissues may involve a down-regulation of the X, similar to the mechanisms of DC observed in many clades (e.g. nematodes and Lepidoptera).

(ii) Sex differences in X:A expression differ between cell-types in somatic tissues

Females and males are significantly different in X:A expression in the majority of somatic tissues (figure 3; electronic supplementary material, figure S9). These differences can either be homogeneous across the tissue or be cell-type-dependent. We therefore investigated X:A expression ratios in different cell-types for all somatic tissues (electronic supplementary material, figures S10–S13). In nine (out of 12) tissues, significant differences in X:A expression occur in less than 60% of the cell-type clusters, i.e. between two and four cell-types.

Interestingly, the cell-type with the most differences between the sexes in X:A expression is epithelial cells, where X:A distributions are significantly different in 8 out of 10 tissues. Other tissue-specific cell-types, including antimicrobial epithelial cells, excretory system cells, salivary gland cells and tracheolar cells, which are found in only one tissue (gut, Malpighian tubule and trachea, respectively), also show significant differences between the sexes. Additionally, somatic precursor cells in the excretory system, also referred to as somatic stem cells, show significant differences in X:A expression between females and males (gut p-value = 1×10^{-4}; Malpighian tubule p-value = 4.61×10^{-4}). These results suggest that X:A expression differences also occur in specialized cell-types.

(iii) X:A expression in the reproductive tissues shows distinct patterns between females and males

In the ovaries, X:A expression is centred around 1 (figure 3A; 'gonad'). Testis, on the other hand, shows a nearly bimodal distribution with one peak corresponding to germline cells (incomplete DC) and another around 1 (figure 3B). This lack of DC in the male germline is well described, namely at the single-cell level [13,16,36]. These differences between male germline and testis somatic cells are clear when comparing X:A distributions at a finer resolution (electronic supplementary material, figure S14). Surprisingly, we also observe a clear bimodal distribution of X:A expression in the male reproductive glands that is shifted towards much lower values with peaks around 0.8 and 0.6 (figure 3A). This result points to a possible lack of DC within a subset of male reproductive gland cells.

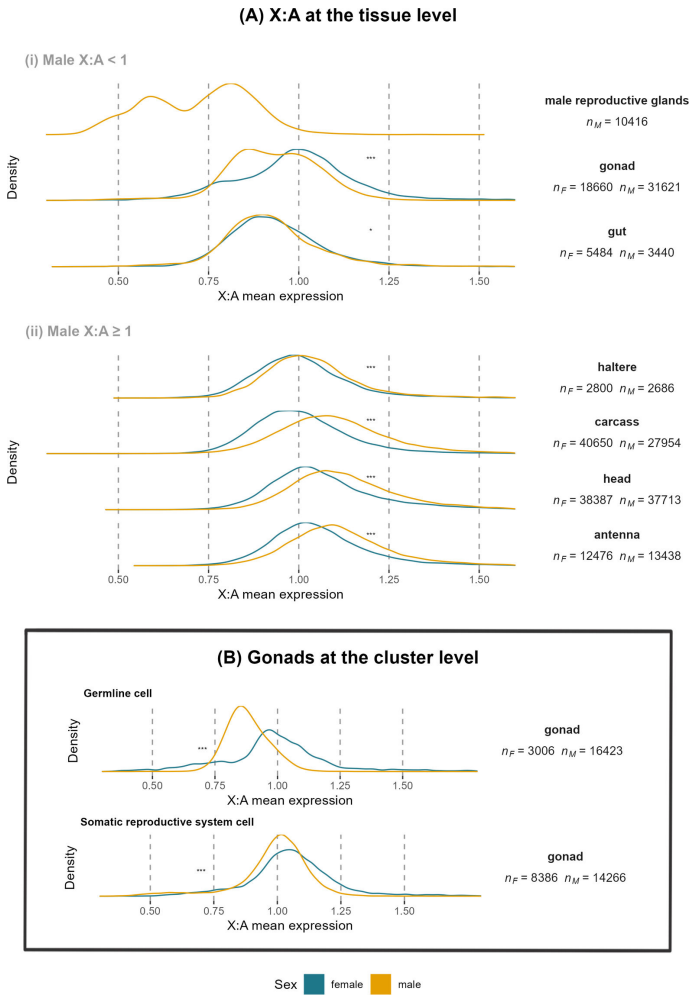

Figure 3. The distribution of average X:A expression ratios varies between females and males. Female (blue) and male (yellow) distribution profiles are plotted separately for (A) five individual tissues plus head and carcass, and (B) for germline and somatic cells in the gonads. Panel (A) is divided into male X:A distributions (i) that exhibit a maximum below 1, and (ii) those whose maximum is greater than or equal to 1. The total number of cells in each distribution is shown on the right-hand side. Gonad refers to ovary and testis. Kolmogorov–Smirnov tests compare female and male X:A expression distributions. Significance levels: *p-value < 0.05; ***p-value < 0.001.

(iv) Main accessory gland cell clusters have reduced roX1 expression

To further investigate the apparent lack of DC in a subset of cells of the male accessory glands, we examined five cell-types: main cells, secondary cells, anterior ejaculatory duct, ejaculatory bulb and ejaculatory bulb epithelium. Main cells are the predominant secreting cell-type in the accessory gland epithelium. By contrast, secondary cells are interspersed among main cells in the apical portion of the gland [37,38]. None of the five cell-types shows completely balanced X:A expression, with median ratios ranging from 0.63 to 0.89 (figure 4). The lowest X:A expression is found among main and secondary cells. In main cells, we observe a marked bimodal distribution similar to that described in figure 3. To identify the cells that show lack of DC, we added X:A mean expression onto a UMAP plot of the male reproductive glands by colouring each cell according to their respective X:A value (figure 4B). Two distinct clusters of main cells were described in an earlier single nucleus RNA-seq dataset of accessory gland [39]. Similarly, cells annotated as 'main cells'—i.e. those expressing main cell marker genes—formed two separate clusters in the Fly Cell Atlas' UMAP based on their remaining gene expression profiles. We observe a decisive split in the degree of DC between these two main cell clusters. The central cluster corresponds to the lowest peak in the bimodal distribution, the first observation of lack of DC in a somatic tissue in *D. melanogaster*.

To assess whether expression of MSL complex genes could explain differences in X:A ratios, we compared expression of lncRNA : roX1 and lncRNA : roX2 in high DC and low DC main cells. Low X:A ratio cells show significantly lower lncRNA : roX1 and lncRNA : roX2 expression when compared to high X:A ratio cells (figure 4C). The expression of other MSL complex genes—*msl-1*, *msl-2*, *msl-3*, *mle* and *mof*—is very low across all accessory gland cells, such that no comparisons across cell-types could be performed (electronic supplementary material, figure S15). To assess whether low DC main cells could be the result of germline contamination, we determined the expression of known spermatogonia, spermatocyte and spermatid cell-type markers across male accessory gland cell clusters. We confirmed the absence of germline marker expression in the accessory glands, particularly in main cells (electronic supplementary material, figure S16), indicating that these cells are not the result of contamination.

We hypothesized that these patterns could be caused by differences in expression of upstream sex determination pathway genes. We started by investigating *dsx*—the gene responsible for the female and male developmental programs in *Drosophila*.

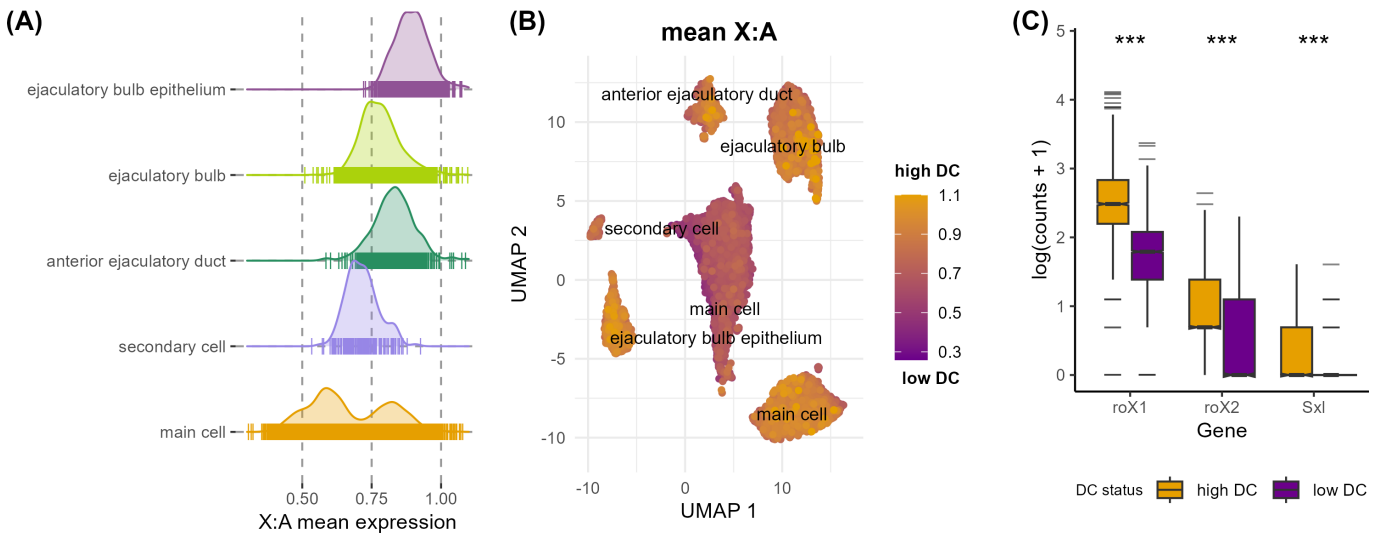

Figure 4. Average X:autosomal gene expression across cell-types in the male accessory glands shows heterogeneity across cell-types that could be related to dosage compensation. (A) Densities of X:A expression for ejaculatory bulb epithelium, ejaculatory bulb, anterior ejaculatory duct, secondary and main cells. Observations are marked as dashes at the bottom of each plot. (B) UMAP plot as a visual representation of the different cell-types within the tissue. Each cell is coloured by X:A expression ratio values (purple: lowest; yellow: highest). (C) Distribution of Sxl, roX1 and roX2 in main cells with high (yellow) and low (purple) X: A expression. Wilcoxon tests compare high and low DC clusters. Significance level: ***p-value < 0.001.

We observe that *dsx* expression is very low ($\text{median}_{dsx} = 0$) among main cells (similarly to most somatic tissues in the electronic supplementary material, figure S17, where $\text{median}_{dsx} = 0$), and that there is higher expression of *dsx* in all other male accessory gland cells ($\text{median}_{dsx} = 2$; electronic supplementary material, figure S15). Finally, we compared the expression of Sex lethal (*Sxl*), transformer (*tra*), transformer 2 (*tra2*) and fruitless (*fru*) across accessory gland cell-types. We found that all cell-types show low *tra*, *tra2* and *fru* expression, and that low X:A ratio cells show lower *Sxl* expression when compared to high X:A ratio main cells, secondary cells and anterior ejaculatory duct cells (figure 4; electronic supplementary material, figure S18).

To corroborate this result, we took advantage of a recent study by Majane *et al*. [39], who generated a single-cell RNA-seq dataset for the accessory gland and ejaculatory duct in three *Drosophila* species: *D. melanogaster*, *Drosophila simulans* and *Drosophila yakuba*. Majane *et al*. [39] detected the two main cell clusters as separate cell-types in the three species, but did not report a difference in DC. After calculating the X:A average expression for each cell in this dataset, we recovered a reduction in X:A expression in main cell cluster 2 compared with main cell cluster 1 and other cell-types. This difference is consistent in all three species (electronic supplementary material, figures S19 and S2), and again associated with reduced expression of lncRNA : roX1 and roX2 in main cell cluster 2 (electronic supplementary material, figure S21), suggesting this is not an artefact of our data.

Finally, we used immunofluorescence imaging to visualize DC in cell nuclei of the accessory gland and to investigate the possibility that the two clusters of main cells could be the result of differential DC between nuclei within the same cell. Each main accessory gland cell has been observed to be binucleated [37,38]. This process has been described as a cell cycle arrest before eclosion [40] that allows for larger amounts of seminal fluid to be ejected [41]. Since the Fly Cell Atlas project sequenced individual nuclei, rather than individual cells, we hypothesized that one nucleus is dosage compensated, while the other lacks the H4K16ac modification, which marks the X chromosome to be upregulated. To test this hypothesis, we performed an immunofluorescence staining protocol on dissected accessory glands (electronic supplementary material, text S1). We could assess the presence of H4K16ac at each nucleus within each cell. Overall, approximately 70% of sampled nuclei (272 out of 387) showed the presence of H4K16ac (red) staining (electronic supplementary material, figure S22(A)). We observed that cells which show H4K16ac do so in both nuclei. As a control, we also assessed the presence of H4K16ac in the testis. Here, we expected developing germline cells to lack H4K16ac as these cells do not require DC. Accordingly, we only observed the presence of H4K16ac in cells at the periphery of the lumen, probably corresponding to somatic cells that support the development of sperm cells (electronic supplementary material, figure S22(B)). In summary, our results show that, while some cells might not exhibit the presence of H4K16ac, those that do, do so across the two nuclei in the cell.

4. Discussion

(a) Comparative analysis of single-nucleus and bulk RNA-seq reveals both shared and unique features

We investigated sex differences in gene expression in the Fly Cell Atlas [24] single-nucleus RNA-seq dataset. Using a pseudo-bulk approach, we found that sex-bias varies at the cell-type level, having identified more sex-biased genes at the cluster- than at the tissue-level in both head and carcass. This is consistent with Darolti & Mank [20] who observed a similar trend in the skin, heart and gonads of guppies and generally supports the idea that cluster-specific bias may be relevant and has been understudied in bulk studies. On the other hand, we found that sex-bias calculated while accounting for differences in

cell-type composition is consistent with sex-bias without correcting for said differences. This is in contrast to Darolti & Mank's [20] findings, who found that allometric scaling differences contributed to the observed patterns of sex-biased expression in tissue- and cluster-level analyses. Our contrasting results might be explained by reduced sex differences in proportion for most cell-type clusters. In fact, we observe that even in the tissues where the most sex-biased clusters are also the largest, sex-bias accounting for differences in cell-type composition is highly correlated to sex-bias without sex differences in cluster proportions. This suggests that, despite sexual dimorphism in body size, proportions of cell-type clusters might display isometric scaling between female and male *D. melanogaster* and not contribute to increased sex-biased gene expression measures.

Encouragingly, the general patterns we observed were largely consistent with previous bulk RNA-seq studies: there is (i) greater sex-bias in the carcass when compared to the head [42]; (ii) an enrichment of X-linked sex-biased genes in the head ([12,32] and [35]); (iii) and a deficit of male-biased genes on the X chromosome previously described by [4,43–47]. On the other hand, our findings on DC reveal substantial variation across tissues and cell-types. While DC has often been described as relatively consistent throughout the body—aside from known exceptions such as the germline [14,48]—other studies have noted context-dependent differences. Our results add to this body of work by highlighting particularly pronounced variation, especially at the cell-type level. These findings suggest that the extent of X:A expression differences might be underestimated in bulk RNA-seq studies of DC. Further studies of DC at the cell-type level will be needed to investigate whether similar variation is possible in species where other molecular mechanisms of DC have evolved.

An advantage to analysing single-nucleus (or single-cell) RNA-seq is that one can investigate this heterogeneity in cellular transcriptomes. Nevertheless, single-nucleus expression data are accompanied by certain caveats related to increased noise and zero-inflated data entries. This captures variability associated with RNA transcription and degradation processes as well as technical sampling error during library preparation. To overcome these limitations, we have chosen a method for count normalization that accounts for overdispersion while removing effects of technical variation on sequencing depth (sctransform [26]). Moreover, we benefited from the replicated experimental design of the Fly Cell Atlas to overcome the low statistical power that is associated with highly variable datasets. Here, we chose a method to infer sex-bias that accounts for potential replicate effects across genes (DESeq2 [29]). While these pseudo-bulk methods are powerful, non-random noise associated with the different replicates may still generate spurious results. Therefore, the use of appropriate normalization methods coupled with stringent significance thresholds is key to ensuring the reliability of the results. Furthermore, well-annotated cell-type clusters are required to draw meaningful biological conclusions. Such information is more readily available for model organisms and their closely related species. Our analyses benefited from the well-characterized gene expression profiles of *D. melanogaster* cells, contributing to the robustness of our study. Overall, the key features of our single-nucleus study enabled us to identify sex-biased expression patterns that align with findings from bulk RNA-seq studies. In addition, we uncovered sex differences that are exclusively detectable at the cluster-level, providing insights only observable at a finer resolution.

(b) Dosage compensation modulates sex-bias in a cell-type-specific manner

Epithelial and muscle cells emerge as the most sex-biased cell-type clusters in the carcass, suggesting that these cell-types might contribute the most to sexual dimorphism at the organismal level. These patterns are consistent with phenotypic studies, where sex differences in epithelial immune response, namely in the gut and in the genital tract [49], and in abdominal pigmentation patterns [50] have been described. In the head, we found that nervous system cells—neurons, sensory neurons and glial cells—rely on few male-biased genes to generate observable differences in behaviour, primarily those associated with male-specific traits such as courtship song (reviewed in [1]). This occurs alongside genes with more localized expression, such as *fru*, expressed in approximately 2% of male neurons [51], where the male splice-form functions as a key switch for courtship behaviour [52]. Interestingly, despite a generally higher X:A expression in males, there is a marked deficit of X-linked male-biased genes in the carcass. This suggests that DC mechanisms may not universally drive male-biased expression. By contrast, this deficit is not observed in the head, where DC appears to facilitate male-biased expression. Together, these findings highlight tissue- and cell-type-specific differences in the impact of DC on the expression and evolution of sex-biased genes. DC may, therefore, provide a permissive environment that enables further upregulation in males in some contexts, while representing an upper limit to male-biased expression in others.

(c) Enhanced resolution at the cluster level highlights dosage compensation heterogeneity

As discussed above, DC varies markedly across various cell-types. In particular, our analysis revealed a surprising reduction of X:A expression in a population of main cells in the male accessory glands. Low DC main cells also showed lower expression of Sxl, a key sex determination gene, and roX1 and roX2, two long non-coding RNAs associated with the MSL complex. The presence of two clusters of main cells with different X:A expression may be linked to 54 protein-coding genes that are differentially expressed between them. These genes are enriched in processes related to cell development and RNA transcription and splicing (electronic supplementary material, table S5). This suggests that low DC main cells may be arresting differentiation where reduced X-linked expression is associated with the deficit of accessory gland-specific genes on the X chromosome [4,47,53]. This process is probably associated with RNA regulation. Additionally, five long non-coding RNAs—roX1, CR44909, CR45097, CR43654 and CR43146—were differentially expressed between the two main cell clusters. All differentially expressed genes showed reduced expression in low X:A expression cells. Lastly, low X:A expression in these main cell clusters is associated with increased autosomal expression when compared to the expression of X-linked genes (electronic supplementary material, figure S23). Recently, reduced roX expression was associated with reduced H4K16ac on the X chromosome and

H3K27me3 on autosomes in *Drosophila* [54]. This resulted in reduced X-linked expression and increased autosomal gene levels. This increase in autosomal expression at low levels of roX gene expression could also contribute to the decreased X:A ratio, especially as accessory gland-specific genes are depleted from the X chromosome.

The insights obtained from our cluster-level analyses are examples of how sex-specific regulation of X-linked and autosomal expression is observable at a finer resolution. This could inform future studies of sex-bias on the importance of context-dependent gene expression.

5. Conclusion

Our study contributes to the discussion on how sex-biased expression varies across the organism. Future research could investigate if the sex differences in expression at the cell-type level we report here contribute to functional differences between the sexes. Phenotypic studies that examine female and male differences in physiological traits and correlate these variations with sex-biased expression could provide valuable insights. In addition, further research is required to understand how context-dependent DC of the male X chromosome is achieved. For example, single-cell profiling of the histone modification landscape, especially for H4K16ac, at the single-cell level might help elucidate the role of the MSL complex in this process. This will further our understanding of the evolution of sexual dimorphism constrained by a genome that is shared between the sexes.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

Data accessibility. No data was generated throughout this project. The RNA-seq data used in this analysis was obtained from [24] and [39]. All scripts used for this analysis can be found in Zenodo [55].

Supplementary material is available online [56].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. C.B.: conceptualization, data curation, formal analysis, investigation, methodology, visualization, writing—original draft; B.V.: conceptualization, methodology, supervision, writing—original draft.

Both authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This work was partly funded by an Austrian Science Foundation FWF ESPRIT fellowship (10.55776/ESP6331524) to C.B.

Acknowledgements. We would like to thank the Vicoso group for their invaluable input and discussions throughout this work. We thank Filip Ruzicka for his insightful comments on the manuscript. All computational resources were provided by the Scientific Computing Unit at ISTA. This research was also supported through resources provided by the Imaging & Optics Facility (IOF) at ISTA.

References

1. Greenspan RJ, Ferveur JF. 2000 Courtship in *Drosophila*. *Annu. Rev. Genet.* **34**, 205–232. (doi:10.1146/annurev.genet.34.1.205)
2. Parsch J, Ellegren H. 2013 The evolutionary causes and consequences of sex-biased gene expression. *Nat. Rev. Genet.* **14**, 83–87. (doi:10.1038/nrg3376)
3. Mank JE. 2017 The transcriptional architecture of phenotypic dimorphism. *Nat. Ecol. Evol.* **1**, 1–7. (doi:10.1038/s41559-016-000610.1038/s41559-016-0006)
4. Meisel RP, Malone JH, Clark AG. 2012 Disentangling the relationship between sex-biased gene expression and X-linkage. *Genome Res.* **22**, 1255–1265. (doi:10.1101/gr.132100.111)
5. Cutter AD. 2023 Sexual conflict, heterochrony and tissue specificity as evolutionary problems of adaptive plasticity in development. *Proc. R. Soc. B* **290**, 20231854. (doi:10.1098/rspb.2023.1854)
6. Ingleby FC, Flis I, Morrow EH. 2014 Sex-biased gene expression and sexual conflict throughout development. *Cold Spring Harb. Perspect. Biol.* **7**, a017632. (doi:10.1101/cshperspect.a017632)
7. Rodríguez-Montes L, Ovchinnikova S, Yuan X, Studer T, Sarropoulos I, Anders S, Kaessmann H, Cardoso-Moreira M. 2023 Sex-biased gene expression across mammalian organ development and evolution. *Science* **382**, eadf1046. (doi:10.1126/science.adf1046)
8. Ayroles JF *et al*. 2009 Systems genetics of complex traits in *Drosophila melanogaster*. *Nat. Genet.* **41**, 299–307. (doi:10.1038/ng.332)
9. Veltsos P, Fang Y, Cossins AR, Snook RR, Ritchie MG. 2017 Mating system manipulation and the evolution of sex-biased gene expression in *Drosophila*. *Nat. Commun.* **8**, s41467–017. (doi:10.1038/s41467-017-02232-6)
10. Zhang Y, Sturgill D, Parisi M, Kumar S, Oliver B. 2007 Constraint and turnover in sex-biased gene expression in the genus *Drosophila*. *Nature* **450**, 233–237. (doi:10.1038/nature06323)
11. Whittle CA, Kulkarni A, Extavour CG. 2021 Evolutionary dynamics of sex-biased genes expressed in cricket brains and gonads. *J. Evol. Biol.* **34**, 1188–1211. (doi:10.1111/jeb.13889)
12. Huylmans AK, Parsch J. 2015 Variation in the X: autosome distribution of male-biased genes among *Drosophila melanogaster* tissues and its relationship with dosage compensation. *Genome Biol. Evol.* **7**, 1960–1971. (doi:10.1093/gbe/evv117)
13. Meiklejohn CD, Landeen EL, Cook JM, Kingan SB, Presgraves DC. 2011 Sex chromosome-specific regulation in the *Drosophila* male germline but little evidence for chromosomal dosage compensation or meiotic inactivation. *PLoS Biol.* **9**, e1001126. (doi:10.1371/journal.pbio.1001126)
14. Argyridou E, Parsch J. 2018 Regulation of the X chromosome in the germline and soma of *Drosophila melanogaster* males. *Genes* **9**, 242. (doi:10.3390/genes9050242)
15. Rayner JG, Hitchcock TJ, Bailey NW. 2021 Variable dosage compensation is associated with female consequences of an X-linked, male-beneficial mutation. *Proc. R. Soc. B* **288**, 20210355. (doi:10.1098/rspb.2021.0355)
16. Witt E, Shao Z, Hu C, Krause HM, Zhao L. 2021 Single-cell RNA-sequencing reveals pre-meiotic X-chromosome dosage compensation in *Drosophila* testis. *PLoS Genet.* **17**, e1009728. (doi:10.1371/journal.pgen.1009728)
17. Julien P, Brawand D, Soumillon M, Necsulea A, Liechti A, Schütz F, Daish T, Grützner F, Kaessmann H. 2012 Mechanisms and evolutionary patterns of mammalian and avian dosage compensation. *PLoS Biol.* **10**, e1001328. (doi:10.1371/journal.pbio.1001328)

18. Uebbing S, Künstner A, Mäkinen H. 2013 Transcriptome sequencing reveals the character of incomplete dosage compensation across multiple tissues in flycatchers. *Genome Biol. Evol.* **5**, 1555–1566. (doi:10.1093/gbe/evt11410.1093/gbe/evt114)
19. Vensko SP, Stone EA. 2015 X-to-autosome expression and msl-2 transcript abundance correlate among *Drosophila melanogaster* somatic tissues. *PeerJ* **3**, 771. (doi:10.7717/peerj.77110.7717/peerj.771)
20. Darolti I, Mank JE. 2023 Sex-biased gene expression at single-cell resolution: cause and consequence of sexual dimorphism. *Evol. Biol* **7**, 148–156. (doi:10.1101/2022.11.08.515642)
21. Gelbart ME, Larschan E, Peng S, Park PJ, Kuroda MI. 2009 *Drosophila* MSL complex globally acetylates H4K16 on the male X chromosome for dosage compensation. *Nat. Struct. Mol. Biol.* **16**, 825–832. (doi:10.1038/nsmb.1644)
22. Lucchesi JC, Kuroda MI. 2015 Dosage compensation in *Drosophila*. *Cold Spring Harb. Perspect. Biol.* **7**, 1–21. (doi:10.1101/cshperspect.a01939810.1101/cshperspect.a019398)
23. Rastelli L, Kuroda MI. 1998 An analysis of maleless and histone H4 acetylation in *Drosophila melanogaster* spermatogenesis. *Mech. Dev.* **71**, 107–117. (doi:10.1016/s0925-4773(98)00009-410.1016/S0925-4773(98)00009-4)
24. Li H *et al*. 2022 Fly Cell Atlas: a single-cell transcriptomic atlas of the adult fruit fly. *Science* **375**, 1–12. (doi:10.1101/2021.07.04.451050)
25. Squair JW *et al*. 2021 Confronting false discoveries in single-cell differential expression. *Nat. Commun.* **12**, 12. (doi:10.1038/s41467-021-25960-2)
26. Choudhary S, Satija R. 2022 Comparison and evaluation of statistical error models for scRNA-seq. *Genome Biol.* **23**, 27. (doi:10.1186/s13059-021-02584-9)
27. Yang S, Corbett SE, Koga Y, Wang Z, Johnson WE, Yajima M, Campbell JD. 2020 Decontamination of ambient RNA in single-cell RNA-seq with DecontX. *Genome Biol* **21**, 1–15. (doi:10.1101/704015)
28. McCarthy DJ, Campbell KR, Lun ATL, Wills QF. 2017 Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* **33**, 1179–1186. (doi:10.1093/bioinformatics/btw777)
29. Love MI, Huber W, Anders S. 2014 Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 1–21. (doi:10.1101/002832)
30. Finak G *et al*. 2015 MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* **16**, s13059–015. (doi:10.1186/s13059-015-0844-5)
31. Fox J, Weisberg S, Price B. 2019 *An R companion to applied regression,* 3rd edn. Thousand Oaks, CA: SAGE Publications, Inc. See https://www.john-fox.ca/Companion/.
32. Catalán A, Hutter S, Parsch J. 2012 Population and sex differences in *Drosophila melanogaster* brain gene expression. *BMC Genom* **13**, 1–12. (doi:10.1186/1471-2164-13-654)
33. Hutter S, Saminadin-Peter SS, Stephan W, Parsch J. 2008 Gene expression variation in African and European populations of *Drosophila melanogaster*. *Genome Biol.* **9**, 12. (doi:10.1186/gb-2008-9-1-r1210.1186/gb-2008-9-1-r12)
34. Assis R, Zhou Q, Bachtrog D. 2012 Sex-biased transcriptome evolution in *Drosophila*. *Genome Biol. Evol.* **4**, 1189–1200. (doi:10.1093/gbe/evs09310.1093/gbe/evs093)
35. Khodursky S, Svetec N, Durki SM. 2020 The evolution of sex-biased gene expression in the *Drosophila* brain. *Genome Res.* **30**, 874–884. (doi:10.1101/gr.259069.11910.1101/gr.259069.119)
36. Mahadevaraju S. 2021 Dynamic sex chromosome expression in Drosophila male germ cells. *Nat. Commun* **12**, s41467–021. (doi:10.1101/2020.03.23.000356)
37. Bairati A. 1968 Structure and ultrastructure of the male reproductive system in *Drosophila melanogaster* Meig. *Monit. Zoöl. Ital. Ital. J. Zoöl.* **2**, 105–182. (doi:10.1080/00269786.1968.10736126)
38. Bertram MJ, Akerkar GA, Ard RL, Gonzalez C, Wolfner MF. 1992 Cell type-specific gene expression in the *Drosophila melanogaster* male accessory gland. *Mech. Dev.* **38**, 33–40. (doi:10.1016/0925-4773(92)90036-j)
39. Majane AC, Cridland JM, Begun DJ. 2022 Single-nucleus transcriptomes reveal functional and evolutionary properties of cell types in the *Drosophila* accessory gland. *Genetics* **220(2)**, 1–15. (doi:10.1101/2021.06.13.448152)
40. Taniguchi K, Kokuryo A, Imano T, Minami R, Nakagoshi H, Adachi-Yamada T. 2014 Isoform-specific functions of Mud/NuMA mediate binucleation of *Drosophila* male accessory gland cells. *BMC Dev. Biol.* **14**, 46. (doi:10.1186/s12861-014-0046-5)
41. Taniguchi K, Kokuryo A, Imano T, Nakagoshi H, Adachi-Yamada T. 2018 Binucleation of accessory gland lobe contributes to effective ejection of seminal fluid in *Drosophila melanogaster*. *Zool. Sci.* **35**, 446. (doi:10.2108/zs170188)
42. Huylmans AK, Parsch J. 2014 Population- and sex-biased gene expression in the excretion organs of *Drosophila melanogaster*. *G3* **4**, 2307–2315. (doi:10.1534/g3.114.013417)
43. Parisi M, Nuttall R, Naiman D, Bouffard G, Malley J, Andrews J, Eastman S, Oliver B. 2003 Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. *Science* **299**, 697–700. (doi:10.1126/science.1079190)
44. Ranz JM, Castillo-Davis CI, Meiklejohn CD, Hartl DL. 2003 Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. *Science* **300**, 1742–1745. (doi:10.1126/science.1085881)
45. Connallon T, Knowles LL. 2005 Intergenomic conflict revealed by patterns of sex-biased gene expression. *Trends Genet.* **21**, 495–499. (doi:10.1016/j.tig.2005.07.006)
46. Sturgill D, Zhang Y, Parisi M, Oliver B. 2007 Demasculinization of X chromosomes in the *Drosophila* genus. *Nature* **450**, 238–241. (doi:10.1038/nature06330)
47. 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**, 1741–7007. (doi:10.1186/1741-7007-9-29)
48. Gupta V, Parisi M, Sturgill D, Nuttall R, Doctolero M, Dudko OK, Malley JD. 2006 Global analysis of X-chromosome dosage compensation. *J. Biol.* **5**, 1–21. (doi:10.1186/jbiol3010.1186/jbiol30)
49. Belmonte RL, Corbally MK, Duneau DF, Regan JC. 2020 Sexual dimorphisms in innate immunity and responses to infection in *Drosophila melanogaster*. *Front. Immunol.* **10**, 1–18. (doi:10.3389/fimmu.2019.03075)
50. Kopp A, Duncan I, Godt D, Carroll SB. 2000 Genetic control and evolution of sexually dimorphic characters in *Drosophila*. *Nature* **408**, 553–559. (doi:10.1038/3504601710.1038/35046017)
51. Lee G, Foss M, Goodwin SF, Carlo T, Taylor BJ, Hall JC. 2000 Spatial, temporal, and sexually dimorphic expression patterns of the *fruitless* gene in the *Drosophila* central nervous system. *J. Neurobiol.* **43**, 404–426. (doi:10.1002/1097-4695(20000615)43:43.0.co;2-d)
52. Demir E, Dickson BJ. 2005 fruitless splicing specifies male courtship behavior in *Drosophila*. *Cell* **121**, 785–794. (doi:10.1016/j.cell.2005.04.02710.1016/j.cell.2005.04.027)
53. Ravi Ram K, Wolfner MF. 2007 Seminal influences: *Drosophila* Acps and the molecular interplay between males and females during reproduction. *Integr. Comp. Biol.* **47**, 427–445. (doi:10.1093/icb/icm046)
54. Li J *et al*. 2025 A noncanonical role of roX RNAs in autosomal epigenetic repression. *Nat. Commun.* **16**, 155. (doi:10.1038/s41467-024-55711-y)
55. Barata C. 2025 carolbarata/snucleus_sex_diffs: updated intermediate files. *Zenodo.* (doi:10.5281/zenodo.17664973)
56. Barata C, Vicoso B. 2025 Supplementary material from: Single-nucleus resolution of sex-biased expression and dosage compensation in *Drosophila melanogaster.* Figshare. (doi:10.6084/m9.figshare.c.8208815)