OXFORD

Sex-biased gene expression at single-cell resolution: cause and consequence of sexual dimorphism

Iulia Darolti[1,2], Judith E. Mank[1]

[1]Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, BC, Canada
[2]Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

Corresponding author: Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, BC V6T 1Z4, Canada. Email: iulia.darolti@unil.ch

Abstract

Gene expression differences between males and females are thought to be key for the evolution of sexual dimorphism, and sex-biased genes are often used to study the molecular footprint of sex-specific selection. However, gene expression is often measured from complex aggregations of diverse cell types, making it difficult to distinguish between sex differences in expression that are due to regulatory rewiring within similar cell types and those that are simply a consequence of developmental differences in cell-type abundance. To determine the role of regulatory versus developmental differences underlying sex-biased gene expression, we use single-cell transcriptomic data from multiple somatic and reproductive tissues of male and female guppies, a species that exhibits extensive phenotypic sexual dimorphism. Our analysis of gene expression at single-cell resolution demonstrates that nonisometric scaling between the cell populations within each tissue and heterogeneity in cell-type abundance between the sexes can influence inferred patterns of sex-biased gene expression by increasing both the false-positive and false-negative rates. Moreover, we show that, at the bulk level, the subset of sex-biased genes that are the product of sex differences in cell-type abundance can significantly confound patterns of coding-sequence evolution. Taken together, our results offer a unique insight into the effects of allometry and cellular heterogeneity on perceived patterns of sex-biased gene expression and highlight the power of single-cell RNA-sequencing in distinguishing between sex-biased genes that are the result of regulatory change and those that stem from sex differences in cell-type abundance, and hence are a consequence rather than a cause of sexual dimorphism.

Keywords: scRNA-sequencing, differential expression, tissue composition, allometry, regulatory evolution

Lay Summary

Males and females of many species display remarkable differences in morphology, physiology, behavior, and other traits, despite sharing an almost identical genome. Differences in gene expression between the sexes, referred to as sex-biased gene expression, are therefore thought to be key for the development of sexually dimorphic traits. However, gene expression is often measured from whole organs or even whole organisms, which represent aggregations of many cell types that can significantly differ in abundance between males and females. Such differences in the relative size or proportion of constituent cell populations may lead to perceived sex-biased gene expression, even if expression within individual cell types is unchanged between the two sexes. As such, it remains unclear to what extent is sex-biased gene expression a consequence of sex differences in cell-type abundance, rather than a cause of sexual dimorphism. To address this, we leverage recent advances in single-cell transcriptomics to compare male and female expression levels across equivalent cell populations from multiple tissues of the guppy, a species that shows extensive sexual dimorphism. We find significant differences in cellular composition between males and females across all sampled tissues and demonstrate that these sex differences in cell-type abundance can alter estimates of sex-biased gene expression. For many genes, expression appears to differ between the sexes at the aggregated whole tissue level but not at the individual cell population level. These sex-biased genes that are a consequence of sex differences in tissue composition can further influence estimates of molecular evolution and studies of selection.

Introduction

Males and females of the same species often exhibit striking differences in a broad range of phenotypic traits, despite sharing the majority of their genome. Changes in patterns of gene expression between the sexes are therefore thought to be key to the evolution of sexual dimorphism (Grath & Parsch, 2016; Mank, 2017), and sex-biased genes are often used as a way to measure the signature of sex-specific selection within the genome. Indeed, a large body of work indicates that the magnitude of transcriptional sex differences correlates with the degree of phenotypic dimorphism across multiple levels of biological diversity (Djordjevic et al., 2022; Ingleby et al., 2015; Magnusson et al., 2011; Mank et al., 2010; Perry et al., 2014; Pointer et al., 2013). However, distinguishing between the causal and the dependent variable has important implications for studying gene expression evolution and for inferences of sexual selection and sexual conflict.

Testing for the causal effect of sex-biased gene expression on phenotypic dimorphism is nontrivial (Abzhanov et al., 2006; Chen et al., 2015; Galouzis & Prud'homme, 2021; Khila et al., 2012; Toubiana et al., 2021). Functional genetic assays are limited by the sheer number of sex-biased loci, the polygenic nature of traits, or simply because the necessary genetic tools are lacking, a particular problem for non-model organisms, which often present the most striking phenotypic differences. Moreover, the assumption that sex-biased genes are the result of regulatory rewiring may not always hold true. Most differential expression analyses have so far relied on bulk RNA-sequencing, comparing expression level between organs (Harrison et al., 2015; Whittle & Extavour, 2019), whole body parts comprising various tissues such as heads or abdomens (Immonen et al., 2017; Standage et al., 2016), or even entire organisms (Djordjevic et al., 2022; Hollis et al., 2014; Ranz et al., 2003; Stuglik et al., 2014). This approach averages expression across a complex aggregate of diverse cell types, ignoring the stochasticity of gene expression across cell populations. In many species, males and females exhibit dimorphism in the relative size of their constituent body parts (Badyaev, 2002), and allometric scaling could influence perception of differential expression in studies where whole organisms are used for RNA preparation (Montgomery & Mank, 2016). In a similar way, whole tissue expression studies may be affected by the heterogeneity in cell-type abundance or composition (Fair et al., 2020; Fuess & Bolnick, 2021; Hunnicutt et al., 2022; Price et al., 2022), and there is indeed evidence that sex differences in cell-type populations exist for many tissues (Mank & Rideout, 2021). For example, differential rates of cell proliferation between males and females seem to underly the development of several sexually dimorphic ornamental traits, such as caudal fins in *Xiphophorus* (Powell et al., 2021; Schartl et al., 2021) and horns in rhinoceros beetles (Emlen et al., 2012). As such, it remains unclear to what extent are sex differences in expression a cause, as opposed to a consequence, of sexual dimorphism.

Sex-biased gene expression is often used to measure the footprint of sex-specific selection and sexual conflict within the genome. However, this approach has produced some discordant results (Grath & Parsch, 2016; Mank, 2017). Male-biased genes in many species tend to exhibit higher rates of evolution at both the coding sequence and the expression level (Khaitovich et al., 2005; Lichilín et al., 2021; Lipinska et al., 2015; Ranz et al., 2003; Sharma et al., 2014; Whittle & Extavour, 2019). While early work in *Drosophila melanogaster* has interpreted this as the result of stronger sexual selection acting in males (Proschel et al., 2006; Sawyer et al., 2007), studies in other species have found that such accelerated patterns of evolution are instead more consistent with relaxed constraint (Dapper & Wade, 2020; Djordjevic et al., 2022; Gershoni & Pietrokovski, 2014; Harrison et al., 2015; Sayadi et al., 2019), and experimental evolution results are mixed (Hollis et al., 2014; Li Richter & Hollis, 2021; Veltsos et al., 2017). Some of the discordance between these different studies may be due to the heterogeneity in tissue composition and varying allometric scaling across species. Distinguishing between sex-biased genes that are due to regulatory changes and those that are simply a consequence of developmental differences in cell-type abundance is important for these types of molecular evolutionary analyses as the former is expected to be more enriched for genes subject to sexual selection and influencing sexual dimorphism.

Single-cell transcriptomics (scRNA-seq) offers the possibility to avoid the challenges posed by measuring expression from a heterogeneous tissue by instead comparing expression level between samples across equivalent cell populations. Here we leverage these recent advances in scRNA-sequencing to determine to what extent sex-biased gene expression is the result of sex differences in cell-type abundance as opposed to regulatory differences between similar cells, and to assess how this impacts inferences of evolutionary divergence.

Methods

Tissue collection and dissociation

We sampled reproductively mature (>3 months of age) male and female guppies from our laboratory population. Sexual maturity in males was detected by the development of a gonopodium, a modified anal fin, and of coloration (Houde, 2019), and in females by the presence of a dark pigmented spot, the gravid spot, close to the anal pore. All fish were raised at a water temperature of 26°C with a 12:12 light:dark schedule and fed a daily diet of flake food and live *Artemia* brine shrimp. Fish were euthanized with a pH-neutralized MS222 solution and dissected immediately. We dissected liver, heart, tail skin (removing any muscle and scales), and gonad (testis in the case of males, and ovaries excluding embryos in the case of pregnant females) tissues and immediately placed them in phosphate-buffered saline (PBS) solution (Corning) on ice. We obtained three replicates for each tissue type and sex, and every replicate contained a nonoverlapping pool of five individuals in order to ensure sufficient material for single-cell dissociation, resulting in a total of 24 samples.

Tissues were incubated at 30°C in a solution of 4 mg/mL collagenase type I (Sigma-Aldrich) and 4 mM $CaCl_2$, gently pipetting every 3 min using a wide-bore pipette tip, until digested. Dissociated tissues were then filter-strained using a Flowmi 40-μm cell strainer. We centrifuged the samples at 300 rcf for 3 min, removing the supernatant and resuspending the cell pellet in PBS containing 0.04% bovine serum albumin (BSA) (Corning). Samples were centrifuged and resuspended in PBS + 0.04% BSA twice and immediately placed on ice before further processing.

Single-cell library preparation and sequencing

We mixed 10 μL of cell solution with 10 μL of the exclusion dye trypan blue 0.4% (Invitrogen) and estimated cell viability and concentration using a Countess II automated cell counter (ThermoFisher). An estimated 8,000 cells from each sample were then loaded onto individual lanes of a 10× Genomic Chromium Controller and barcoded 3′ single-cell libraries were prepared using the 10X Genomics Chromium Next GEM Single Cell 3' kit v3.1 following the manufacturer's instructions (#CG0000204). We assessed the quality and concentration of cDNA and of libraries using an Agilent 4200 TapeStation and the Agilent High Sensitivity D5000 ScreenTape system. Libraries were sequenced on an Illumina NovaSeq 6000 sequencer, with an average sequencing depth of 20,000 read pairs per cell.

scRNA-seq data processing

We used CellRanger v5.0.1 with the "mkref" function (Zheng et al., 2017) to build a reference index using the Ensembl *P. reticulata* genome (GCA_000633615.2) and annotation (release-103) files. Using the CellRanger v5.0.1 "count" function, we then aligned sequencing reads from fastq files to the reference index, identified cell-associated barcodes, and extracted gene-by-cell count matrices. Data filtering and downstream analyses were performed using Seurat v4.1.0 (Satija et al., 2015) in R v4.0.5 (R Core Team, 2021). We filtered the count data by keeping genes expressed in at least three cells and cells with at least 100 expressed genes. Raw counts were then normalized and scaled for each sample

to account for differences in sequencing depth per cell using the "SCTransfrom" function (Hafemeister & Satija, 2019). For each cell type of each tissue, we calculated and plotted Pearson correlations between the number of genes and the number of reads per cell, before and after normalization, using the "FeatureScatter" function in Seurat. We used the function "doubletFinder_v3" from the DoubletFinder v2.0.3 package (McGinnis et al., 2019) in R to identify and remove doublets, which are the result of random encapsulation of more than one cell within the single-cell bead as part of the microfluidics process.

scRNA-seq data clustering

We performed the principal component analysis (PCA) using the "RunPCA" function and identified the significant PCs using the "ElbowPlot" function. We then constructed the nearest-neighbor graph with the "FindNeighbors" function and performed graph-based clustering with the "FindClusters" function. We used the clustree v0.4.4 package in R to guide the decision of the optimal resolution to choose for the clustering analysis (Zappia & Oshlack, 2018). Clusters were then visualized using uniform manifold approximation and projection (UMAP) embedding with "RunUMAP" (Supplementary Figure S1). For each tissue, we used the "FindAllMarkers" function in Seurat to identify differentially expressed genes for each cell cluster. We identified and annotated the different cell types within each tissue based on marker gene information from the Zebrafish Cell Landscape (Jiang et al., 2021; Wang et al., 2023) and other published work (Liu et al., 2022; Morrison et al., 2021; Qian et al., 2022) (Supplementary Figure S2).

Differential cell-type abundance analysis

For each tissue type, we tested for significant differences between male and female samples in the abundance of the identified cell types. We first quantified the number of cells assigned to each cell type and sample. We then merged counts for each sex and calculated the proportion of cells of each cell type out of the total number of cells. We added 1e-10 to each proportion value to avoid infinitely high numbers associated with $\log_2 0$ and calculated the female-to-male fold change (FC) in cell-type abundance as $\log_2$ (female proportion/male proportion). To identify significant differences in cell-type abundance between the sexes, we computed two-proportion z-tests ($p < .01$) with the "prop.test" function in R.

Sex-biased gene expression analysis

To identify sex-biased genes at the cell level, for each identified cell type, we aggregated raw counts across all cells to the sample level using the "aggregate.Matrix" function in R. We then used DESeq2 to normalize the count data accounting for differences in library size between samples and applying a log2-transformation with the "rlog" function (Love et al., 2014). Lastly, we performed differential expression analysis using the "DESeq" function. Sex-biased genes were called based on $|\log_2 FC| \geq 1$ and a false discovery rate (FDR) adjusted p-value $< .05$ to correct for multiple testing. To identify sex-biased genes at the bulk level, we followed the same steps described above, but instead aggregated counts across all cells from all clusters together to obtain a single expression value for each gene and sample.

In addition, we tested the extent to which expression estimates from bulked single-cell RNA-seq data correlate with those from bulk RNA-seq data. For this, we analyzed publicly available bulk RNA-seq data from adult male and female gonad (Sharma et al., 2014) and male skin tissue (NCBI SRA PRJDB11269). We used HISAT2 v2.0.4 (Kim et al., 2015) to map reads to the *P. reticulata* genome keeping paired (--no-mixed) and concordant

(--no-discordant) alignments only. We estimated gene expression by extracting read counts for each gene using HTSeq-count (Anders et al., 2015). We log2-transformed counts and normalized with regards to library size using the "rlog" function in DESeq2. We performed Spearman correlations of gene expression with the "cor.test" function in R.

Rates of coding-sequence evolution analysis

We obtained coding sequences from the outgroup species *Gambusia affinis* (ASM309773v1), *Xiphophorus maculatus* (Xipmac4.4.2), and *Oryzias latipes* (MEDAKA1) from Ensembl 104 and extracted the longest isoform for each gene. We used reciprocal BLASTn v2.7.1 (Altschul et al., 1990) with an *e*-value cutoff of 10e-10 and a minimum percentage identity of 30% to determine orthology across the *P. reticulata* genes and outgroup sequences. For genes with multiple blast hits, we chose the top hit based on the highest BLAST score.

We used *O. latipes* (MEDAKA1) protein-coding sequences from Ensembl 104 and BLASTx v2.3.0 with an *e*-value cutoff of 10e-10 and a minimum percentage identity of 30% to obtain open reading frames. We excluded orthogroups without BLASTx hits or valid protein-coding sequences. We aligned orthologous gene sequences with PRANK v170427 (Loytynoja & Goldman, 2008) and filtered alignments to remove gaps. We also masked poorly aligned and error-rich regions with SWAMP (Harrison et al., 2014) with a threshold of 4 misalignments in a window size of 5 bp and a minimum sequence length of 100 bp.

To obtain divergence estimates, we used the branch model (model = 2, nssites = 0) in the CODEML package in PAML v4.8 (Yang, 2007). Genes with $d_S > 2$ were removed from subsequent analyses to avoid inaccurate divergence estimates due to mutation saturation and double hits (Axelsson et al., 2008). We divided genes into different sex-bias categories (see Figure 3) and extracted the number of nonsynonymous (D_N) and synonymous substitutions (D_S) and the number of nonsynonymous (N) and synonymous (S) sites. For each group of genes, we then calculated the mean rate of nonsynonymous substitutions (d_N) and mean rate of synonymous substitutions (d_S) as the number of substitutions across all genes divided by the number of sites ($d_N = D_N/N$; $d_S = D_S/S$). We used bootstrapping with 1,000 replicates to determine the 95% confidence intervals for divergence estimates in each gene group and tested for differences in d_N, d_S, and d_N/d_S estimates between the different gene groups based on 1,000 replicates permutation tests.

Results

We generated 24 scRNA-seq data sets from skin, heart, liver, and gonad tissue from adult male and female guppies, with three replicates for each sex and tissue. Guppies are highly sexually dimorphic, displaying sex differences in size (Houde, 2019), sexual ornaments (Reznick & Endler, 1982), life history (Reznick & Endler, 1982), and behavior (Reznick, 1989), among other traits, and we chose these four tissues to reflect a range of phenotypic dimorphism. Following quality control and filtering (see Methods), we recovered between 503–3,747, 1,163–8,966, 4,985–6,835, and 1,492–7,973 cells (Supplementary Table S1) and 9,275, 9,818, 11,540, and 13,617 genes expressed in skin, heart, liver, and gonad tissues, respectively (Supplementary Table S2). Following UMAP dimensionality reduction and marker-based annotation, we identified 13, 11, 9, and 8 distinctly expressed cell clusters for skin, heart, liver, and gonad, respectively (Supplementary Figures S1 and S2). These clusters are representative of most major cell types found in these tissues (Jiang et al., 2021; Wang et al., 2023).

Sex differences in cell-type abundance across tissues

We examined each tissue for differences in cell composition between the sexes based on the estimated abundances of each of the identified cell types. We found significant sex differences ($p < .01$, two-proportion z-tests) in abundance for many cell populations (Figure 1; Supplementary Table S3), the most extreme differences being present in the reproductive tissue, where most of the cell types are sex limited (Figure 1D, Supplementary Figure S3). Notably, the vast majority of cells in the male gonad are related to sperm, while the female gonad represents a broader mix of gametic and somatic cells (Supplementary Figure S3). However, we observe a range of dimorphism in cell composition among somatic tissues as well. The liver plays a critical role in several physiological processes, including hormonal regulation, metabolism, digestion, and immune response (Bruslé & Anadon, 1996), and shows marked sex differences (Marcos et al., 2015; Ullah et al., 2021). Consistent with the more sexually dimorphic nature of this somatic tissue, we found extensive sex differences in cellular composition, with all but one of the identified cell types

showing a significant sex-bias in abundance (Figure 1C). Although the heart and skin tissues are less sexually dimorphic compared with the liver, we also recovered sex differences in abundance for more than half of the cell populations in these two tissues (Figure 1A and B). Markedly, guppies exhibit striking sexual dimorphism in skin pigmentation, with male-specific ornamental color patterns (Reznick & Endler, 1982), which we also find reflected here through the strong male-biased abundance of melanocytes in males (Figure 1).

Sex-biased gene expression at bulk and cell level

Dropout events are a well-known challenge with scRNA-seq data, where the prevalence of genes with nonbiological zero expression measurements can confound gene expression analyses. We employed a normalization and variance stabilization approach to reduce the effect of technical noise on estimates of gene expression (Hafemeister & Satija, 2019). Indeed, for most cell types across each tissue, we find a significantly decreased correlation between the expression level of a gene and the total sequencing depth of a cell following normalization (Supplementary Figure
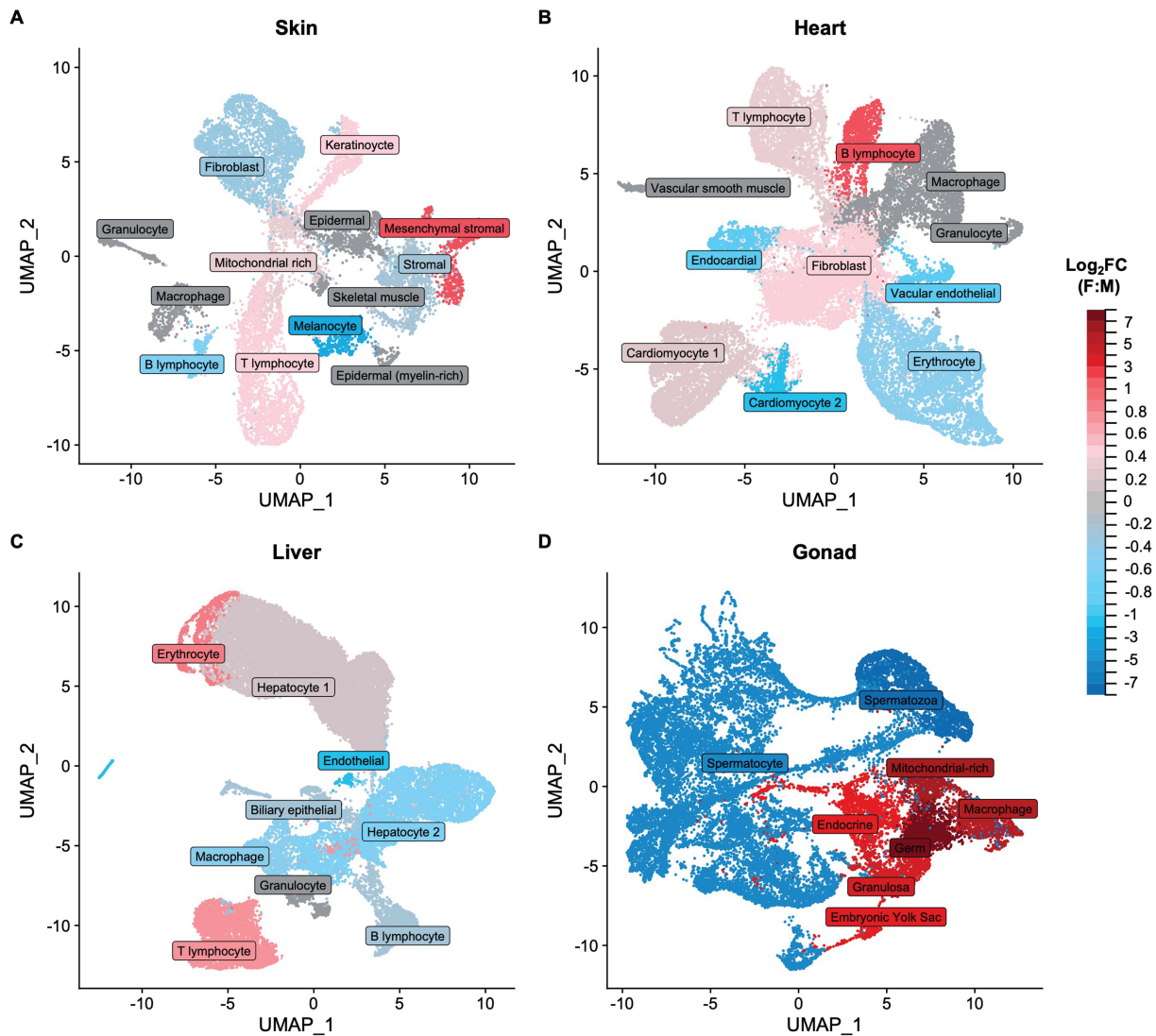

Figure 1. Differential cell-type abundance between males and females for skin (A), heart (B), liver (C), and gonad (D) tissues. Cell types that are significantly more abundant in females are shown in red, those that are significantly male-biased in abundance are shown in blue, while unbiased cell types are in gray. Significance based on two-proportions z-tests ($p < .01$).

S4). This indicates that our gene expression analyses should not be influenced by variation in sequencing depth and that differences in gene expression across cells primarily reflect biological heterogeneity (Hafemeister & Satija, 2019).

For each tissue, we aggregated expression measures across all cells to obtain a single value for each gene and sample, thus reflecting expression at the bulk tissue level. Using the bulked scRNA-seq data, we identified genes with a sex-biased expression profile ($|\log_2$ fold change$| \geq 1$, FDR-corrected p-value < .05). Consistent with previous estimates in guppies (Sharma et al., 2014), we find that, at the bulk level, male and female gonads exhibit strong patterns of sex-biased gene expression, with 9,524 genes identified as differentially expressed, representing 70% of all genes expressed in the gonads (Figure 2D, Supplementary Table S2). Top male-biased genes included those encoding for spermatogenesis processes (*march11*, *morn3*, *spata18*, *spatc1l*), ciliary- and flagellar-associated proteins (*cfap52*, *ropn1l*, *tekt1*), male-specific development (*dmrt1*), and calcium-binding proteins (*efcab2*), while female-biased genes encoded for, among others, zona pellucida glicoproteins (*zp2l2*, *zp3f.1*, *zp3d.2*), oocyte-specific proteins (*zar1*, *zar1l*), and ovarian folliculogenesis processes (*gdf9*, *cx43.4*, *pcdh18a*) (Supplementary Table S2).

Given their sex-specific roles, reproductive organs are by far the most transcriptionally dimorphic tissues; however, there is substantial variation in sex differences in expression across somatic tissues as well (Khodursky et al., 2020; Ma et al., 2018; Mank et al., 2007; Oliva et al., 2020; Yang et al., 2006), which we also observe here. The liver was the most transcriptionally dimorphic somatic tissue, with 349 sex-biased genes, more than twice as many compared with the skin (114) and heart (155) tissues (Figure 2A–C). The liver has been shown to be one of the somatic tissues with the most sex-biased gene expression profiles in other fish species (Qiao et al., 2016; Rose et al., 2015; Taboada et al., 2012; Zheng et al., 2013). In fish and other oviparous vertebrates, the liver carries an important role in the vitellogenesis process of egg yolk protein synthesis, transport, and uptake in the maturing oocyte (Arukwe & Goksøyr, 2003; Hara et al., 2016), and, in line with this, we also find vitellogenins (*vtg1*, *vtg2*, *vtg3*) to be some of the top female-biased genes in the liver (Supplementary Table S2). While skin and heart tissues express fewer sex-biased genes, we also note highly male-biased genes in skin with a role in pigmentation (*fhl2a*, *pnp4a*).

To test whether bulked scRNA-seq expression data correlates with bulk RNA-seq data, we analyzed publicly available bulk RNA-seq data for adult male and female gonad and male skin tissue. Unfortunately, bulk data for the other tissues were not available. We found a strong correlation in gene expression estimates between the bulk RNA-seq and the bulk scRNA-seq datasets for both male gonad and male skin tissues (Supplementary Figure S5). Multiple factors, such as technical variation and
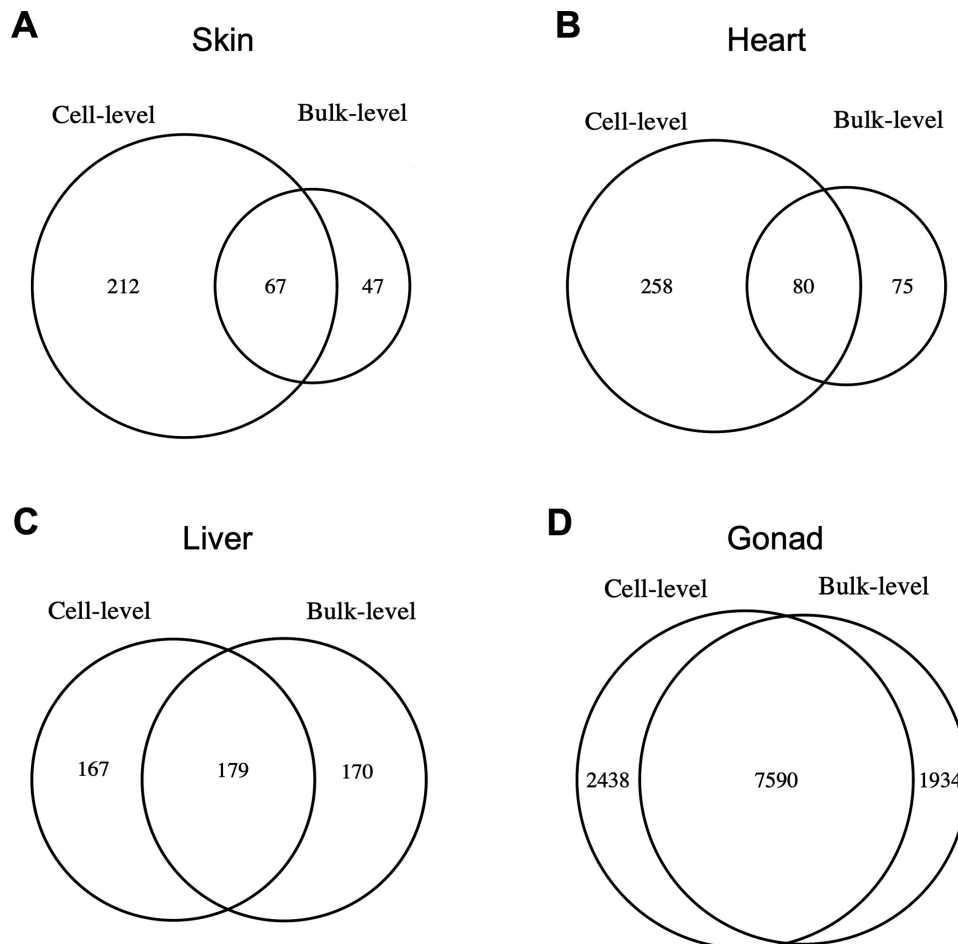

Figure 2. Number of genes showing sex-biased gene expression at the cell and bulk level in skin (A), heart (B), liver (C), and gonad (D) tissues. Numbers at the "Cell-level" represent the union of genes identified as sex-biased across all the different cell types. Significant differential expression between the sexes was based on $|\log_2$ FC$| \geq 1$ and an FDR-corrected p-value < .05.
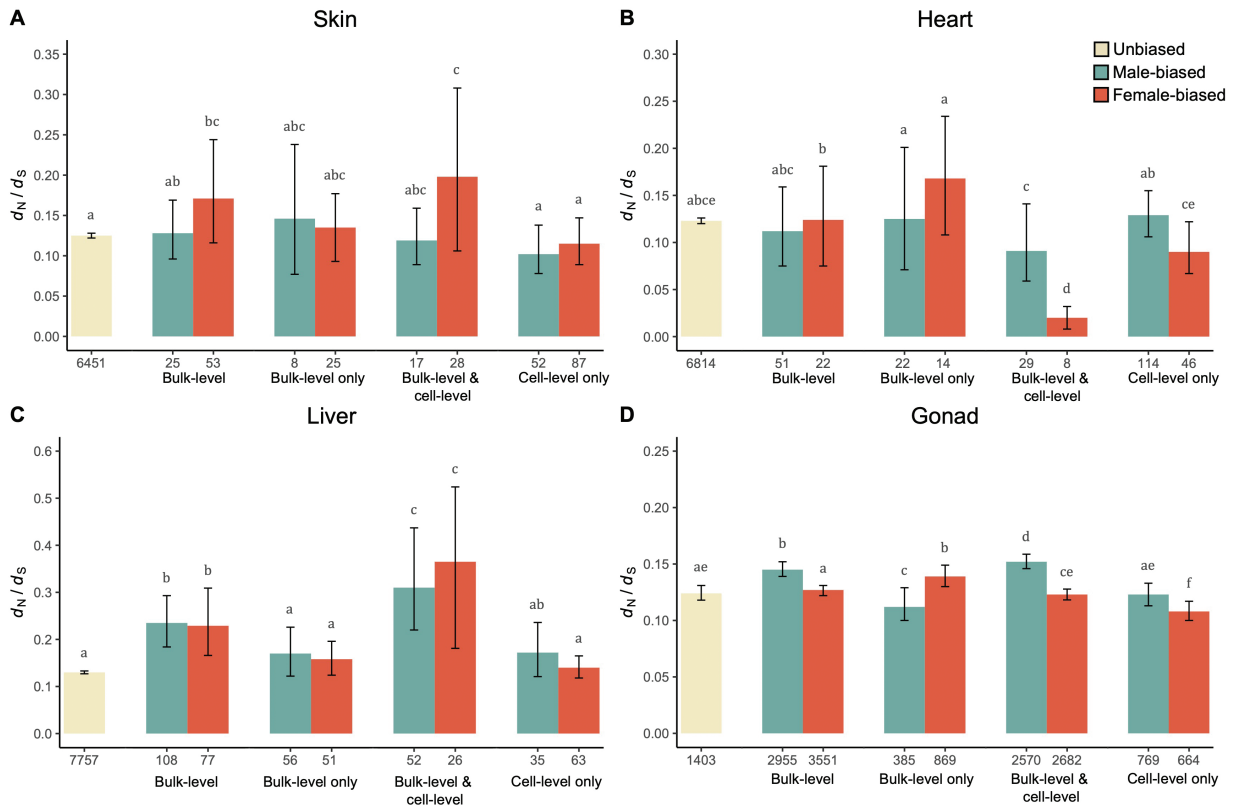
Figure 3. Rates of coding-sequence evolution (d_N/d_S) for genes in skin (A), heart (B), liver (C), and gonad (D). Shown are estimates of divergence for genes that are unbiased at both the bulk level and cell level (yellow), all genes identified as male-biased (green) and female-biased (orange) at the bulk level (Bulk-level), genes with a male-biased and female-biased expression at the bulk level but that are unbiased in every cell type within that tissue (Bulk-level only), genes that show a sex-biased expression profile both at the bulk level and in at least one of the identified cell types in that tissue (Bulk-level and Cell-level), and sex-biased genes at the cell level but not at the bulk level (Cell-level only). Numbers on the x-axis represent the number of genes in each group. Letters above bars indicate pairwise significance between groups based on 1,000 replicates permutations, where bars that share a letter are not significantly different, while bars with no overlapping letters are significantly different ($p < .05$).

differences in sample preparation, can explain the lack of correlation for the female gonad tissue. We observe a very strong correlation between the three bulk scRNA-seq female gonad samples (Supplementary Figure S5). However, there was only one bulk RNA-seq female gonad sample available which prevented us from testing for the effects of technical noise on the correlation analysis. Nevertheless, previous work in other systems has shown a strong correlation between expression estimates from bulk RNA-seq and bulk scRNA-seq data across various tissues (Karlsson et al., 2021; Marinov et al., 2014; Trapnell et al., 2014).

We next compared patterns of differential gene expression between males and females at the whole tissue level with those at the cell level. As such, in addition to the aggregated bulk-level expression, for each identified cell type we aggregated expression data across all cells to identify sex-biased genes in each cell population. Overall, we discovered more sex-biased genes at the cell level than at the bulk level in all tissues, except for liver. However, what is even more striking is that across the somatic tissues between 49% and 77% of the identified sex-biased genes at the cell level have unbiased expression profiles at the bulk level (Figure 2). By contrast, in the gonad, only about 24% of the differentially expressed genes at the cell level show no sex differences in expression at the bulk level. These results indicate that, outside of the reproductive tissues, differential gene expression analyses based on bulk RNA-sequencing data are limited in their ability to comprehensively identify sex-biased genes. Nonisometric-scaling relationships between cellular subcomponents of a tissue may

cause patterns of differential expression in low abundance cell types to be concealed when expression is aggregated across various cell types in standard bulk RNA-sequencing experiments, thus creating a false-negative inference of sex-biased gene expression.

Moreover, in all somatic tissues, many genes are characterized as sex-biased at the bulk level, but not at the cellular level. We hypothesized that these differential expression patterns are not due to regulatory rewiring within cell types but instead a consequence of differences in cell-type abundance between males and females. Indeed, these genes tend to have a significantly lower magnitude of expression fold change (Supplementary Figure S6) and are more highly expressed in cell types that exhibit significant sex differences in abundance (Supplementary Figure S7). In skin, genes that are female-biased at the bulk level only are predominantly expressed in the mesenchymal stromal cell population which is more abundant in females. Similarly in liver, genes that are female- and male-biased in expression at the bulk level but not at the cell level have a significantly higher expression in the female-abundant T lymphocytes and the male-abundant endothelial cells, respectively. Although in the gonad a much smaller percentage of genes are sex-biased only at the bulk level, these too appear to be the result of sex differences in cell-type abundance (Supplementary Figure S7). These results show that heterogeneity in cellular scaling relationships, due to developmental differences in cell proliferation, between males and females

can influence perceived patterns of differential gene expression in the absence of regulatory changes.

Rates of evolution of sex-biased genes

We tested whether the cause of sex-biased gene expression influences estimates of coding-sequence evolution. We hypothesized that sex-biased genes that are due to sex differences in cell-type abundance, and thus a consequence of sex-specific developmental trajectories, may not be direct targets of sex-specific selection and may not show elevated rates of sequence divergence compared with unbiased genes. By contrast, loci exhibiting regulatory differences between males and females are more likely to underlyly sexually dimorphic phenotypes and would be expected to show stronger signatures of sex-specific selection, if indeed these genes as a class are more likely to be the focus of sexual selection.

When considering all the differentially expressed genes at the bulk level, we found that male-biased genes in liver and gonad tissues and female-biased genes in skin, liver, and gonad tissues evolve significantly faster than unbiased genes (Figure 3), owing to higher rates of nonsynonymous substitutions (Supplementary Table S4). To further understand what drives these elevated divergence patterns at the bulk level, we split genes into those that are sex-biased at the bulk level but unbiased at the cell level, which are more likely to result from sex differences in cell-type abundance (Supplementary Figure S7), and those that are sex-biased at both the bulk and cell level, and which are due to regulatory sex differences. We found that the high divergence rates observed at the bulk level are driven by the subset of genes that are sex-biased at both the bulk and the cell level, as these genes exhibited an even more accelerated rate of evolution (Figure 3) compared to unbiased genes. These loci are more likely to be underlying sexually dimorphic phenotypes as they also exhibit elevated $\log_2$FC estimates (Supplementary Figure S6). On the other hand, the subset of genes that are sex-biased at the bulk level only, and thus, the outcome of differential tissue composition between the sexes, showed similar rates of coding-sequence evolution to unbiased genes, further suggesting that they are not the main targets of selection. An exception to this may be highly differentially expressed genes in the gonad that have important reproductive functions, such as sperm-specific genes, and which would therefore be subject to sexual selection. In general, however, we observe that false-positive differentially expressed genes can obscure true patterns of evolutionary divergence for sex-biased genes at the bulk level.

Previous work has noted that the level of sequence divergence increases with the degree of sex-biased gene expression (Djordjevic et al., 2022; Harrison et al., 2015; Lichilín et al., 2021). To test this, we split sex-biased genes from the gonad tissue into different groups based on their level of sex-bias. We also find that highly male-biased and female-biased genes exhibit greater divergence than lowly male-biased and female-biased genes (Supplementary Figure S8). Considering that we observe a higher expression fold change for genes that are sex-biased at both the bulk and the cell level (Supplementary Figure S6), this suggests that the accelerated rates of sequence divergence for this group of genes may at least in part be due to their more extreme sex-bias in expression. However, the small number of sex-biased genes for which we could estimate rates of coding-sequence evolution prevented us from investigating this further in the somatic tissues.

Discussion

To determine the relative role of regulatory versus developmental differences underlying sex-biased gene expression, we used single-cell transcriptomic data from multiple somatic and reproductive tissues of male and female guppies, a species that exhibits extensive phenotypic sexual dimorphism. Sex-biased genes can be a cause of sexual dimorphism, as is often assumed in studies that use them to study the molecular genetic footprint of sexual selection and sexual conflict (e.g., Harrison et al., 2015; Mank, 2017; Sayadi et al., 2019). Alternatively, sex-biased genes may be a consequence of developmental differences in cell proliferation between the sexes that result in differences in cell-type abundances. Sex differences in cell populations are known to exist in many tissues (Mank & Rideout, 2021), and differential rates of cell proliferation underly the development of several sexually dimorphic ornamental traits (Emlen et al., 2012; Powell et al., 2021; Schartl et al., 2021). The question of whether sex-biased genes are a cause or consequence of sexual dimorphism is critical, as the former might be subject to differences in sex-specific selection and therefore useful in studies of sexual selection, sexual dimorphism, and sexual conflict, while the latter are largely a consequence of differences in developmental programming.

Our analysis of gene expression at single-cell resolution illustrates that cellular heterogeneity within tissues and allometric scaling differences of cell types between males and females can have a major influence on inferred patterns of sex-biased gene expression. Such scaling effects can generate both false-negative detection of differentially expressed genes, which may preponderantly concern genes that are expressed in low abundance cell types, as well as false-positive patterns of differential gene regulation, as is the case for genes that are sex-biased at the bulk tissue level but not at the cell level. These false-positive results have the potential to affect inferences on the evolution of sex-biased gene expression in several ways.

Allometry can be a confounding factor in estimates of rates of coding-sequence evolution for sex-biased genes, as we show that false-positive sex-biased genes diminish the signal of elevated rates of divergence for both male- and female-biased genes. Variation in the fraction of genes that are erroneously classified as sex-biased due to sex differences in tissue composition may explain some of the discordance in patterns of rapid rates of coding-sequence evolution and sex-specific selection observed across studies (Harrison et al., 2015; Khaitovich et al., 2005; Lipinska et al., 2015; Mank et al., 2010; Ranz et al., 2003; Whittle & Johannesson, 2013). Sex differences in gene co-expression networks are thought to contribute to sexually dimorphic phenotypes and potentially alleviate sexually antagonistic selection (Lopes-Ramos et al., 2020; Rago et al., 2020; Sutherland et al., 2019). Yet fundamental differences in gene regulatory networks can exist between cell types and variation in cell-type abundance between the sexes can also affect gene co-expression measurements (Ribeiro et al., 2022).

The degree of sex differences in cell-type abundance likely varies across species as a function of phenotypic sexual dimorphism, and this can influence patterns of rapid turnover of sex-bias across species. We observed large differences in cell-type abundance between the sexes (Supplementary Table S3, Figure 1, Supplementary Figure S3), with the liver showing the least overall fold change in cell-type abundance, moderate levels in the skin and heart, and the greatest observed in the gonad. In some cases, cell-type abundance differences were consistent with visible phenotypic differences, such as the male-bias in melanocytes in male skin, as might be expected from male coloration. However, many cell types with sex differences might not necessarily be predicted from phenotypic differences, such as the male-bias in heart cardiomyocytes, or the female-bias in skin mesenchymal

stromal cells. Identifying distinct cell populations for non-model organisms relies largely on databases of cell-type-specific marker genes from model species, which may be distantly related. While our analysis did not exhaustively identify all the different cell subpopulations, male and female homologous cell clusters were annotated in the same way, and we were able to recover all the major cell types expected within each tissue (Jiang et al., 2021; Wang et al., 2023).

Notably, spermatocytes and spermatozoa combined make up 97% of male gonad cells. In contrast, the female gonad is comprised of only 20% germ cells, with the remainder a mix of somatic cells. These patterns are consistent with previous findings in zebrafish (Jiang et al., 2021) and point to the fact that bulk comparisons between male and female gonads, or whole bodies where the majority of expression differences are due to the gonad (Parisi et al., 2004), are in practice comparing expression related to sperm in males with a range of cell functions in females. Male-biased genes identified from whole-organism or bulk gonad preparations in many species show rapid rates of protein evolution (Ellegren & Parsch, 2007; Parsch & Ellegren, 2013; Ranz et al., 2003) and exhibit high rates of turnover (Harrison et al., 2015; Khodursky et al., 2020; Papa et al., 2017; Whittle & Extavour, 2019; Zhang et al., 2007), both of which might be expected from sexual selection predictions. Indeed, in our bulk analysis, male-biased genes in the gonad show elevated rates of evolution (Figure 3D). However, it is important to note that genes with restricted expression (e.g., restricted to sperm) often experience fewer adaptive constraints (Yannai et al., 2005) than those with broader expression. Indeed, when comparing genes that are sex-biased within gonad cell types, rather than those that differ as a result of cell-type abundance, male-biased genes in the gonad evolve at the same rate as unbiased genes (Figure 3D).

Overall, our results suggest that sex differences in cell-type abundance scale with visible sexual dimorphism, suggesting that bulk RNA-seq approaches may be increasingly problematic in more dimorphic species. Single-cell transcriptomics, such as that employed here, offer a promising way to correct for any potential false inferences of differential gene expression and patterns of sequence divergence that are associated with standard bulk-level sequencing of heterogeneous tissue samples. However, although scRNA-sequencing approaches are increasingly employed for disentangling the role of regulatory changes in the evolution of intra- and interspecific phenotypic variation (Fuess & Bolnick, 2021; Murat et al., 2023), the associated costs are still substantially higher compared with standard bulk RNA-sequencing and many analytical challenges remain (Alfieri et al., 2022). Cell isolation protocols, especially for non-model organisms, are nontrivial as cell physiology may differ across tissues and species, and often require organism-specific technical knowledge to avoid cell death or bias in expression profiles (Reichard & Asosingh, 2019; Wang et al., 2021). Incomplete or poorly annotated genomes can additionally limit the identification of cell identities and gene expression patterns (Healey et al., 2022), although several *k*-mer-based methods have been developed to aid with de novo transcriptome assembly and identification of cell populations in cases where a reference genome is lacking (Nip et al., 2020; Sun et al., 2021). Compared with bulk RNA-seq data, scRNA-seq data are also affected by a greater sparsity, where a high proportion of genes have zero expression measurements within cells (Jiang et al., 2022). This sparsity is often due to cell-to-cell variation in technical factors, such as the number of molecules detected per cell and stochastic sampling, which can confound results (Stegle

et al., 2015). However, normalization and variance stabilization approaches, such as the one employed here, can reduce the effect of technical noise on estimates of gene expression (Hafemeister & Satija, 2019). Moreover, performing differential gene expression analyses on aggregate cell data within each cell type, rather than on individual cells, reduces the number of zeroes in the data and offers better performance especially for lowly expressed genes (Squair et al., 2021).

In the absence of scRNA-sequencing data or information regarding the cellular composition of tissues, we suggest that adopting more stringent fold change thresholds for calling differentially expressed genes has the potential to substantially reduce the false-positive rates associated with bulk-level sequencing of heterogeneous tissues. Although a few studies have accounted for increasing fold change thresholds when studying the evolution of differential gene expression (Darolti et al., 2018; Djordjevic et al., 2022; Harrison et al., 2015; Ma et al., 2018, 2020; Mariani et al., 2003; Perry et al., 2014), others rely on a $\log_2$ fold change of 1 or lower, or on statistical significance alone. However, our results from bulk-level analyses show that such lowered thresholds can result in the inclusion of many sex-biased genes that are due to differences in cell-type abundance between males and females (Supplementary Figure S6). Although this attempt to reduce false-positive effects may also inadvertently remove some true-positive sex-biased genes, strong patterns of transcriptional dimorphism will remain unaffected. This approach would be particularly recommended for studies in which estimating additional tissue scaling parameters is limited due to the small size of the organism or difficulty in precise tissue dissection. Alternatively, deconvolution methods may also prove useful (Aguirre-Gamboa et al., 2020; Monaco et al., 2019).

Taken together, our findings offer an important insight into the effect of allometry and heterogeneity in tissue composition on inferred patterns of sex-biased gene expression and demonstrate the power of single-cell RNA-sequencing in differentiating sex-biased genes that stem from regulatory rewiring between males and females from those that are due to differences in cell-type abundance resulting from sex-specific developmental trajectories.

Supplementary material

Supplementary material is available online at *Evolution Letters* (https://academic.oup.com/evlett/qrad013).

Data availability

scRNA-sequencing reads have been deposited to the NCBI Short Read Archive (BioProject ID PRJNA902547). Scripts used for data processing and analysis are available on GitHub (https://github.com/manklab/Darolti_and_Mank_2022_Guppy_SingleCellExpression) and are archived on Zenodo (https://doi.org/10.5281/zenodo.7788730).

Author contributions

I.D. and J.E.M. conceived the study, performed the analysis and wrote the manuscript.

Conflict of interest: The authors declare no conflict of interest. Editorial processing of the manuscript was done independently of J.E.M. who is an associate editor of *Evolution Letters*.

Acknowledgments

References

Abzhanov, A., Kuo, W. P., Hartmann, C., Grant, B. R., Grant, P. R., & Tabin, C. J. (2006). The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. *Nature*, **442**(7102), 563–567. https://doi.org/10.1038/nature04843

Aguirre-Gamboa, R., de Klein, N., di Tommaso, J., Claringbould, A., van der Wijst, M. G. P., de Vries, D., Brugge, H., Oelen, R., Võsa, U., Zorro, M. M., Chu, X., Bakker, O. B., Borek, Z., Ricaño-Ponce, I., Deelen, P., Xu, C. J., Swertz, M., Jonkers, I., Withoff, S., ... Li., Y. (2020). Deconvolution of bulk blood eQTL effects into immune cell subpopulations. *BMC Bioinformatics*, **21**, 1–23.

Alfieri, J. M., Wang, G., Jonika, M. M., Gill, C. A., Blackmon, H., & Athrey, G. N. (2022). A primer for single-cell sequencing in non-model organisms. *Genes*, **13**(2), 380. https://doi.org/10.3390/genes13020380

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, **215**(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**(2), 166–169. https://doi.org/10.1093/bioinformatics/btu638

Arukwe, A., & Goksøyr, A. (2003). Eggshell and egg yolk proteins in fish: Hepatic proteins for the next generation: Oogenetic, population, and evolutionary implications of endocrine disruption. *Comparative hepatology*, **2**, 1–21.

Axelsson, E., Hultin-Rosenberg, L. I. N. A., Brandström, M., Zwahlen, M., Clayton, D. F., & Ellegren, H. (2008). Natural selection in avian protein-coding genes expressed in brain. *Molecular Ecology*, **17**, 3008–3017.

Badyaev, A. V. (2002). Growing apart: An ontogenetic perspective on the evolution of sexual size dimorphism. *Trends in Ecology & Evolution*, **17**, 369–378.

Bruslé, J., & Anadon, G. G. (1996). The structure and function of fish liver. In Munshi, J. S. D., & Dutta, H. M. (Eds.), *Fish morphology*. Science Publishers Inc.

Chen, Y. C., Harrison, P. W., Kotrschal, A., Kolm, N., Mank, J. E., & Panula, P. (2015). Expression change in Angiopoietin-1 underlies change in relative brain size in fish. *Proceedings of the Royal Society B: Biological Sciences*, **282**(1810), 20150872. https://doi.org/10.1098/rspb.2015.0872

Dapper, A. L., & Wade, M. J. (2020). Relaxed selection and the rapid evolution of reproductive genes. *Trends in Genetics*, **36**(9), 640–649. https://doi.org/10.1016/j.tig.2020.06.014

Darolti, I., Wright, A. E., Pucholt, P., Berlin, S., & Mank, J. E. (2018). Slow evolution of sex-biased genes in the reproductive tissue of the dioecious plant Salix viminalis. *Molecular Ecology*, **27**(3), 694–708. https://doi.org/10.1111/mec.14466

Djordjevic, J., Dumas, Z., Robinson-Rechavi, M., Schwander, T., & Parker, D. J. (2022). Dynamics of sex-biased gene expression during development in the stick insect *Timema californicum*. *Heredity*, **129**, 113–122.

Ellegren, H., & Parsch, J. (2007). The evolution of sex-biased genes and sex-biased gene expression. *Nature Reviews Genetics*, **8**(9), 689–698. https://doi.org/10.1038/nrg2167

Emlen, D. J., Warren, I. A., Johns, A., Dworkin, I., & Lavine, L. C (2012). A mechanism of extreme growth and reliable signaling in sexually selected ornaments and weapons. *Science*, **337**(6096), 860–864. https://doi.org/10.1126/science.1224286

Fair, B. J., Blake, L. E., Sarkar, A., Pavlovic, B. J., Cuevas, C., & Gilad, Y. (2020). Gene expression variability in human and chimpanzee populations share common determinants. *Elife*, **9**, e59929. https://doi.org/10.7554/eLife.59929

Fuess, L. E., & Bolnick, D. I (2021). Single-cell RNA sequencing reveals micro-evolution of the stickleback immune system. *bioRxiv*. https://doi.org/10.1101/2021.12.20.473470

Galouzis, C. C., & Prud'homme, B. (2021). Transvection regulates the sex-biased expression of a fly X-linked gene. *Science*, **371**(6527), 396–400. https://doi.org/10.1126/science.abc2745

Gershoni, M., & Pietrokovski, S. (2014). Reduced selection and accumulation of deleterious mutations in genes exclusively expressed in men. *Nature Communications*, **5**, 1–10.

Grath, S., & Parsch, J. (2016). Sex-biased gene expression. *Annual Review of Genetics*, **50**, 29–44. https://doi.org/10.1146/annurev-genet-120215-035429

Hafemeister, C., & Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biology*, **20**, 1–15.

Hara, A., Hiramatsu, N., & Fujita, T. (2016). Vitellogenesis and choriogenesis in fishes. *Fisheries Science*, **82**(2), 187–202. https://doi.org/10.1007/s12562-015-0957-5

Harrison, P. W., Jordan, G. E., & Montgomery, S. H. (2014). SWAMP: Sliding window alignment masker for PAML. *Evolutionary Bioinformatics*, **10**, EBO.S18193EBO–EBO.S18193. https://doi.org/10.4137/ebo.s18193

Harrison, P. W., Wright, A. E., Zimmer, F., Dean, R., Montgomery, S. H., Pointer, M. A., & Mank, J. E. (2015). Sexual selection drives evolution and rapid turnover of male gene expression. *Proceedings of the National Academy of Sciences*, **112**(14), 4393–4398. https://doi.org/10.1073/pnas.1501339112

Healey, H. M., Bassham, S., & Cresko, W. A. (2022). Single-cell iso-sequencing enables rapid genome annotation for scRNAseq analysis. *Genetics*, **220**(3), iyac017. https://doi.org/10.1093/genetics/iyac017

Hollis, B., Houle, D., Yan, Z., Kawecki, T. J., & Keller, L. (2014). Evolution under monogamy feminizes gene expression in *Drosophila melanogaster*. *Nature Communications*, **5**, 1–5.

Houde, A. (2019). *Sex, color, and mate choice in guppies*. Princeton University Press.

Hunnicutt, K. E., Good, J. M., & Larson, E. L (2022). Unraveling patterns of disrupted gene expression across a complex tissue. *Evolution*, **76**(2), 275–291. https://doi.org/10.1111/evo.14420

Ingleby, F. C., Flis, I., & Morrow, E. H. (2015). Sex-biased gene expression and sexual conflict throughout development. *Cold Spring Harbor Perspectives in Biology*, **7**, a017632.

Immonen, E., Sayadi, A., Bayram, H., & Arnqvist, G. (2017). Mating changes sexually dimorphic gene expression in the seed beetle Callosobruchus maculatus. *Genome Biology and Evolution*, **9**(3), 677–699. https://doi.org/10.1093/gbe/evx029

Jiang, M., Xiao, Y., E, W., Ma, L., Wang, J., Chen, H., Gao, C., Liao, Y., Guo, Q., Peng, J., Han, X., & Guo, G. (2021). Characterization of the zebrafish cell landscape at single-cell resolution. *Frontiers in Cell and Developmental Biology*, **9**, 2734.

Jiang, R., Sun, T., Song, D., & Li, J. J. (2022). Statistics or biology: The zero-inflation controversy about scRNA-seq data. *Genome Biology*, **23**, 1–24.

Karlsson, M., Zhang, C., Méar, L., Zhong, W., Digre, A., Katona, B., Sjöstedt, E., Butler, L., Odeberg, J., Dusard, P., & Edfors, F. (2021). A single-cell type transcriptomics map of human tissues. *Science Advances*, **31**, eabh2169.

Khaitovich, P., Hellmann, I., Enard, W., Nowick, K., Leinweber, M., Franz, H., Weiss, G., Lachman, M., & Paabo, S. (2005). Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science*, **309**, 1850–1854.

Khila, A., Abouheif, E., & Rowe, L. (2012). Function, developmental genetics, and fitness consequences of a sexually antagonistic trait. *Science*, **336**(6081), 585–589. https://doi.org/10.1126/science.1217258

Khodursky, S., Svetec, N., Durkin, S. M., & Zhao, L. (2020). The evolution of sex-biased gene expression in the *Drosophila* brain. *Genome Research*, **30**(6), 874–884. https://doi.org/10.1101/gr.259069.119

Kim, D., Langmead, B., & Salzbergg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, **12**, 357–360.

Li Richter, X. Y., & Hollis, B. (2021). Softness of selection and mating system interact to shape trait evolution under sexual conflict. *Evolution*, **75**(10), 2335–2347. https://doi.org/10.1111/evo.14329

Lichilín, N., El Taher, A., & Böhne, A. (2021). Sex-biased gene expression and recent sex chromosome turnover. *Philosophical Transactions of the Royal Society B*, **376**, 20200107.

Lipinska, A., Cormier, A., Luthringer, R., Peters, A. F., Corre, E., Gachon, C. M., Cock, J., M., & Coelho, S. M. (2015). Sexual dimorphism and the evolution of sex-biased gene expression in the brown alga Ectocarpus. *Molecular Biology and Evolution*, **32**, 1581–1597.

Liu, Y., Kossack, M. E., McFaul, M. E., Christensen, L. N., Siebert, S., Wyatt, S. R., Kamei, C. N., Horst, S., Arroyo, N., Drummond, I. A., Juliano, C. E., & Draper, B. W. (2022). Single-cell transcriptome reveals insights into the development and function of the zebrafish ovary. *Elife*, **11**, e76014.

Lopes-Ramos, C. M., Chen, C. Y., Kuijjer, M. L., Paulson, J. N., Sonawane, A. R., Fagny, M., Platig, J., Glass, K., Quackenbush, J., & DeMeo, D. L (2020). Sex differences in gene expression and regulatory networks across 29 human tissues. *Cell reports*, **31**, 107795.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**, 1–21.

Loytynoja, A., & Goldman, N. (2008). Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. *Science*, **320**(5883), 1632–1635. https://doi.org/10.1126/science.1158395

Ma, W. J., Carpentier, F., Giraud, T., & Hood, M. E. (2020). Differential gene expression between fungal mating types is associated with sequence degeneration. *Genome Biology and Evolution*, **12**(4), 243–258. https://doi.org/10.1093/gbe/evaa028

Ma, W. J., Veltsos, P., Toups, M. A., Rodrigues, N., Sermier, R., Jeffries, D. L., & Perrin, N. (2018). Tissue specificity and dynamics of sex-biased gene expression in a common frog population with differentiated, yet homomorphic, sex chromosomes. *Genes*, **9**(6), 294. https://doi.org/10.3390/genes9060294

Magnusson, K., Mendes, A. M., Windbichler, N., Papathanos, P. A., Nolan, T., Dottorini, T., Rizzi, E., Christophides, G. K., & Crisanti, A. (2011). Transcription regulation of sex-biased genes during ontogeny in the malaria vector *Anopheles gambiae*. *PLoS One*, **6**, e21572.

Mank, J. E. (2017). The transcriptional architecture of phenotypic dimorphism. *Nature Ecology & Evolution*, **1**, 1–7.

Mank, J. E., Hultin-Rosenberg, L., Axelsson, E., & Ellegren, H. (2007). Rapid evolution of female-biased, but not male-biased, genes expressed in the avian brain. *Molecular Biology and Evolution*, **24**(12), 2698–2706. https://doi.org/10.1093/molbev/msm208

Mank, J. E., Nam, K., Brunström, B., & Ellegren, H. (2010). Ontogenetic complexity of sexual dimorphism and sex-specific selection. *Molecular Biology and Evolution*, **27**(7), 1570–1578. https://doi.org/10.1093/molbev/msq042

Mank, J. E., & Rideout, E. J. (2021). Developmental mechanisms of sex differences: From cells to organisms. *Development*, **148**(19), dev199750. https://doi.org/10.1242/dev.199750

Marcos, R., Correia-Gomes, C., Miranda, H., & Carneiro, F. (2015). Liver gender dimorphism: Insights from quantitative morphology. *Histology and Histopathology*, **30**(12), 1431–1437. https://doi.org/10.14670/HH-11-648

Mariani, T. J., Budhraja, V., Mecham, B. H., Gu, C. C., Watson, M. A., & Sadovsky, Y. (2003). A variable fold-change threshold determines significance for expression microarrays. *The FASEB Journal*, **17**, 321–323.

Marinov, G. K., Williams, B. A., McCue, K., Schroth, G. P., Gertz, J., Myers, R. M., & Wold, B. J. (2014). From single-cell to cell-pool transcriptomes: Stochasticity in gene expression and RNA splicing. *Genome Research*, **24**(3), 496–510. https://doi.org/10.1101/gr.161034.113

McGinnis, C. S., Murrow, L. M., & Gartner, Z. J. (2019). DoubletFinder: Doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. *Cell systems*, **8**(4), 329–337.e4. https://doi.org/10.1016/j.cels.2019.03.003

Monaco, G., Lee, B., Xu, W., Mustadah, S., Hwang, Y. Y., Carre, C., Burdin, N., Visan, L., Ceccarelli, M., Poidinger, M., Zippelius, A., de Magalhaes, J. P., & Larbi, A. (2019). RNA-Seq signatures normalized by mRNA abundance allow absolute deconvolution of human immune cell types. *Cell Reports*, **26**, 1627–1640.

Montgomery, S. H., & Mank, J. E. (2016). Inferring regulatory change from gene expression: The confounding effects of tissue scaling. *Molecular Ecology*, **25**(20), 5114–5128. https://doi.org/10.1111/mec.13824

Morrison, J. K., DeRossi, C., Alter, I. L., Nayar, S., Giri, M., Zhang, C., Cho, J. H., & Chu, J. (2021). Single-cell transcriptomic profiling of healthy and fibrotic adult zebrafish liver reveals conserved cell identities and stellate cell activation phenotypes with human liver. *bioRxiv*. https://doi.org/10.1101/2021.08.06.455422

Murat, F., Mbengue, N., Winge, S. B., Trefzer, T., Leushkin, E., Sepp, M., Cardoso-Moreira, M., Schmidt, J., Schneider, C., Mößinger, K., Brüning, T., Lamanna, F., Belles, M. R., Conrad, C., Kondova, I., Bontrop, R., Behr, R., Khaitovich, P., Pääbo, S., & Kaessmann, H. (2023). The molecular evolution of spermatogenesis across mammals. *Nature*, **613**, 308–316.

Nip, K. M., Chiu, R., Yang, C., Chu, J., Mohamadi, H., Warren, R. L., & Birol, I. (2020). RNA-Bloom enables reference-free and reference-guided sequence assembly for single-cell transcriptomes. *Genome Research*, **30**(8), 1191–1200. https://doi.org/10.1101/gr.260174.119

Oliva, M., Muñoz-Aguirre, M., Kim-Hellmuth, S., Wucher, V., Gewirtz, A. D., Cotter, D. J., Parsana, P., Kasela, S., Balliu, B., Viñuela, A., Castel, S. E., Mohammadi, P., Aguet, F., Zou, Y., Khramtsova, E. A., Skol, A. D., Garrido-Martín, D., Reverter, F., Brown, A., & Stranger, B. E. (2020). The impact of sex on gene expression across human tissues. *Science*, **369**, eaba3066.

Papa, F., Windbichler, N., Waterhouse, R. M., Cagnetti, A., D'Amato, R., Persampieri, T., Lawniczak, M. K. N., Nolan, T., & Papathanos, P. A. (2017). Rapid evolution of female-biased genes among four species of *Anopheles* malaria mosquitoes. *Genome Research*, **27**, 1536–1548.

Parisi, M., Nuttall, R., Edwards, P., Minor, J., Naiman, D., Lu, J., Doctolero, M., Vanier, M., Chan, C., Malley, J., Eastman, S., & Oliver, B. (2004). A survey of ovary-, testis- and soma-biased gene expression in *Drosophila melanogaster* adults. *Genome Biology*, **5**, 1–18.

Parsch, J., & Ellegren, H. (2013). The evolutionary causes and consequences of sex-biased gene expression. *Nature Reviews Genetics*, **14**(2), 83–87. https://doi.org/10.1038/nrg3376

Perry, J. C., Harrison, P. W., & Mank, J. E. (2014). The ontogeny and evolution of sex-biased gene expression in *Drosophila melanogaster*. *Molecular Biology and Evolution*, **31**(5), 1206–1219. https://doi.org/10.1093/molbev/msu072

Pointer, M. A., Harrison, P. W., Wright, A. E., & Mank, J. E. (2013). Masculinization of gene expression is associated with exaggeration of male sexual dimorphism. *PLoS Genetics*, **9**(8), e1003697. https://doi.org/10.1371/journal.pgen.1003697

Powell, D. L., Payne, C., Banerjee, S. M., Keegan, M., Bashkirova, E., Cui, R., Andolfatto, P., Rosenthal, G. G., & Schumer, M. (2021). The genetic architecture of variation in the sexually selected sword ornament and its evolution in hybrid populations. *Current Biology*, **31**, 923–935.

Price, P. D., Palmer Droguett, D. H., Taylor, J. A., Kim, D. W., Place, E. S., Rogers, T. F., Mank, J. E., Cooney, C. R., & Wright, A. E. (2022). Detecting signatures of selection on gene expression. *Nature Ecology & Evolution*, **6**, 1035–1045.

Proschel, M., Zhang, Z., & Parsch, J. (2006). Widespread adaptive evolution of *Drosophila* genes with sex-biased expression. *Genetics*, **174**(2), 893–900. https://doi.org/10.1534/genetics.106.058008

Qian, P., Kang, J., Liu, D., & Xie, G. (2022). Single cell transcriptome sequencing of Zebrafish testis revealed novel spermatogenesis marker genes and stronger Leydig-germ cell paracrine interactions. *Frontiers in Genetics*, **13**. https://doi.org/10.3389/fgene.2022.851719

Qiao, Q., Le Manach, S., Sotton, B., Huet, H., Duvernois-Berthet, E., Paris, A., Duval, C., Ponger, L., Marie, A., Blond, A., Mathéron, L., Vinh, J., Bolbach, G., Djediat, C., Bernard, C., Edery, M., & Marie, B. (2016). Deep sexual dimorphism in adult medaka fish liver highlighted by multi-omic approach. *Scientific Reports*, **6**, 1–12.

Rago, A., Werren, J. H., & Colbourne, J. K. (2020). Sex biased expression and co-expression networks in development, using the hymenopteran *Nasonia vitripennis*. *PLoS Genetics*, **16**(1), e1008518. https://doi.org/10.1371/journal.pgen.1008518

Ranz, J. M., Castillo-Davis, C. I., Meiklejohn, C. D., & Hartl, D. L (2003). Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. *Science*, **300**(5626), 1742–1745. https://doi.org/10.1126/science.1085881

Reichard, A., & Asosingh, K. (2019). Best practices for preparing a single cell suspension from solid tissues for flow cytometry. *Cytometry. Part A*, **95**, 219–226.

Reznick, D., & Endler, J. A. (1982). The impact of predation on life history evolution in Trinidadian guppies (*Poecilia reticulata*). *Evolution*, **36**, 160–177.

Reznick, D. N. (1989). Life-history evolution in guppies: 2. Repeatability of held observations and the effects of season on life histories. *Evolution*, **43**(6), 1285–1297. https://doi.org/10.1111/j.1558-5646.1989.tb02575.x

Ribeiro, D. M., Ziyani, C., & Delaneau, O. (2022). Shared regulation and functional relevance of local gene co-expression revealed by single cell analysis. *Communications Biology*, **5**, 1–11.

Rose, E., Flanagan, S. P., & Jones, A. G. (2015). The effects of synthetic estrogen exposure on the sexually dimorphic liver transcriptome of the sex-role-reversed Gulf pipefish. *PLoS One*, **10**(10), e0139401. https://doi.org/10.1371/journal.pone.0139401

Satija, R., Farrell, J. A., Gennert, D., Schier, A. F., & Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology*, **33**(5), 495–502. https://doi.org/10.1038/nbt.3192

Sawyer, S. A., Parsch, J., Zhang, Z., & Hartl, D. L (2007). Prevalence of positive selection among nearly neutral amino acid replacements in *Drosophila*. *Proceedings of the National Academy of Sciences*, **104**(16), 6504–6510. https://doi.org/10.1073/pnas.0701572104

Sayadi, A., Martinez Barrio, A., Immonen, E., Dainat, J., Berger, D., Tellgren-Roth, C., Nystedt, B., & Arnqvist, G. (2019). The genomic footprint of sexual conflict. *Nature Ecology & Evolution*, **3**, 1725–1730.

Schartl, M., Kneitz, S., Ormanns, J., Schmidt, C., Anderson, J. L., Amores, A., Catchen, J., Wilson, C., Geiger, D., Du, K., Garcia-Olazábal, M., Sudaram, S., Winkler, C., Hedrich, R., Warren, W. C., Walter, R., Meyer, A., & Postlethwait, J. H. (2021). The developmental and genetic architecture of the sexually selected male ornament of swordtails. *Current Biology*, **31**, 911–922.

Sharma, E., Künstner, A., Fraser, B. A., Zipprich, G., Kottler, V. A., Henz, S. R., Weigel, D., & Dreyer, C. (2014). Transcriptome assemblies for studying sex-biased gene expression in the guppy, *Poecilia reticulata*. *BMC Genomics*, **15**, 1–21.

Squair, J. W., Gautier, M., Kathe, C., Anderson, M. A., James, N. D., Hutson, T. H., Hudelle, R., Qaiser, T., Matson, K. J. E., Barraud, Q., Levine, A. J., La Manno, G., Skinnider, M. A., & Courtine, G. (2021). Confronting false discoveries in single-cell differential expression. *Nature Communications*, **12**(1), 5692. https://doi.org/10.1038/s41467-021-25960-2

Standage, D. S., Berens, A. J., Glastad, K. M., Severin, A. J., Brendel, V. P., & Toth, A. L (2016). Genome, transcriptome and methylome sequencing of a primitively eusocial wasp reveal a greatly reduced DNA methylation system in a social insect. *Molecular Ecology*, **25**(8), 1769–1784. https://doi.org/10.1111/mec.13578

Stegle, O., Teichmann, S. A., & Marioni, J. C (2015). Computations and analytical challenges in single-cell transcriptomics. *Nature Reviews Genetics*, **16**(3), 133–145. https://doi.org/10.1038/nrg3833

Stuglik, M. T., Babik, W., Prokop, Z., & Radwan, J. (2014). Alternative reproductive tactics and sex-biased gene expression: the study of the bulb mite transcriptome. *Ecology and Evolution*, **4**(5), 623–632. https://doi.org/10.1002/ece3.965

Sun, Q., Peng, Y., & Liu, J. (2021). A reference-free approach for cell type classification with scRNA-seq. *Iscience*, **24**(8), 102855. https://doi.org/10.1016/j.isci.2021.102855

Sutherland, B. J., Prokkola, J. M., Audet, C., & Bernatchez, L. (2019). Sex-specific co-expression networks and sex-biased gene expression in the salmonid Brook Charr *Salvelinus fontinalis*. *G3: Genes, Genomes, Genetics*, **9**, 955–968.

Taboada, X., Robledo, D., Del Palacio, L., Rodeiro, A., Felip, A., Martínez, P., & Viñas, A. (2012). Comparative expression analysis in mature gonads, liver and brain of turbot (*Scophthalmus maximus*) by cDNA-AFLPS. *Gene*, **492**(1), 250–261. https://doi.org/10.1016/j.gene.2011.10.020

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Toubiana, W., Armisén, D., Viala, S., Decaras, A., & Khila, A. (2021). The growth factor BMP11 is required for the development and evolution of a male exaggerated weapon and its associated fighting behavior in a water strider. *PLoS Biology*, **19**(5), e3001157. https://doi.org/10.1371/journal.pbio.3001157

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N. J., Livak, K. J., Mikkelsen, T. S., & Rinn, J. L (2014). Pseudo-temporal ordering of individual cells reveals dynamics and regulators of cell fate decisions. *Nature Biotechnology*, **32**(4), 381–386. https://doi.org/10.1038/nbt.2859

Ullah, I., Shin, Y., Kim, Y., Oh, K. B., Hwang, S., Kim, Y. I., Lee, J. W., Hur, T. Y., Lee, S., & Ock, S. A. (2021). Effect of sex-specific differences on function of induced hepatocyte-like cells generated from male and female mouse embryonic fibroblasts. *Stem Cell Research & Therapy*, **12**, 1–14.

Veltsos, P., Fang, Y., Cossins, A. R., Snook, R. R., & Ritchie, M. G. (2017). Mating system manipulation and the evolution of sex-biased gene expression in *Drosophila*. *Nature Communications*, **8**, 1–11.

Wang, R., Zhang, P., Wang, J., Ma, L., E, W., Suo, S., Jiang, M., Li, J., Chen, H., Sun, H., Fei, L., Zhou, Z., Zhou, Y., Chen, Y., Zhang, W., Wang, X., Mei, Y., Sun, Z., Yu, C., ... Han, X. (2023). Construction of a cross-species cell landscape at single-cell level. *Nucleic Acids Research*, **51**, 501–506.

Wang, X., Yu, L., & Wu, A. R. (2021). The effect of methanol fixation on single-cell RNA sequencing data. *BMC Genomics*, **22**, 1–16.

Whittle, C. A., & Extavour, C. G. (2019). Selection shapes turnover and magnitude of sex-biased expression in Drosophila gonads. *BMC Evolutionary Biology*, **19**, 1–20.

Whittle, C. A., & Johannesson, H. (2013). Evolutionary dynamics of sex-biased genes in a hermaphrodite fungus. *Molecular Biology and Evolution*, **30**(11), 2435–2446. https://doi.org/10.1093/molbev/mst143

Yang, X., Schadt, E. E., Wang, S., Wang, H., Arnold, A. P., Ingram-Drake, L., Drake, T. A., & Lusis, A. J. (2006). Tissue-specific expression and regulation of sexually dimorphic genes in mice. *Genome Research*, **16**, 995–1004.

Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, **24**(8), 1586–1591. https://doi.org/10.1093/molbev/msm088

Yannai, J., Su, A. I., & Li, W. H. (2005). Gene expression evolves faster in narrowly than broadly expressed mammalian genes. *Molecular Biology and Evolution*, **22**(10), 2113–2118.

Zappia, L., & Oshlack, A. (2018). Clustering trees: A visualization for evaluating clusterings at multiple resolutions. *GigaScience*, **7**(7), giy083. https://doi.org/10.1093/gigascience/giy083

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**(7167), 233–237. https://doi.org/10.1038/nature06323

Zheng, G. X., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B., Wheeler, T. D., McDermott, G. P., Zhu, J., Gregory, M. T., Shuga, J., Montesclaros, L., Underwood, J. G., Masquelier, D. A., Nishimura, S. Y., Schnall-Levin, M., Wyatt, P. W., Hindson, C. M., & Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nature Communications*, **8**, 1–12.

Zheng, W., Xu, H., Lam, S. H., Luo, H., Karuturi, R. K. M., & Gong, Z. (2013). Transcriptomic analyses of sexual dimorphism of the zebrafish liver and the effect of sex hormones. *PLoS One*, **8**(1), e53562. https://doi.org/10.1371/journal.pone.0053562