

UC Irvine

UC Irvine Previously Published Works

Title

Comparative single-nucleus transcriptomics reveals asymmetric evolution of the *Drosophila* male and female germlines.

Permalink

<https://escholarship.org/uc/item/5p37v8r4>

Journal

PLoS biology, 24(7)

ISSN

1544-9173

Authors

Hariyani, Imtiyaz E

Das, Sudeshna

Le, Emma M

et al.

Publication Date

2026-07-01

DOI

10.1371/journal.pbio.3003869

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

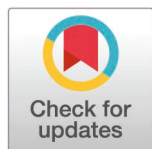
RESEARCH ARTICLE

Comparative single-nucleus transcriptomics reveals asymmetric evolution of the *Drosophila* male and female germlines

Imtiyaz E. Hariyani^{1*}, Sudeshna Das², Emma M. Le¹, Carmen Gamero-Castano¹, Tina Soroudi¹, Joshua Choi¹, Spring Momeni¹, Justin C. Kim¹, Rongying Lu¹, Vivek Swarup², José M. Ranz^{1,3*}

1 Department of Ecology and Evolutionary Biology, University of California Irvine, Irvine, California, United States of America, **2** Department of Neurobiology and Behavior, University of California Irvine, Irvine, California, United States of America, **3** Department of Systems Biology, University of California Irvine, Irvine, California, United States of America

* ihariyan@uci.edu (IEH); jranz@uci.edu (JMR)



OPEN ACCESS

Citation: Hariyani IE, Das S, Le EM, Gamero-Castano C, Soroudi T, Choi J, et al. (2026) Comparative single-nucleus transcriptomics reveals asymmetric evolution of the *Drosophila* male and female germlines. PLoS Biol 24(7): e3003869. <https://doi.org/10.1371/journal.pbio.3003869>

Academic Editor: Harmit S. Malik, Fred Hutchinson Cancer Research Center, UNITED STATES OF AMERICA

Received: November 18, 2025

Accepted: June 5, 2026

Published: July 20, 2026

Copyright: © 2026 Hariyani et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: Raw sequencing data, demultiplexed by sample for each sublibrary sequenced, were deposited as part of the NCBI BioProject PRJNA1273697. Further results are available in the supplementary files

Abstract

Reproductive organs vary widely across species yet share conserved cell types that produce gametes, sustaining species' perpetuation. However, tissue-level comparisons mask critical differences among cell types, obscuring where evolutionary divergence occurs even between closely related species. We quantified expression divergence at cell-type resolution between two sibling species, *Drosophila melanogaster* and *D. simulans*, while disentangling adaptive and nonadaptive evolutionary mechanisms. We built a comparative single-nucleus transcriptomic atlas of over 100,000 nuclei from testes and ovaries of both species. Our analysis revealed sharply heterogeneous divergence across testis cell types, contrasting with a broader conservation across ovary cell types. Notably, in both organs, ~40% of genes showing interspecific differences did so in only one cell type. In the testis, spermatogonia were largely conserved, whereas divergence peaked in primary spermatocytes with extensive rewiring of coexpression modules linked to microtubule and mitochondrial functions. In the ovary, expression was largely conserved, except in early germline and late follicle cells, which showed shifts in oogenesis and cell-cycle-related coexpression modules. Divergent cell types in both tissues were enriched for evolutionarily young genes with narrow expression breadth and faster protein evolution rates. Additionally, the ovary exhibited a faster-X effect consistent with adaptive evolution. These findings reveal a fundamental asymmetry in how male and female germlines evolve, with functional constraints relaxed in specific testis cell types but broadly maintained across the ovary. Our work provides an evolutionary framework explaining how core reproductive functions are safeguarded during species diversification while identifying germline cells that drive evolutionary change.

([Supporting information](#)), and the processed expression data are browsable in a web app at https://ranzlab-server.bio.uci.edu/Drosophila_snRNA-seq_Gene_Expression_Browser_Testis/ and https://ranzlab-server.bio.uci.edu/Drosophila_snRNA-seq_Gene_Expression_Browser_Ovary/. The custom scripts used for this study are available at Zenodo [147] (<https://doi.org/10.5281/zenodo.16622709>).

Funding: This work was supported by a National Science Foundation (<https://www.nsf.gov/>) Award to J.M.R. (MCB-2129845), and a National Institutes of Health (<https://www.nih.gov/>) grant to V.S. (5R01AG071683-03). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: ART, aligned rank transform; CC, cyst cells; C12choriogenic MBFCs stage 12C14choriogenic MBFCs stage 14DE, differentially expressed; DEGs, differentially expressed genes; DME, differential module eigengene; ECepithelial cellsEFC, early follicle cells; ES, early spermatogonia; EpSepithelium somaticESpermatocytesFCA, Fly Cell Atlas; FSC/preFCsfollicle stem cells and pre-follicle cellsGgermlineGO, gene ontology; GS, germarium somatic; G2a-2bgermarium region 2a and 2b cellsG2b-3germarium region 2b and 3 cellsGSC/G1-2agermline stem cells and germarium region 1 and 2a cellsH-Chub cellsLS, late spermatogonia; LSplate spermatocytesMBFCs, main-body follicle cells; Memeiotic/post-meioticMEs, module eigengenes; MFCmitotic follicle cells stage 1-5MimitoticMSpmid spermatocytesMPSpma-turing primary spermatocytesMSCI, meiotic sex chromosome inactivation; OoviductOSM, ovarian sheath muscle; PCA, Principal component analysis; PMFCpost-mitotic follicle cells stage 6ROIs, regions of interest; SsomaticSpermatidsSC, stretch cells; SPC, stalk and polar cells; TCLC, terminal corpus luteum cells; TE, transposable element; TOM, topological overlap matrix; UMAP, uniform manifold approximation and projection. UnunannotatedV7vitellogenic main-body follicle cells (MBFCs) stage 7V8vitellogenic MBFCs stage 8V9-10Avitellogenic MBFCs stage 9-10A.

Introduction

Gene expression changes are a key driver of species adaptation and phenotypic diversification [1–4]. While early comparative transcriptomic studies using microarrays or bulk RNA-sequencing helped characterize large-scale expression differences at the tissue-level, they lacked cell type resolution and were confounded by differences in tissue composition between species [5,6]. Single-cell and single-nucleus RNA sequencing (scRNA-seq and snRNA-seq, respectively) overcome these limitations by enabling direct comparisons of homologous cell types at unprecedented granularity within complex tissues [7,8]. These approaches have led to more precise evolutionary inferences—for example, in mammalian tissues such as the testis [5,9] and various regions of the brain [10,11]—by clarifying which cell types are most conserved versus those contributing disproportionately to expression divergence.

The gonads are central to understanding expression evolution, as they harbor most sex-biased genes [12] and exhibit distinct patterns of gene expression [13–15]. *Drosophila melanogaster* has served as a premier model organism for studying gametogenesis, primarily using bulk transcriptomic data, often derived from whole-body samples [16–21]. More recently, single-cell studies in *Drosophila* have revealed stage- and cell-type-specific expression patterns across both testis and ovary [7,22–30]. For example, DNA damage response genes are enriched in early germline stages of both gonadal tissues [22,27], and late spermatogenesis is marked by an enrichment of de novo genes [22]. However, reliance on single-strain datasets has limited our ability to capture the intraspecific genetic diversity of *D. melanogaster* [31–33], which is crucial for understanding functional constraints and identifying expression changes that truly contribute to species divergence [34,35]. Key questions remain unresolved: which gonadal cell types drive interspecific divergence? Do the testis and ovary follow parallel evolutionary trajectories, or do their constraints and modes of evolution fundamentally differ? What coexpression networks and functional pathways underlie this divergence? And do chromosomal location or genomic features explain putative cell-type-specific patterns?

Despite evolutionary divergence, gametogenesis in *Drosophila* and mammals shares conserved developmental features, including similar differentiation stages and core molecular pathways [36–38]. In mammals, transcriptome divergence is not uniform across the testis: it peaks in meiotic and post-meiotic germ cells of the testis, while early germline cells remain highly conserved—likely due to strong functional constraints tied to their pleiotropic roles [9,39,40]. In the mammalian ovary, the theca cells—endocrine somatic cells that surround the follicle—accumulate greater expression divergence between species compared to other ovarian somatic cell types [41]. Whether *Drosophila* exhibits similar patterns of stage- or cell-type-specific divergence in either gonadal tissue remains unknown. Comparative single-cell analyses of closely related *Drosophila* species remain scarce, making it difficult to discern whether previously reported expression divergence reflects true biological differences or methodological artifacts such as confounding or compensatory effects [13,16–18]. If substantial interspecific expression differences exist in *Drosophila*, it is unclear

whether they are broadly distributed across cell types or concentrated in terminally differentiated stages, similar to patterns observed in the mammalian testis.

To address these questions, we used snRNA-seq to compare the testis and ovary transcriptomes of three strains spanning two closely related species, *D. melanogaster* and *D. simulans*, which diverged ~1.4 mya [42]. We characterized expression divergence across cell types by integrating two complementary analyses: differential expression at the level of individual genes and network divergence at a systems level. We also examined how differentiation trajectories in spermatogenesis and oogenesis relate to pleiotropic constraint, using proxies such as expression breadth and phylogenetic gene age. Our results reveal a dichotomy in the tempo and mode of transcriptome evolution between the male and female germline. Testis transcriptomes evolve rapidly and in a highly cell-type-specific manner, with divergence in meiotic cells driven by a combination of adaptive and nonadaptive mechanisms. In contrast, the ovary exhibits more conserved expression properties, consistent with stronger pleiotropic constraint—except in a subset of germline and late-stage follicle cells, which show pronounced divergence. We further show that the cell types accumulating divergent genes also tend to have narrow expression breadth and younger phylogenetic age, especially in the testis, and that divergent genes are more likely to be on the X chromosome in the ovary. Together, these findings uncover how distinct evolutionary mechanisms shape gene expression programs across cell types in the male and female germlines.

Results

Combinatorial barcoding enables cell-type-resolved gonadal transcriptomes across *Drosophila* species

To characterize gene expression divergence between closely related *Drosophila* species at cell-type resolution, we performed single-nucleus RNA-sequencing (Parse Biosciences Evercode v2 [43]) on gonadal tissues from two strains of *Drosophila melanogaster* (A4, ISO1) and one strain of *D. simulans* (*w*⁵⁰¹), selected to capture intra- and interspecific divergence within the *D. melanogaster* species subgroup. We profiled testes from 1 to 2-day-old naïve males and ovaries from 3-day-old virgin females using split-pool combinatorial barcoding [43]. After quality control and doublet removal, we retained 106,328 high-quality nuclei—40,274 nuclei from testis and 66,054 nuclei from ovary—across 12 biological samples (3 strains × 2 tissues × 2 biological replicates). To minimize mapping artifacts due to nucleotide differences, we aligned sequencing reads to strain-specific genome assemblies. We then integrated cross-strain data for each tissue using Harmony [44], which best preserved cell-type and species structure among the tools tested (Methods). With an average of 44,742 reads per nucleus (S1 Table), our dataset is comparable to or exceeds those of previous studies (S2 and S3 Tables), enabling high-resolution comparisons of conserved and divergent expression patterns within and between species.

Cellular composition of *Drosophila* gonads is conserved across species

We identified conserved germline and somatic cell populations in both testis and ovary, whose proportions and marker expression were highly conserved across all three strains. Using uniform manifold approximation and projection (UMAP) and graph-based clustering, we found 10 testis and 17 ovary cell types and annotated them using known marker genes (S4 Table). In the testis, we identified cells undergoing mitosis (Mi; GSC/early spermatogonia (ES) and late spermatogonia (LS)) and meiosis (Me; early spermatocytes (ESp), mid spermatocytes (MSp), late spermatocytes (LSp), maturing primary spermatocytes (MPSp), and post-meiotic spermatids (S)), along with three somatic populations (S; epithelial (EC), cyst (CC), and hub (HC) cells). In the ovary, we distinguished early germline cells in the germarium (G; germline stem cells and germarium region 1 and 2a (GSC/G1-2a), germarium region 2a and 2b (G2a-2b), and germarium region 2b and 3 cells (G2b-3)) from somatic cells in the germarium (GS; follicle stem cells (FSC/pre-FCs), early follicle cells (EFC), and stalk and polar cells (SPC)), and multiple somatic populations in the epithelium (EpS; main body follicle cells (V7, V8, V9-10A, C12, C14), oviduct (O), ovarian sheath muscle (OSM), stretch cells (SC), and terminal corpus luteum cells (TCLC)).

Furthermore, marker gene expression remained remarkably consistent across strains for all annotated cell types (S4 and S5 Figs). Principal component analysis revealed that expression profiles clustered by species and by broad cell type category—mitotic, meiotic/post-meiotic, and somatic in the testis, and along a developmental trajectory continuum in the ovary (Fig 2A and 2B). Biological replicates consistently clustered together, and the two strains of *D. melanogaster* always grouped more closely with each other than with *D. simulans*, reflecting their phylogenetic relatedness and confirming the

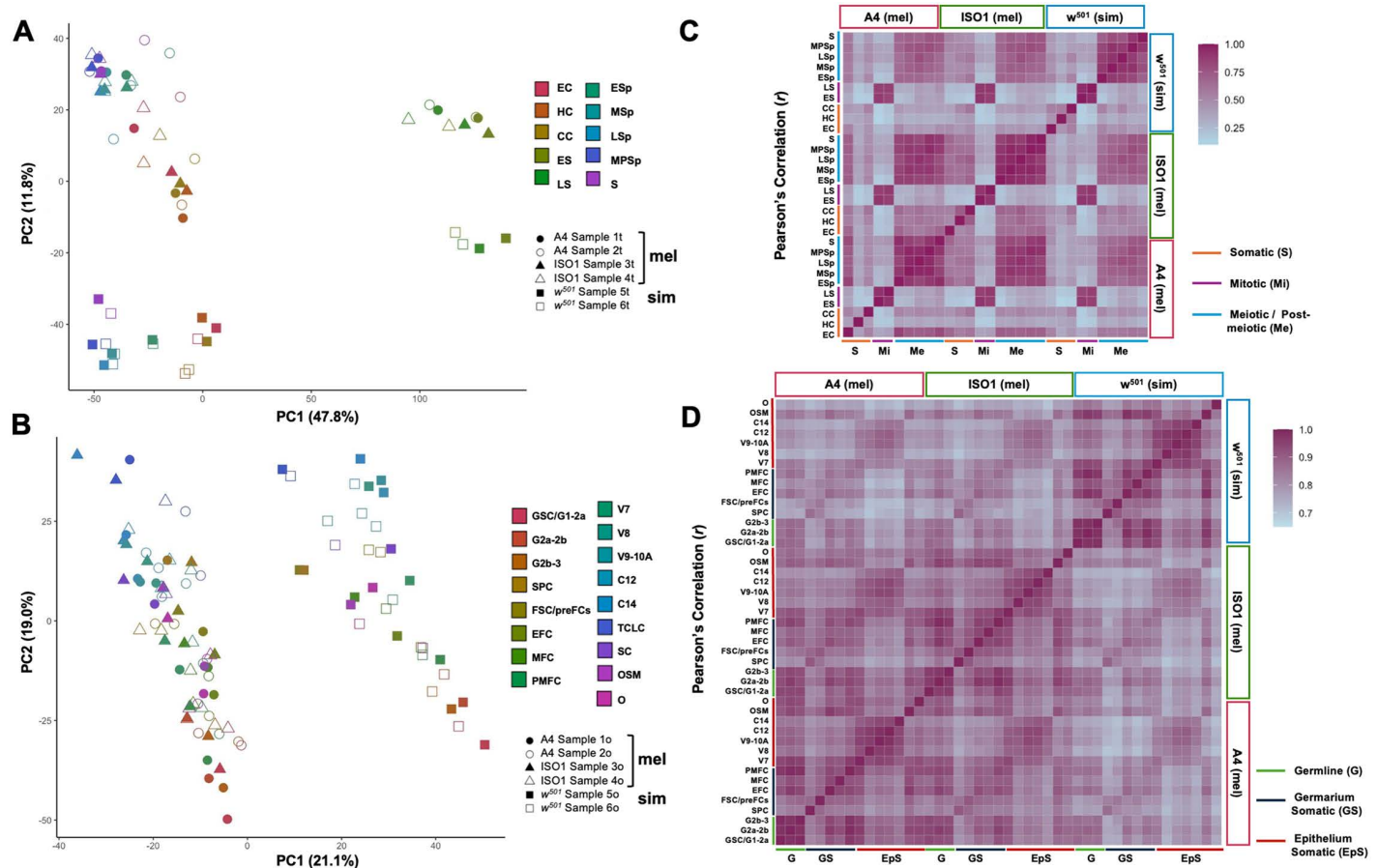


Fig 2. Intra- and interspecific expression variation and correlations across cell types in *Drosophila* gonads. Principal component analysis (PCA) of cell-type pseudo-bulks. **(A)** Testis. PC1 separates the mitotic cells from the meiotic and somatic cells. PC2 shows separation between species. **(B)** Ovary. PC1 shows the separation between species. PC2 shows cell progression along the oogenesis trajectory. Correlation matrices depicting Pearson's correlation coefficients between the expression levels of genes found expressed in common between any two cell types across the three strains assayed in **(C)** testis and **(D)** ovary. A4 and ISO1, *D. melanogaster* strains; w⁵⁰¹, *D. simulans* strain. In this correlation analysis, only cell types with at least 50 cells per strain and 350 cells across strains were included; as a result, the TCLC and SC ovarian cell types were excluded. Genes were required to be expressed in at least 1% of cells with a minimum average expression of 0.01. The order of cell types along both axes is identical, with broadly defined and precise cell types indicated on the x-axis and y-axis, respectively. Broad categories: somatic (S), mitotic (Mi), and meiotic (Me) categories for testis; and germine (G), germarium somatic (GS), and epithelium somatic (EpS) for ovary. Testis cell types: EC, epithelial cells; HC, hub cells; CC, cyst cells; ES, germline stem cells and early spermatogonia; LS, late spermatogonia; ES, early spermatocytes; MS, mid spermatocytes; LSp, late spermatocytes; MPSP, maturing primary spermatocytes; and S, spermatids. Ovary cell types: GSC/G1-2a, germline stem cells and germarium region 1 and 2a cells; G2a-2b, germarium region 2a and 2b cells; G2b-3, germarium region 2b and 3 cells; SPC, stalk and polar cells; FSC/preFCs, follicle stem cells and pre-follicle cells; EFC, early follicle cells; MFC, mitotic follicle cells stage 1-5; PMFC, post-mitotic follicle cells stage 6; V7, vitellogenic main-body follicle cells (MBFCs) stage 7; V8, vitellogenic MBFCs stage 8; V9-10A, vitellogenic MBFCs stage 9-10A; C12, choriogenic MBFCs stage 12; C14, choriogenic MBFCs stage 14; TCLC, terminal corpus luteum cells; SC, stretch cells; OSM, ovarian sheath muscle; and O, oviduct.

<https://doi.org/10.1371/journal.pbio.3003869.g002>

validity of our sample preparation and sequencing pipeline. Together, these findings indicate the conserved developmental architecture of gametogenesis, enabling cell-type-level comparisons of gene expression evolution across species.

Contrasting patterns of cell-type-specific expression divergence in the testis and ovary

The conserved cellular architecture of the gonads extended to the number of genes expressed per cell type, although we noted sharp within-tissue variation (S6 Fig). In the testis, late and maturing primary spermatocytes (LSp and MPSP, respectively), exhibited the highest median number of expressed genes, whereas mitotic cell types (ES and LS) had the fewest (Kruskal-Wallis chi-squared, $P < 0.001$ for each strain; S7 Table for post-hoc tests). In the ovary, germline stem cells and stage 1 and 2a cells in the germarium (GSC/G1-2a), and stage 12 choriogenic MBFCs (C12) expressed the highest median number of genes (Kruskal-Wallis chi-squared, $P < 0.001$ for each strain; S7 Table for post-hoc tests). These patterns remained stable across all three strains (S6 Fig).

Testis and ovary cell types displayed distinct trajectories of transcriptome differentiation. We quantified this differentiation at three levels: within strain, within species (A4 versus ISO1), and between species (*D. melanogaster* versus *D. simulans*). To do so, we calculated the average expression of each expressed gene for each gonadal cell type in each strain (S1 and S2 Data) and constructed gene expression correlation matrices for each tissue (Fig 2C and 2D). Within strains, median correlation values among broad cell-type categories differed significantly in both the testis and ovary, although this variation was larger in the testis (S7 Fig). This pattern indicates that transcriptome differentiation is more cell-type-specific in the testis than in the ovary (S8 Table). In the testis, median correlation values were significantly higher among mitotic cell types (Mi-Mi), among meiotic cell types (Me-Me), and between post-meiotic and meiotic categories (PMe-Me) (Kruskal-Wallis rank sum test; S7 Fig; S8 Table). In the ovary, the highest correlation values were found among germline cell types (G-G) (Kruskal-Wallis rank sum test; S7 Fig; S8 Table). When all pairwise cell type comparisons within each tissue were examined, within-strain median correlation values were significantly lower in the testis than in the ovary (two-way aligned rank transform (ART) ANOVA; S8 Fig; S9 Table).

Between strains, median expression correlations were also significantly higher among ovary cell types than among testis cell types, both within and between species (two-way ART ANOVA; S9 Fig; S10 Table). A closer examination by tissue revealed that the testis harbored the widest range of expression correlation values (min = 0.31, somatic cells, A4 versus *w⁵⁰¹*; max = 0.94, meiotic cells, A4 versus ISO1). Among broad cell-type categories, meiotic cells exhibited the largest and the only significant difference between intra- and interspecific contrasts. In contrast, mitotic cell types—germline stem cells/early spermatogonia (ES) and late spermatogonia (LS)—displayed similar median correlation values within and between species, suggesting a more conserved early transcriptional program. Somatic cell populations exhibited significantly lower correlation values than all other broad cell categories (two-way ART ANOVA; S10 Fig; S11 Table). By contrast, the ovary featured high intra- and interspecific correlation values (>0.75) across all broad cell-type categories, indicating more conserved expression patterns across both germline and somatic categories. Within each broad cell-type category, intraspecific median correlation values were significantly higher than interspecific ones ($P_{\text{adj}} < 0.05$). However, for any given pair of strains, median correlation values did not differ significantly among broad cell-type categories (two-way ART ANOVA; S11 Fig; S12 Table).

Together, these results show that expression correlations are highest within strain, lower within species, and lowest between species, but that this decline is highly tissue- and stage-specific. The ovary is more uniformly conserved, whereas the testis spans a much wider range of transcriptome differentiation, with meiotic cells showing the largest differences between intra- and interspecific comparisons.

Testis cell types harbor more intra- and interspecific expression differences than ovary cell types

To identify the genes driving patterns of expression correlation within and between species, we performed differential expression analyses, focusing on 11,481 one-to-one orthologous genes [48]. These analyses included 9,218 genes

expressed in testis cell types and 8,533 genes expressed in ovary cell types, with 8,084 genes expressed in both (S3 and S4 Data). We defined differentially expressed genes (DEGs) using a 1% FDR and a $\log_2$ fold-change $\geq |1|$, based on pairwise comparisons at both the intraspecific (ISO1 versus A4) and interspecific (ISO1 versus w^{501} , and A4 versus w^{501}) levels for each cell type and tissue. Consistent with previous work on the interspecific evolution of sex-biased expression [16,17,20,21], the testis showed greater expression divergence than the ovary (3,253 DEGs or 35.3% versus 1,748 DEGs or 20.5% of the expressed genes in each tissue, respectively; 2-sample test for equality of proportions with continuity correction or 2STEP, $\chi^2=479.2$, d.f. = 1, $P<2.2 \times 10^{-16}$) (Fig 3A and 3E).

The two interspecific contrasts revealed overlapping but nonidentical DEG sets (Fig 3B and 3F), with a slightly higher percentage shared between contrasts in the testis compared to the ovary (47.3% or 1,414/2,987 versus 43.3% or 738/1,705, respectively; 2STEP, $\chi^2=7.0$, d.f. = 1, $P=8.05 \times 10^{-3}$), a difference that remained when DEGs at the intra-specific level were omitted (52.7% or 1,227/2,328 versus 46.4% or 654/1,408, respectively; 2STEP, $\chi^2=13.5$, d.f. = 1, $P=2.4 \times 10^{-4}$). For both tissues, and akin to our previous correlation analysis, the number of DEGs varied across cell types (Friedman rank sum test; testis: $\chi^2=25.07$, d.f. = 9, $P=2.9 \times 10^{-3}$; ovary: $\chi^2=45.39$, d.f. = 16, $P=1.2 \times 10^{-4}$), peaking in the late and maturing primary spermatocytes (LSp and MPSP, respectively) in the testis, and in the germline stem cells and germarium region 1 and 2a (GSC/G1-2a) and 2a-2b (G2a-2b) in the ovary (Figs 3C, 3G and S12). This pattern sharply contrasts with the mitotic and somatic cell types in the testis, and most somatic cell types in the ovary (e.g., the follicle (FSC/pre-FCs and EFC) and oviduct (O) cell populations), which showed the lowest DEG counts between species. The number of DEGs per cell type was highly correlated across pairwise contrasts between strains (testis, minimum $r^2=0.78$, $P=7.7 \times 10^{-4}$; ovary, minimum $r^2=0.93$, $P=6.4 \times 10^{-10}$; S13 Table).

Expression differentiation across interspecific pairwise contrasts was also highly consistent at the cell type level. In the testis, 82.1% (1,007/1,227) of the DEGs were found in the same cell type across comparisons, with an average consistency of 63.7% across cell types, although varying significantly among them (test for equality of proportions, $\chi^2=191.85$, d.f. = 9, $P=2.2 \times 10^{-16}$; cyst cells (CC) and mid spermatocytes (MSP), the lowest and highest deviation relative to the random expectation; S14 Table) (Fig 3D). The ovary showed comparatively higher consistency (92.5%, 605/654; 2STEP, $\chi^2=37.07$, d.f. = 1, $P=1.14 \times 10^{-9}$), with an average of 74.5%, but also with significant differences across cell types (test for equality of proportions, $\chi^2=89.87$, d.f.=16, $P=2.6 \times 10^{-12}$; stretch cells (SC), and germarium 2a-2b (G2a-2b), the lowest and highest deviation relative to the random expectation; S14 Table) (Fig 3H). Across both tissues, about 40% of consistent DEGs—those differentially expressed in the same cell type for both interspecific contrasts—were differentially expressed in only one cell type (S13 Fig). In the case of differential expression in more than one cell type, species directionality (e.g., upregulation in *D. simulans* relative to *D. melanogaster*) was preserved across all affected cell types for nearly all DEGs (1,004 and 604 DEGs in the testis and ovary, respectively). Furthermore, we detected a slight but significant tendency toward upregulation in *D. melanogaster* relative to *D. simulans* for DEGs found only in testis, compared to those found only in ovary (two-tailed Fisher's exact test or FET, $P=0.033$, odds ratio=0.77). However, within each tissue, DEGs specific to one or shared between both tissues did not differ significantly in directional bias (FET, $P>0.05$) (S14 Fig).

A sizable fraction of DEGs at the intraspecific level displayed interspecific differences (Fig 3A and 3E). This overlap was significantly greater in the ovary (87.4% or 297/340) than in the testis (71.2% or 659/925; 2STEP, $\chi^2=34.08$, d.f. = 1, $P=5.28 \times 10^{-9}$). After excluding DEGs both at the intra and interspecific levels, we detected significantly more DEGs per cell type in the testis than in the ovary, and substantially fewer DEGs at the intraspecific level (2-way ART ANOVA, S15 Fig and S15 Table; and one subsequent one-sided permutation test for the counts of DEGs at the interspecific level between tissues due to acute overdispersion in testis, $P=0.048$). However, the global ratio of the number of interspecific to intraspecific DEGs was significantly higher in the ovary than in testis (14.07 versus 3.79; 2STEP, $\chi^2=63.64$, d.f. = 1, $P=1.5 \times 10^{-15}$). Because this ratio compares between-species divergence to standing within-species differentiation, lower values indicate that interspecific expression differences accumulate only modestly relative to intraspecific differences, which is suggestive of weaker functional constraints on the transcriptome program of a

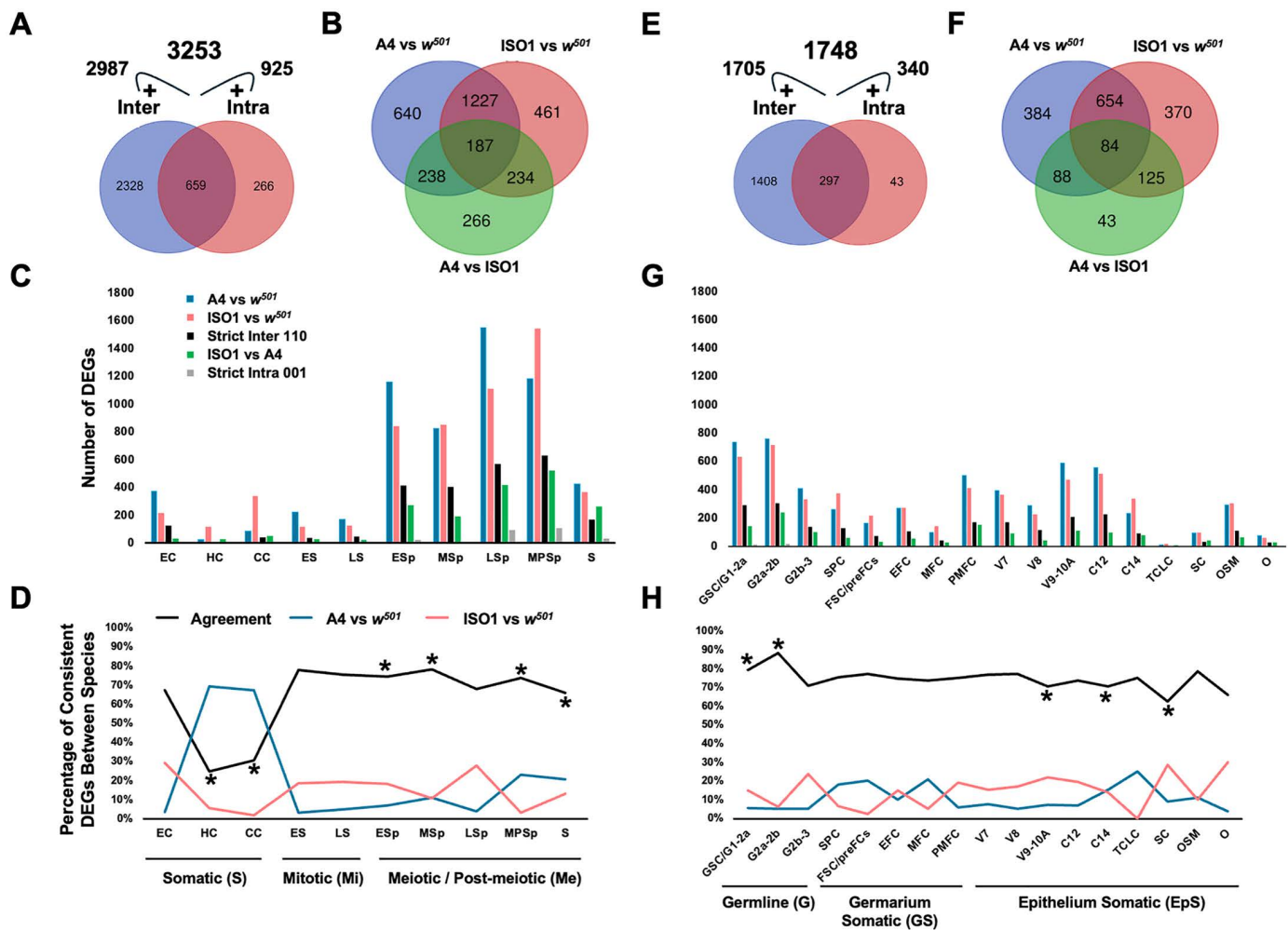


Fig 3. Magnitude of expression differentiation across testis and ovary cell types at two evolutionary timescales. Venn diagrams showing the number of differentially expressed genes (DEGs) unique to and shared between *D. melanogaster* and *D. simulans* in the testis: (A) at the inter- and intraspecific levels; and (B) across the three pairwise comparisons involving the *D. melanogaster* strains A4 and ISO1, and the *D. simulans* strain *w*⁵⁰¹. (C) Patterns of differential expression across pairwise comparisons and testis cell types. Strict Inter (pattern 110): DEGs in both interspecific contrasts in the same cell type without being differentially expressed at the intraspecific level. Strict Intra (pattern 001): DEGs at the intraspecific level without being differentially expressed at the interspecific level. (D) Percentage of DEGs per testis cell type between the *D. melanogaster* strains A4 and ISO1, shown relative to the *D. simulans* strain *w*⁵⁰¹. DEGs are categorized based on whether they show consistent differential expression in both interspecific comparisons (A4 vs. *w*⁵⁰¹ and ISO1 vs. *w*⁵⁰¹) or are unique to only one of the comparisons. (E–H) Equivalent plots for the ovary. Genes were required to be expressed in at least 1% of cells with a minimum average expression of 0.01. Asterisks, cell types showing the lowest and highest deviation relative to the random expectation for DEGs consistent across both interspecific contrasts (S14 Table). Testis cell types: EC, epithelial cells; HC, hub cells; CC, cyst cells; ES, germline stem cells and early spermatogonia; LS, late spermatogonia; ESp, early spermatocytes; MSp, mid spermatocytes; LSp, late spermatocytes; MPSP, maturing primary spermatocytes; and S, spermatids. Ovary cell types: GSC/G1-2a, germline stem cells and germarium region 1 and 2a cells; G2a-2b, germarium region 2a and 2b cells; G2b-3, germarium region 2b and 3 cells; SPC, stalk and polar cells; FSC/preFCs, follicle stem cells and pre-follicle cells; EFC, early follicle cells; MFC, mitotic follicle cells stage 1-5; PMFC, post-mitotic follicle cells stage 6; V7, vitellogenic main-body follicle cells (MBFCs) stage 7; V8, vitellogenic MBFCs stage 8; V9-10A, vitellogenic MBFCs stage 9-10A; C12, choriogenic MBFCs stage 12; C14, choriogenic MBFCs stage 14; TCLC, terminal corpus luteum cells; SC, stretch cells; OSM, ovarian sheath muscle; and O, oviduct.

<https://doi.org/10.1371/journal.pbio.3003869.g003>

given tissue. Accordingly, the lower ratio observed in the testis denotes more relaxed functional constraints compared to the ovary. This ratio also varied substantially across cell types (S16 Fig). By explicitly incorporating intraspecific variation, our experimental design provides a baseline reference for interpreting interspecific divergence relative to

standing expression variation, yielding a more accurate assessment of the selective constraints operating collectively on transcriptome programs of different tissues.

The increased resolution of snRNA-seq also uncovered fine-scale expression dynamics. For example, the gene *bol*, which encodes an RNA-binding protein essential for meiosis and spermatid differentiation [49,50], showed stable expression across strains and cell types (Fig 4A), consistent with the action of stabilizing selection. Conversely, the gene *eIF4G1*, which encodes a translation initiation factor involved in spermatogenesis [51], was downregulated in both *D. melanogaster* strains relative to *D. simulans* (Fig 4B), but only in late and maturing primary spermatocytes and spermatids (LSp, MPSp, and S, respectively). This pattern is compatible with the action of lineage-specific selection, although additional intraspecific expression data from both species and outgroup species are needed to substantiate this possibility. The patterns of expression for both genes, *bol* and *eIF4G1*, were independently validated by smFISH (S17 and S18 Figs; S5 and S6 Data), confirming the conserved expression of *bol* across cell types, and the downregulation of *eIF4G1* in a cell-type-specific manner. Equivalent patterns were observed for key genes during oogenesis such as *orb* and *ovo* [52,53] (S19 Fig). Some genes showed more complex patterns of gene expression evolution, diverging in one or both *D. melanogaster* strains relative to *D. simulans*, with or without intraspecific changes (Fig 4C).

In total, we documented 1,452 DEGs consistently diverged across both interspecific contrasts in at least one common cell type in the testis or ovary. Only 11.0% of these DEGs were shared across both tissues (Fig 4D). For this subset, the expression directionality between *D. melanogaster* and *D. simulans* was largely consistent across tissues. Only 21 genes displayed opposing patterns: 13 upregulated in *D. melanogaster* ovaries but downregulated in testes, and 8 showing the reversed pattern. Collectively, our findings suggest that the testis transcriptional program is more evolutionarily labile than that of the ovary, while providing enhanced resolution into how intra- and interspecific expression differences manifest at the level of individual cell types.

Coexpression network modules diverge across species in a cell-type-specific manner

To determine whether gene expression divergence within individual cell types extended to broader coordinated changes at the network-level, we performed a coexpression network analysis (hdWGCNA [54]). Our analysis revealed cell-type-specific modules of coexpressed genes encompassing 3,700 genes in the testis, and 4,730 genes in the ovary (S16 Table; S7 and S8 Data). Roughly half of the constituent genes (1,642 in the testis and 2,655 in the ovary) were assigned to modules in more than half of the cell types, including known gametogenesis genes such as *polo* [55], *CycE* [56], and *mei-P26* [57]. Conversely, a smaller set of genes (396 genes in the testis; 320 in the ovary) was uniquely assigned to modules in 20% or fewer cell types, highlighting potential cell-type-specific roles (S20A Fig).

To assess the functionality of these coexpression networks, we identified hub genes, defined as the top 10 genes within each module ranked by their intramodular connectivity, kME [58]. These included known regulators of spermatogenesis and oogenesis, including *dj*, which is involved in mitochondrial differentiation within the flagellum during sperm individualization [59], *S-Lap8*, a protease-encoding gene crucial for spermatogenesis [60], and the previously mentioned *orb* and *ovo*, which are required for egg chamber formation and polarity establishment, and the activation of multiple maternally expressed genes during oocyte development, respectively [52,53]. As expected, hub genes were enriched for GO terms tied to key aspects of germline development and gamete function, including spermatid development (GO:0007286), sperm motility (GO:0097722), and cytoplasmic translation (GO:0002181) (S17 Table).

Next, we examined whether the coexpression modules showed species-specific expression patterns by conducting a differential module eigengene (DME) analysis [54]. We found that about half of the coexpression modules exhibited interspecific divergence across testis and ovary cell types, with at least one divergent module per cell type (S20B Fig; S16 Table). In the testis, divergent modules were enriched during the later stages of spermatogenesis, consistent with our findings that species-specific transcriptional differences are most pronounced in the terminally differentiating meiotic cell types. A similar but nonsignificant trend was observed in the ovary, where terminal cell types—defined as the main

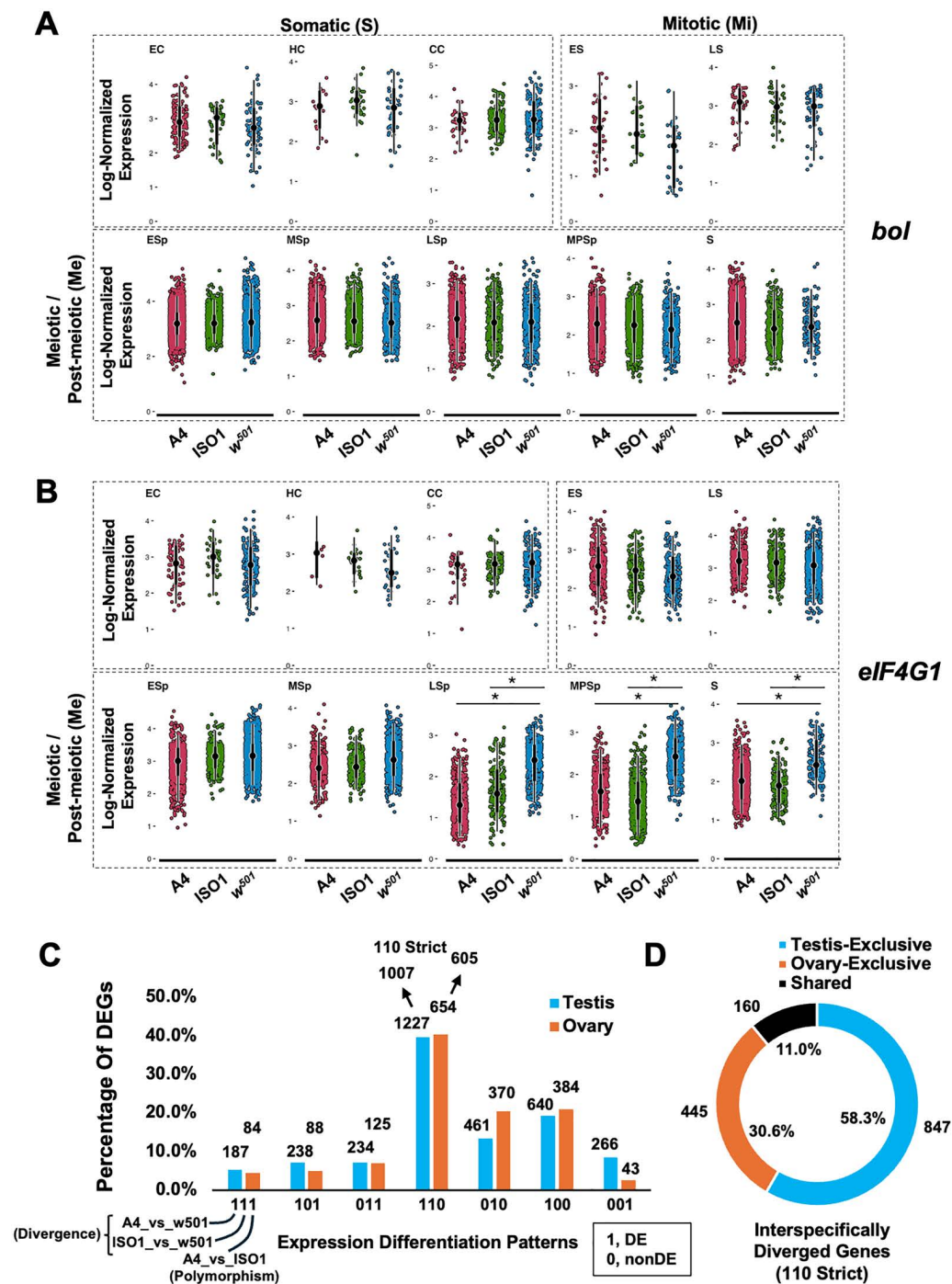


Fig 4. Expression differentiation patterns at the cell type level in *Drosophila*. (A) The gene *bol*, a non-DEG across testis cell types and strains, illustrates the stability in mRNA expression levels among strains, a pattern consistent with functional constraints. (B) The gene *eIF4G1* exemplifies a testis expressed gene that has diverged consistently at the interspecific level, but not at the intraspecific level. A4 and ISO1, *D. melanogaster* strains; *w*⁵⁰¹, *D. simulans* strain. Asterisks denote statistically significant differences in mRNA levels at 1% FDR and a log₂ fold-change ≥ |1|. (C) Histogram showing the percentage and the number (above bars) of DEGs in at least one cell type in at least one of the pairwise contrasts among the three strains. The combination of presence (1) or absence (0) of expression differences in each of the three pairwise contrasts results in seven possible patterns of expression differentiation. The categorization of each gene—including those showing no significant expression differences (000)—is provided in S3 and S4 Data. '110 Strict' refers to the subset of DEGs entailing expression changes in at least one same cell type in both interspecific contrasts; most DEGs following

a 110 pattern are part of this more stringent subset. **(D)** Donut chart showing the number and percentage of consistent DEGs at the interspecific level, exclusive to either testis or ovary, or shared between both tissues. Genes expressed in at least 1% of cells with a minimum average expression of 0.01 were considered. Testis cell types: EC, epithelial cells; HC, hub cells; CC, cyst cells; ES, germline stem cells and early spermatogonia; LS, late spermatogonia; ESp, early spermatocytes; MSp, mid spermatocytes; LSp, late spermatocytes; MPSP, maturing primary spermatocytes; and S, spermatids.

<https://doi.org/10.1371/journal.pbio.3003869.g004>

body follicle cells—were compared against all other cell types (terminal cell types versus the rest, 2STEP; testis, $\chi^2=4.62$, d.f.=1, $P=3.2 \times 10^{-2}$; ovary, $\chi^2=1.92$, d.f. = 1, $P=0.17$).

We further assessed the association between DEGs and the number of divergent coexpression modules across cell types. Under a random distribution of DEGs across modules, we expect that cell types with more divergent modules would also contain proportionally more DEGs, resulting in a positive correlation between these metrics. Alternatively, if DEGs are concentrated within a limited subset of divergent modules, no such proportional relationship would be observed, indicating uneven modular contributions to expression network differentiation. Our results indicate that divergence in the testis is broadly distributed across modules, whereas in the ovary, divergence is concentrated in a restricted subset of coexpression modules (Spearman's correlation coefficient; testis, $\rho=0.782$, $P=7.5 \times 10^{-3}$; ovary, $\rho=-0.084$, $P=0.766$; [S21 Fig](#)). Notably, in the testis, 13 DEGs within interspecifically divergent modules are also found in an ancient spermatocyte protein interaction network conserved from *Drosophila* to humans [\[61\]](#). None of the 13 genes was part of our list of top 1% hub genes, and only the knock-down of one of them resulted in impairment of male fertility [\[61\]](#). These properties suggest that even genes part of this metazoan male germ cell network can accommodate transcriptional divergence at short phylogenetic distances with no detrimental consequences.

The results of the DME and DEG analyses reinforced one another, providing a coherent picture of cell-type-specific expression divergence between *D. melanogaster* and *D. simulans*. For example, in late spermatocytes (LSp), the DME analysis revealed several divergent coexpression modules, with Module M3 and Module M1 showing a marked upregulation in *D. simulans* and in *D. melanogaster*, respectively ([Fig 5A](#)). Examination of DEG distributions within these modules showed that 29 constituent genes were upregulated in *D. simulans* in Module M3 ([Fig 5B](#), nodes in red), while none were downregulated, aligning with the overall pattern of network-level divergence. GO enrichment analysis of this module highlighted processes such as microtubule cytoskeleton organization (GO:0000226) and nuclear division (GO:0000280) ([Fig 5C](#); [S18 Table](#)). In contrast, module M1 contained 27 genes upregulated in *D. melanogaster* and none downregulated ([S22A Fig](#), nodes in green). This module was enriched for processes such as mitochondrial transport (GO:0006839), male gamete generation (GO:0048232), and spermatogenesis (GO:0007283) ([S18 Table](#); [S22B Fig](#)). Representative genes from both coexpression modules, such as the microtubule-associated gene *Klp10A* [\[62\]](#) and the somatic sex differentiation gene *janB* [\[63\]](#), further illustrate the cell-type-specific nature of these species-specific expression shifts ([Figs 5D and S22C](#)). As independently validated by smFISH, *Klp10A* exhibits upregulation in late spermatocytes (LSp) of *D. simulans*, while its expression remains conserved in nonmeiotic cell types such as late spermatogonia (LS) ([S23 Fig](#); [S9 Data](#)). This aligns with the canonical role of *Klp10A* as a centriole-length regulator in mitotic cells of the testis [\[62\]](#), while suggesting that it may perform an additional meiotic function in *D. simulans*. Such species-specific regulatory roles at the level of individual genes and entire subnetworks are suggestive of functional specialization between *D. melanogaster* and *D. simulans*.

Another example of cell-type-specific network-level divergence was observed in the germarium region 2a-2b cells (G2a-2b) of the ovary. The DME analysis uncovered key coexpression modules underlying early oogenesis divergence, with two modules, M1 and M2, displaying distinct coexpression patterns between species ([S24A–S24C Fig](#)). Module M1 was upregulated in *D. simulans* and was enriched for pathways related to germ cell development (GO:0007281) and oogenesis (GO:0048477) ([S19 Table](#); [S24D Fig](#)). A notable example is the gene *rhi* ([S24F Fig](#)), which encodes a germline-restricted heterochromatin protein essential for fertility and transposable element (TE) silencing [\[64\]](#), and proposed to be implicated in female hybrid sterility between *D. simulans* and *D. melanogaster* [\[65,66\]](#). Conversely, Module

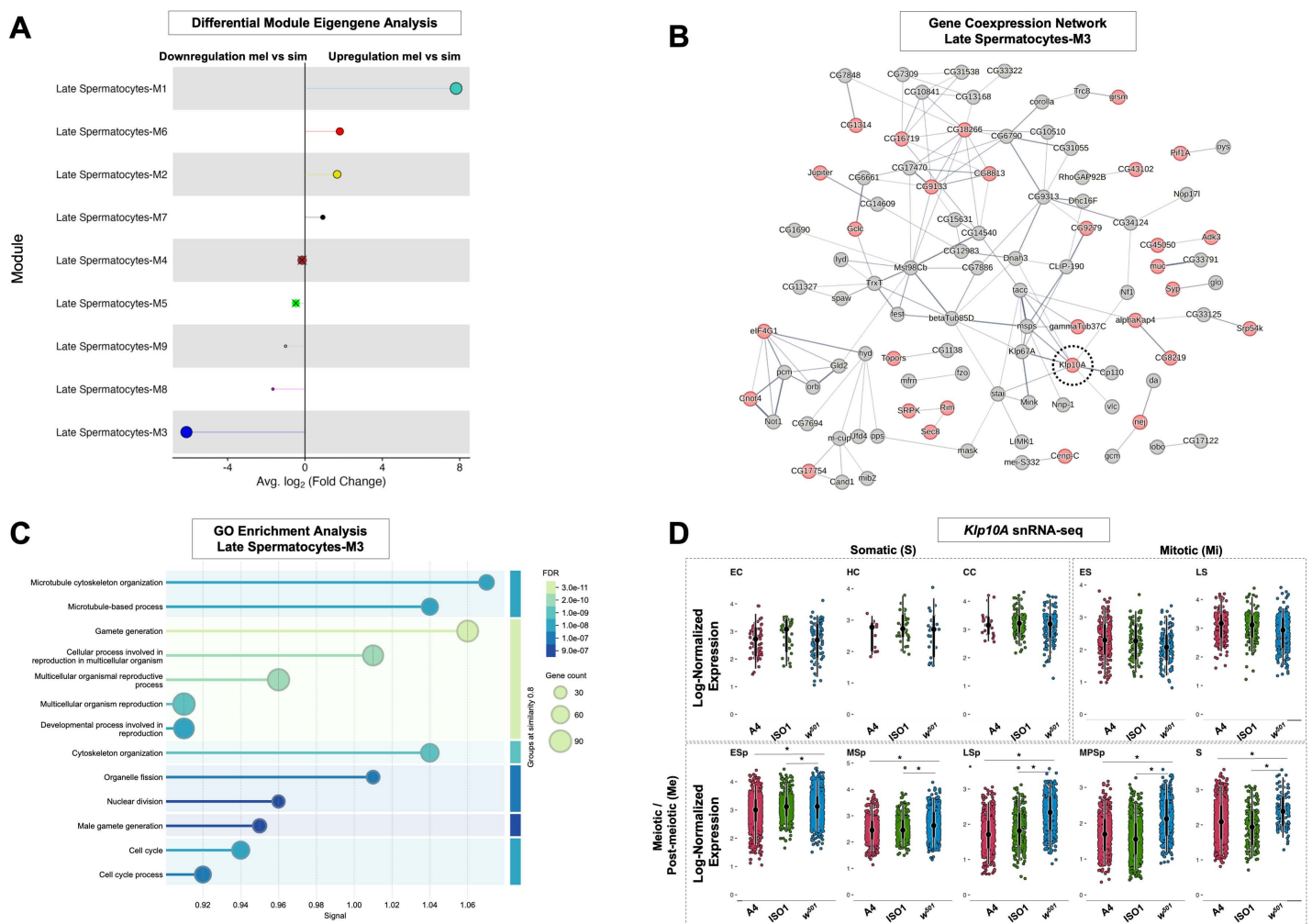


Fig 5. Divergent coexpression network in late spermatocytes between *D. melanogaster* and *D. simulans*. (A) Differential module eigengene (DME) analysis showing the average log₂ fold change in coexpression module activity between *D. melanogaster* and *D. simulans* in late spermatocytes (minimum module size = 100 genes). Module M3 is upregulated in *D. simulans*, while Module M1 is upregulated in *D. melanogaster*. (B) Gene coexpression network structure of the top 160 genes in Module M3 in late spermatocytes (medium confidence threshold = 0.4; FDR = 5%). Red nodes represent genes that are significantly downregulated in *D. melanogaster* relative to *D. simulans* according to the DEG analysis. Network edges denote the level of confidence, and disconnected nodes are omitted. *Klp10A*, a DEG part of this subnetwork, is highlighted with a dashed circle. (C) Gene Ontology enrichment results (Biological Process category) for the top 500 genes in Module M3. (D) Single-nucleus RNA-seq expression patterns of *Klp10A*, a representative gene from Module 3 (panel B), across testis cell types. The cell type of interest, late spermatocytes (LSp), is highlighted. Asterisks denote statistically significant differences in mRNA levels at 1% FDR and a log₂ fold-change ≥ |1|. Testis cell types: EC, epithelial cells; HC, hub cells; CC, cyst cells; ES, germline stem cells and early spermatogonia; LS, late spermatogonia; ESp, early spermatocytes; MSp, mid spermatocytes; LSp, late spermatocytes; MPSP, maturing primary spermatocytes; and S, spermatids.

<https://doi.org/10.1371/journal.pbio.3003869.g005>

M2 was overexpressed in *D. melanogaster* and predominantly enriched for cell-cycle (GO:0007049) and related GO terms. One of the hub genes in this module, *mei-W68*, mediates double-strand break formation during meiosis [67] (S19 Table; S24E and S24G Fig). Notably, DEGs within both divergent modules exhibited typical connectivity relative to non-DEGs, indicating that divergence is not confined to peripheral genes, particularly for M1 where DEGs were in fact significantly more connected than non-DEGs (Kruskal-Wallis rank sum test; M1: $\chi^2 = 17.94$, d.f. = 1, $P = 2.28 \times 10^{-5}$; M2: $\chi^2 = 1.11$, d.f. = 1, $P = 0.291$). These findings demonstrate species-specific divergence in germline differentiation and cell-cycle

regulation during oogenesis, supported by both DME and DEG analyses, which together reveal cell-type-specific transcriptional changes between *D. melanogaster* and *D. simulans*.

Functional constraint across gonadal cell types is stage-dependent in the testis but uniform in the ovary

The differences in expression divergence across cell types—particularly the greater evolutionary divergence of meiotic cells in the testis—suggest that functional constraints vary across gametogenesis. To test this hypothesis, we examined whether genes expressed in different cell types exhibit varying degrees of pleiotropy, given that pleiotropy is linked to the magnitude of functional constraint on a gene [68–70]. We used the tau index [71], calculated from FlyAtlas2 tissue-level expression data [72], to estimate expression breadth—a proxy for pleiotropy and functional constraint [68,69,73]—across genes expressed in different gonadal cell types. In the testis, tau values differed significantly across cell types (Kruskal-Wallis test, $P < 1 \times 10^{-100}$ for all three strains), with a stepwise increase from mitotic to meiotic stages (Fig 6A; S20 Table). In contrast, tau values in the ovary were lower overall and exhibited less variation across cell types. Among ovarian cell types, three somatic populations—stalk and polar cells (SPC), and choriogenic stage 12 (C12) and stage 14 (C14) MBFCs, the latter two involved in the late stages of oogenesis—showed the highest tau values (Fig 6B; S20 Table). Overall, the meiotic cells of the testis displayed the highest tau values of all gonadal cell types, suggesting weaker functional constraint, which is hypothesized to be associated with a more permissive chromatin environment found in these cell types [74–76].

We next examined expression breadth at the cell-type level within each tissue. This allowed us to ask whether tissue-level patterns were also observed at finer resolution. We repurposed the tau index to quantify cell-type specificity in our single-nucleus dataset for each tissue and strain. In the ovary, cell-type-level tau values recapitulated the patterns observed in the tissue-level analysis (S25 Fig). In the testis, however, the pattern was reversed: genes expressed in early spermatocytes, which showed high tau values in the tissue-level analysis, now displayed the lowest tau values when expression breadth was assessed across testis cell types. This indicates that although these genes are testis-biased at the tissue-level, they are broadly expressed across cell types within the testis, consistent with stronger functional constraint within that tissue (S21 Table; S25 Fig). These results suggest that patterns of functional constraint depend on the biological scale at which expression breadth is assessed and underscore the importance of accounting for within-tissue variation.

Given that expression breadth is also correlated with phylogenetic gene age [77], we hypothesized that young genes (age categories B–E, *i.e.*, genes that arose during the radiation of the genus *Drosophila*) would be enriched in late spermatogenesis stages, while older genes (*i.e.*, age category A, predating the radiation of the genus *Drosophila*) would be overrepresented in mitotic cell types. Using a precisely curated gene age categorization in flies [78], we found substantial support for these hypotheses (S26 Fig). In contrast, the ovary revealed a contrasting pattern with only the germarium region 2b–3 cells (G2b–3) showing enrichment for ancient genes, while the three aforementioned somatic cell types with higher tau exhibited a significantly elevated presence of younger genes from multiple age categories (S27 Fig).

The rate of protein evolution (ω , the ratio of nonsynonymous to synonymous substitutions) provides an additional proxy for the strength of functional constraint and is known to covary with expression divergence, particularly in sex-biased genes [16], which are predominantly expressed in germline tissues [12]. To calculate ω , we used divergence and polymorphism data from *D. simulans* and two *D. melanogaster* populations—one from the species' ancestral range and one from a derived, post-colonization population—to account for potential effects of differing demographic histories [79]. In the testis, genes expressed in meiotic cell types exhibited significantly higher ω values than those expressed in mitotic cell types (S22 Table; S28A Fig), which was consistent across all strains. In the ovary, ω was relatively uniform across cell types, with the exception of stage-12 choriogenic MBFCs (C12), which exhibited slightly elevated ω compared to most other cell types for all three strains (S23 Table; S28B Fig).

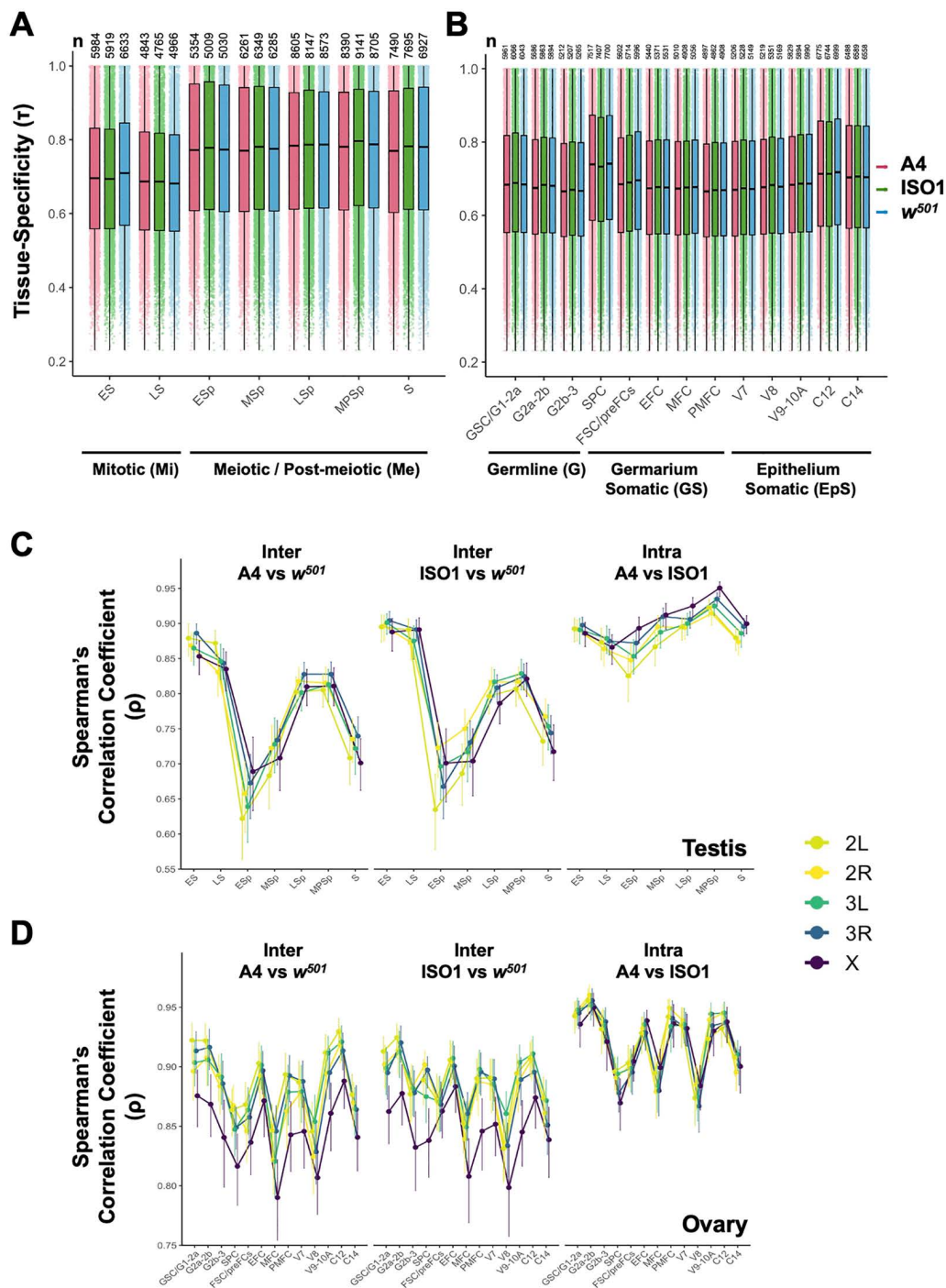


Fig 6. Expression breadth and chromosome expression correlation across testis and ovary cell types. (A, B) Tau index (τ) distributions, a measure of expression specificity, across testis and ovary, respectively, for all expressed genes in specific cell types. Higher τ values indicate more restricted expression across tissues, *i.e.*, a proxy for lower pleiotropy. Numbers above each boxplot denote the number of expressed genes per cell type. Only cell types directly involved in gametogenesis and with at least 100 cells per strain and 350 across all strains were included. Genes expressed in at least 1% of each cell type and with a minimum average expression of 0.01 are considered. Boxes represent the interquartile range around the median (black horizontal line) and whiskers extend to 1.5 times the interquartile range. Each point represents the tau value of a particular gene, and values are grouped by strain and cell type. A4 and ISO1, *D. melanogaster* strains; w^{501} , *D. simulans* strain. (C, D) Spearman's correlation coefficient (ρ) of gene expression between strains across cell types in the testis and ovary, respectively, calculated separately for each major chromosome arm (2L, 2R, 3L, 3R, and X).

Error bars represent the 95% confidence interval across replicates. Expression values above the 95th percentile within each cell type were omitted to reduce the impact of outliers in the expression correlation analyses.

<https://doi.org/10.1371/journal.pbio.3003869.g006>

To disentangle the contributions of adaptive and nonadaptive mechanisms to these differences, we partitioned ω into its adaptive (ω_a) and nonadaptive (ω_{na}) components [80–82]. In the testis, meiotic cells had elevated ω_a values, particularly in the early and mid-spermatocytes (ESp and MSp, respectively), suggesting a stronger role for positive selection (S22 Table; S28C Fig). In contrast, ω_{na} peaked in late spermatocytes (LSp), indicating relaxed purifying selection during the later stages of spermatogenesis (S22 Table; S28E Fig). In the ovary, both ω_a and ω_{na} were more uniform across cell types (S23 Table; S28D and S28F Fig). In both tissues, X-linked genes consistently displayed higher ω_a values than autosomal genes, suggesting a faster-X effect at the sequence level, driven by positive selection (S29A and S29B Fig). Since the median ω_a for autosomal genes was significantly lower relative to ω_{na} for all strains and tissues (paired, two-sided Wilcoxon signed-rank test, $V=0$, $P_{adj} < 0.02$ for all comparisons), the sequence evolution of autosomal genes is impacted more by nonadaptive than by adaptive mechanisms (S29C and S29D Fig).

Collectively, patterns of functional constraint differ starkly between the testis and ovary. The testis exhibits stage-dependent shifts in constraint across cell types, with early spermatocytes simultaneously showing the lowest tau (i.e., broader within-testis expression) and elevated rates of adaptive evolution (ω_a), suggesting a dual role in hosting core functions while enabling innovation. Genes expressed in later meiotic stages, in contrast, showed higher nonadaptive rates of sequence evolution (ω_{na}), suggesting more relaxed functional constraints. In the ovary, both expression breadth and evolutionary rates were more uniform across cell types, with only a few somatic populations (SPC, C12, C14) showing modest deviations and increased expression of younger genes. Genes on the X-chromosome showed consistently higher adaptive rates, indicating that chromosomal location is associated with differences in protein evolution of genes expressed in the gonadal tissues. Together, these results highlight the distinct evolutionary regimes operating in male and female gonads.

Genomic distribution of interspecifically divergent genes is biased in the female germline

The faster evolution of X-linked protein-coding genes expressed in the *Drosophila* gonads at the sequence-level raises the question of whether expression divergence also shows a faster-X effect in the testis and ovary. The hemizygosity of the X chromosome has been hypothesized to result in greater divergence at the coding and expression levels compared to autosomes [83]. Although the evidence is mixed [13, 19, 84–86], faster-X effects have been primarily reported for male-biased genes in expression [19, 85], many of which are preferentially expressed in the testis. We investigated how divergent genes between species at the cell type level were distributed across the genome (S30A Fig). Consistent DEGs between species exclusive to the ovary, or shared across both tissues, deviated from the random expectation (chi-squared goodness of fit test, $\chi^2=51.35$, d.f. = 4, $P_{adj}=1.88 \times 10^{-10}$, and $\chi^2=26.74$, d.f. = 4, $P_{adj}=2.24 \times 10^{-5}$, respectively). Post-hoc tests revealed that both categories of divergent genes were enriched on the X chromosome, with ovary-exclusive DEGs also significantly depleted from several autosomes (S24 Table; S30B Fig). In stark contrast, testis-exclusive DEGs showed a distribution not significantly different from the random expectation (chi-squared goodness of fit test, $\chi^2=5.96$, d.f. = 4, $P_{adj}=0.202$) (S30B Fig).

To test the faster-X effect more broadly by including both DEGs and non-DEGs, we analyzed the correlation between expression levels per chromosome and cell type [19, 84, 86]. Interspecific comparisons confirmed greater X-linked expression divergence (i.e., lower Spearman's correlation coefficients) in the ovary but not in the testis, mirroring the DEG distribution results (Fig 6C and 6D; S25 Table). At the intraspecific level, we found no differences across chromosomes, and all patterns were recapitulated at the pseudo-bulk level (S31A Fig).

Both adaptive and nonadaptive mechanisms can generate a faster-X effect [87]. To distinguish between these possibilities, we assessed the strength of purifying selection on gene expression by analyzing the correlation of gene expression

between the testis and ovary in each strain across chromosomes at the bulk-level [19]. Across strains and species, the X chromosome consistently exhibited higher testis-ovary correlation values than most autosomes (S31B Fig). These findings suggest that the pronounced expression divergence on the X is unlikely to result from relaxed functional constraints, instead pointing to positive selection as a more plausible driver of the observed faster-X effect.

Overall, our findings challenge traditional views from previous *Drosophila* studies, as we do not find evidence for a faster-X effect associated with gene expression divergence in testis [19,85,86]. The discrepancy between an observed faster-X effect at the coding-level and the absence of this effect at the expression level in testis has also been found in mammals [39,88]. The paradox may be explained by the unique biology of the testis, with permissive chromatin enabling widespread baseline expression, which is less sensitive to regulatory changes [75], combined with potential regulatory constraints associated with meiotic sex chromosome inactivation (MSCI) or a shutdown of dosage compensation in the meiotic cells [23,89]. The enrichment of DEGs exclusive to the ovary and those shared between tissues on the X chromosome points to a substantial, persistent role for positive selection acting on X-linked genes in females—the sex in which the X chromosome resides two-thirds of the time in any given generation—, presumably through dominant or partially dominant mutations [90,91].

Discussion

Our analysis of 106,328 testis and ovary nuclei from three *Drosophila* strains resolves cell type-specific patterns of transcriptional divergence at short phylogenetic distances. Traditional bulk transcriptomic studies in *Drosophila* lacked cell type resolution [13,16,17,20,92,93], limiting inferences about fine-scale expression differences. We identified numerous key genes involved in spermatogenesis (*Jupiter*, *janB*, and *eIF4G1* [63,94,95]) and oogenesis (*ovo*, *rhi*, and *mei-W68* [52,64,67]) that display interspecific expression differences in a cell-type-specific manner, extending similar observations from mammalian gonads [9] and other tissues [10,41,96–99]. Our results also reveal new insights into gene functionality. For example, *ovo*, known for its central role in oocyte development [52], shows species-specific expression in main-body follicle cells (MBFCs), suggesting a somatic role consistent with isoform-specific activity [100].

Transcriptome evolution during spermatogenesis is especially dynamic, with divergence concentrated in meiotic and post-meiotic germ cells. These stages harbor more DEGs and divergent coexpression networks than earlier stages, providing an increased opportunity for regulatory divergence. Our findings align with prior studies on sectioned testis from *D. melanogaster* and *D. simulans* [101] and mammalian snRNA-seq that identified late spermatogenesis as a hotspot of expression divergence [9]. In mammals, this has been linked to narrower spatiotemporal expression profiles and a greater prevalence of evolutionarily young genes [9], facilitated in part by a more permissive chromatin state [74–76]. These factors may allow both neutral and adaptive expression changes to accumulate—neutral changes due to weak or no purifying selection [102], and adaptive changes in response to different selective pressures, most notably sperm competition [76,103].

We found evidence for reduced functional constraints as spermatogenesis progresses in *Drosophila*. Phylogenetically younger genes with narrower expression breadth and faster rates of sequence evolution are more prevalently expressed at later stages, which also display lower ratios of interspecific to intraspecific expression changes. Consistent with these patterns, our network-level analyses identified mitochondrial transport and microtubule-based processes—both essential for germ cell development [104,105]—as major targets of transcriptome divergence in late spermatogenesis. This divergence may reflect species-specific adaptations shaped by sexual selection, particularly sperm competition [106]. Notably, sperm length—a key factor in sperm competition—depends on the microtubule cytoskeleton for elongation [105], a trait that differs markedly between *D. melanogaster* and *D. simulans* [107]. Together, these findings support a model of accelerated transcriptome evolution in late spermatogenesis, driven by a combination of adaptive and nonadaptive evolutionary mechanisms.

In the ovary, although overall divergence is more limited than in the testis (605 DEGs in the ovary versus 1,007 DEGs in the testis), cell-type-specific differences are still prominent. Genes expressed in the ovary tend to show more uniform

evolutionary properties across cell types—older age, broader expression, and slower sequence evolution—reflecting stronger pleiotropic properties and therefore stronger functional and evolutionary constraints. However, the global ratio of interspecific to intraspecific expression changes is significantly higher in the ovary than in the testis, perhaps reflecting a higher dependency of the ovary transcriptional program on transcription factor activity [75], and suggesting a more moderate role for nonadaptive mechanisms driving interspecific transcriptome divergence in the ovary. Among the most divergent cell types are terminally differentiating main-body follicle cells (MBFCs), which provide structural and regulatory support to the oocyte [108], and early-stage germline cells. For example, in germarium region 2a-2b germline cells (G2a-2b), which contain a mix of meiotic and nurse cells that have exited meiosis [109], we identified divergent coexpression modules associated with TE suppression, cell-cycle processes, and meiotic progression. This mirrors observations in mammals, where species-specific divergence in early oogenesis is linked to variation in cell-cycle timing [110], and is consistent with recent findings suggesting that the female meiotic germ cells may be especially evolutionarily dynamic [111]. The presence of meiotic gene expression in early germline cells of the germarium may help explain why divergence is detectable at these early stages [112]. These observations suggest that fine-scale regulation of mitotic-to-meiotic transitions—or meiosis itself—may be an important target of expression divergence in the ovary, as in the testis, with implications for oocyte quality and reproductive success. Furthermore, we find that X-linked genes show higher expression divergence and faster evolution at the protein-coding level than autosomal genes regardless of cell type, which seems to be primarily driven by adaptive mechanisms. Together, our findings suggest that the ovary evolves more slowly than the testis but is far from being transcriptionally static, undergoing cell-type-specific divergence even in core biological processes.

Overall, our findings reveal a striking asymmetry in the evolution of the male and female germlines. The testis exhibits stage-dependent expression divergence across cell types, with mitotic spermatogonia remaining highly conserved and meiotic spermatocytes emerging as a hotspot of divergence. In contrast, the ovary shows more constrained expression evolution along its developmental trajectory, with X-linked genes exhibiting elevated expression divergence that is consistent with a primary role of adaptive mechanisms. Across both tissues, gene age and pleiotropy are associated with the degree of cell-type-specific divergence, underscoring the influence of both adaptive and nonadaptive forces. These results support a more granular model of transcriptome evolution than previously recognized—one that reflects distinct biological demands and varying exposure to selection mechanisms [113,114]. Finally, we provide a publicly accessible resource for further exploring interspecific expression divergence during gametogenesis (https://ranzlab-server.bio.uci.edu/Drosophila_snRNA-seq_Gene_Expression_Browser_Ovary/ and https://ranzlab-server.bio.uci.edu/Drosophila_snRNA-seq_Gene_Expression_Browser_Testis/).

Materials and methods

Fly husbandry

Two strains of *D. melanogaster* (A4 and the reference strain ISO1; Bloomington *Drosophila* stock center ids #3852 [115] and 2057 [116], respectively), and the strain *w⁵⁰¹* of *D. simulans* (*Drosophila* species stock center #id: 14021-0251.195 [117]) were used. Flies were reared on a standard dextrose-cornmeal-yeast medium at room temperature (~25 °C) under continuous light. All manipulations of flies were performed under CO₂ anesthesia.

Gonadal tissue dissection

Testis dissections were performed on 1–2-day-old virgin males and ovary dissections on 3-day-old virgin females using fine-tipped forceps in cold Schneider's *Drosophila* Medium (ThermoFisher, #21720024). The selected age prevented the overrepresentation of mature sperm that characterizes older males [7] and ensured that females were fully fertile and receptive [118]. Dissected tissues were immediately transferred to 200 µL of Schneider's Medium in nuclease-free 1.5 mL Eppendorf tubes, pooled (10–40 tissues per tube), and kept on ice for less than 2 h. Separate petri dishes were used

for dissections of different strains and tissues to avoid cross-contamination and ensure sample integrity. After dissections were completed, samples were flash-frozen by submerging the tubes in liquid nitrogen and immediately stored at -80°C for long-term preservation. To maintain RNA integrity, samples were thawed only prior to nuclei isolation. Two biological replicates for each dissected tissue and strain were generated. In total, we dissected 1,200 pairs of testes (200 pairs per biological replicate) and 240 pairs of ovaries (40 pairs per biological replicate).

Nuclei isolation

Nuclei isolation from *Drosophila* ovary and testis was performed using a two-step dissociation process. Initially, tissues were incubated in HBSS Buffer containing 2.5 mg/mL collagenase (Invitrogen, #17018-029) at 37°C for 15 min. The tissues were then homogenized in EZ Lysis buffer (Sigma-Aldrich, #NUC101-1KT) and incubated on ice for 10 min, followed by filtration through a $70\text{ }\mu\text{m}$ filter. The filtered homogenate was centrifuged at 750 g for 5 min at 4°C and resuspended in 1 mL of lysis buffer. After a second centrifugation, the nuclei were incubated in Nuclei Wash and Resuspension buffer (1xPBS, 1% BSA, $0.5\text{U}/\mu\text{L}$ RNase inhibitor) for 5 min. To remove debris, a debris removal solution (Mitenyi Biotec, #130-109-398) was added to the nuclei suspension, and the mixture was centrifuged at $3,000\text{ g}$ for 10 min at 4°C . The purified nuclei were fixed and permeabilized using the Nuclei Fixation Kit (Parse Biosciences) and cryopreserved in DMSO until library preparation.

Preparation of snRNA-seq library and sequencing

Libraries were prepared using the EVERCODE WT V2 kit (Parse Biosciences) and quantified with the Qubit dsDNA HS assay kit (Invitrogen, #Q32851). The average fragment length of each sub-library was measured using the D5000 HS kit (Agilent, #5067-5592, #5067-5593). Finally, the eight sub-libraries were sequenced on the Illumina Novaseq 6000 S4 platform using paired-end sequencing at the UCI Genomics Research and Technology Hub, with an estimated sequencing depth of 50,000 read pairs/nuclei.

snRNA-seq read alignment

Raw sequencing data were assessed using FastQC (version 0.11.9) [119] to evaluate read quality, sequence composition, and potential sequencing artifacts. Read alignment was performed against the respective genome assembly of each *Drosophila* strain: ISO1 [120], A4 [32], and w^{501} [121]. To facilitate comparative genomic analysis, we employed LiftOff (version 1.6.3) [122] to annotate the A4 genome by mapping gene annotations from the ISO1 reference genome to the A4 assembly using default parameters. We filtered each annotation in the GTF file to retain protein-coding genes only using 10X Genomics' CellRanger mkgtf (version 8.0.1) [123] prior to creating the references using the 'mkref' mode of Parse Biosciences' split-pipe pipeline (version 1.1.2) [124]. Reads from each of the eight sub-libraries were demultiplexed by strain and aligned independently to the three respective references using the 'all' mode of split-pipe. The alignments were configured to incorporate each sample's designated wells, in addition to specifying the V2 chemistry of the library preparation protocol and a post_min_map_frac of 0.01. Lastly, we combined the resulting alignments from the eight sub-libraries for each sample using the 'comb' mode of split-pipe to generate the final count matrices.

Quality control, normalization, and integration

We considered previously delineated orthologous calls between *D. melanogaster* to *D. simulans* [48] and replaced the *D. simulans* gene names with those from *D. melanogaster* while retaining the genes without any one-to-one mapping. We combined the data from our two biological replicates to create one Seurat object for each tissue and strain, applying a stringent quality control criteria with the following thresholds: min.features = 200, min.cells = 3, and a minimum nCount_RNA of 300 for testis and 500 for ovary (version 5.1.0) [125]. We removed nuclei with mitochondrial reads above

5% and limited the dataset to nuclei with a maximum of 7,500 expressed genes and 100,000 counts to remove outliers, including nuclei whose gene counts are inflated by doublets/multiplets or other technical artifacts. These thresholds were implemented while inspecting violin and feature scatter plots to assess the distribution of the metrics across samples and using the Fly Cell Atlas (FCA) snRNA-seq dataset as a benchmark reference generated from comparable tissues [7]. Each of these datasets was then log-normalized (NormalizeData) and scaled (ScaleData) before obtaining a tissue- and strain-specific UMAP using Seurat's built-in functions (RunPCA, FindNeighbors, FindClusters, RunUMAP).

DoubletFinder (version 2.0.4) [126] was employed to minimize the potential effect of doublets by using the recommended iterative parameter sweeping approach to identify the optimal pK (a neighborhood-size parameter) value for each Seurat object. We assumed a default doublet formation rate of 3% (<https://support.parsebiosciences.com/hc/en-us/articles/360053107311-What-is-the-expected-doublet-rate>) and used a homotypic doublet proportion estimation based on initial clustering results. Nuclei classified as doublets were removed from downstream analyses.

The post-filtered Seurat objects were merged into a single object for each tissue, and normalization (NormalizeData) and variable feature identification (FindVariableFeatures) were performed for each layer independently. The cross-strain layers were integrated using Harmony (version 1.2.1) [44], which performed the best among other tools tested (LIGER [127], RPCA -https://satijalab.org/seurat/articles/integration_rpca.html-, CCA [128], scVI [129]), and is also recommended for closely related species, carrying out species mixing while preserving biological heterogeneity [130].

Clustering and cell type annotation

The integrated layers were subsequently joined for each tissue, and an ElbowPlot was used to determine the number of PCs for the FindNeighbors command for both testis (dims = 1:15) and ovary (dims = 1:20). We used clustree (version 0.5.1) [131] for interrogating the clusters at increasing resolution before running FindClusters on the testis (resolution = 1.2) and ovary (resolution = 1) Seurat objects (S32 Fig). We generated a combined UMAP for each tissue using the harmony reductions and visualized the plots using scpubR (version 2.0.2) [132]. Cell type annotations were inferred using known marker genes specifically expressed in each cell type (S4 Table; S4 and S5 Figs), as documented in recently published single-cell and single-nucleus RNA-seq studies of the gonadal tissue of *Drosophila* (testis [7,22–24]; ovary [7,25,26,28]).

Several control analyses were implemented to confirm expected species clustering and validate our cell type annotation. First, we performed a pseudo-bulk analysis and generated an MDS plot for each cell type of each tissue, confirming that biological replicates clustered by strain. Second, the proportions of different cell types were compared across strains using pairwise Pearson's correlations (S1 Text). Lastly, a trajectory analysis using Monocle3 (version 1.3.5) [133–135] was performed to demonstrate that the annotated cell types were arranged in the expected developmental sequence of gametogenesis for each tissue (S1 and S2 Figs).

Principal component analysis of cell-type pseudo-bulks

To examine expression differences within each tissue, we aggregated raw gene expression counts by cell type and sample to generate pseudo-bulk profiles. We combined these profiles into a single matrix with corresponding cell type and sample metadata, removed genes with zero counts across all pseudo-bulks, and applied variance-stabilizing transformation using the rlog function in DESeq2. We then performed principal component analysis on the transformed matrix and visualized the top two components to assess variation between cell types and strains (Fig 2A and 2B).

Correlation of gene expression

To assess the degree of gene expression correlation for the same cell type across *Drosophila* strains, and among different cell types within strains, we identified orthologous genes present in all three strains (ISO1, A4, and *w*⁵⁰¹). We extracted the

average gene expression levels of all 1-to-1 orthologous genes between *D. melanogaster* and *D. simulans* and generated a correlation matrix using Pearson correlation coefficients, which was visualized using a heatmap.

Differential gene expression analysis

We performed pairwise differential gene expression analyses for each interspecific (A4 versus w^{501} ; ISO1 versus w^{501}) and intraspecific (ISO1 versus A4) comparison for testis and ovary using MAST (version 1.28.0) [136]. Differential expression was evaluated for each gene and cell type by comparing expression levels between two given strains under the following criteria: $|\log_2 \text{fold change}| \geq 1$ and $\text{FDR} \leq 1\%$. We further corrected for multiple testing across the three comparisons using the Benjamini-Hochberg method ($\text{FDR} \leq 1\%$) and classified genes as differentially expressed (DE) or nondifferentially expressed (nDE). Genes that did not meet a minimum expression threshold of $\text{min.pct} = 0.01$ in a given cell type for both strains being compared were excluded from the analysis, thereby mitigating possible artifacts associated with low expression.

We adopted a systematic approach to characterize expression differentiation patterns across the three pairwise comparisons in the testis and ovary using custom in-house Python scripts. We implemented an encoding system where each gene received a three-digit binary code representing its differential expression status across these comparisons (A4 versus w^{501} ; ISO1 versus w^{501} ; ISO1 versus A4) for each tissue. Each digit can either be '0' for an nDE gene or '1' for a DE gene in at least one of the cell types considered (10 for the testis and 17 for ovary, respectively). Thus, interspecifically diverged genes had a code of '110', which was compared between testis and ovary to investigate tissue-specific gene expression evolution. A similar 10- or 17-digit trinary code (0, no difference; +1, overexpression in strain 1 versus strain 2; and -1, underexpression in strain 1 versus strain 2) was applied for summarizing patterns of differential expression across cell types within each comparison for the testis and ovary, respectively.

Weighted gene co-expression network analysis (hdWGCNA)

To identify modules of significantly co-expressed genes, we performed weighted gene co-expression network analysis (WGCNA) using the hdWGCNA framework (version 0.3.01) [54]. The analysis was carried out independently for each cell type, incorporating cells from all three strains, focusing exclusively on genes shared across strains and expressed in at least 5% of cells. Metacells were constructed by grouping cells with nearest-neighbor ($k = 25$) aggregation using the Harmony-reduced space. Metacells shared by more than 15 cells were excluded, groups with fewer than 50 cells were discarded, and the resulting metacell expression matrix was log normalized. The "signed" network type was used for optimizing the soft-thresholding power and maximize the scale-free topology fit. A topological overlap matrix (TOM) was computed for each cell type, and hierarchical clustering with dynamic tree cutting ($\text{deepSplit} = 4$, minimum module size = 25) was applied to identify co-expression modules. Module eigengenes (MEs) were calculated for each module and hub genes within each module were identified based on eigengene connectivity (kME) by ranking genes by kME. Modules assigned to the "gray" category, indicating unassigned genes, were excluded from downstream analyses.

To assess interspecies differences in module activity, we performed differential module eigengene (DME) analysis using hdWGCNA's 'FindDMEs' function. Module eigengene expression profiles were compared between *D. melanogaster* and *D. simulans*, applying $\text{min.pct} = 0.01$ to exclude lowly expressed genes. We identified significantly divergent modules for each cell type using the Wilcoxon rank-sum test and compared the constituent genes of these divergent modules with the differentially expressed genes to examine congruency between network-level and gene-level patterns of expression divergence.

Functional enrichment analysis

We analyzed gene ontology (GO) term enrichment among genes present in co-expressed gene modules for each cell type and tissue using clusterProfiler (version 4.0) [137] with genome-wide annotations from org.Dm.eg.db (version 3.18.0) [138]. Enrichment results were restricted to the top 500 genes per module and we applied a 5% FDR [139]. To identify

shared and cell-type-specific functional enrichments, we used the compareCluster function in clusterProfiler to compare GO term enrichment across modules.

Single-molecule RNA fluorescence in situ hybridization

We followed the smFISH protocol as described previously for the *Drosophila* testis [140]. Briefly, 10 pairs of testes from 1-day-old males were dissected on ice in 1×PBS and collected in 1.5 mL Eppendorf tubes. After centrifugation, PBS was removed, and tissues were fixed in 4% formaldehyde (in PBS) for 30 min on a nutator. Samples were then washed twice in PBS, transferred to 70% ethanol, and incubated overnight at 4 °C. Ethanol was removed and samples were briefly incubated in wash buffer A (2×SSC, 10% formamide). Hybridization was carried out overnight at 37 °C in hybridization buffer containing 2×SSC, 10% dextran sulfate, 1 mg/mL *E. coli* tRNA, 2 mM vanadyl ribonucleoside complex, 0.5% BSA, and 10% formamide, with 1 µL of 12.5 µM Stellaris probe solution targeting each focal gene (*bol*, *eIF4G1*, and *Klp10A*). On Day 3, without removing the hybridization solution, 1 mL of wash buffer A was added into the tube and incubated for 30 min at 37 °C. The tube was then centrifuged at 750 *g* for 5 mins, followed by a final wash in wash buffer B for 30 min at 37 °C. Samples were mounted in VECTASHIELD with DAPI (Vector Labs) and imaged immediately. Imaging was performed on a Nikon Ti2-U Inverted Fluorescence Microscope. All solutions used for RNA FISH were RNase-free. Stellaris RNA FISH probes were synthesized by LGC Biosearch Technologies. Probes were designed following blastn homology searches [141] to avoid nucleotide sequence differences between the open reading frames of *D. melanogaster* and *D. simulans* orthologs for the focal genes.

Quantification of smFISH signal was performed in Fiji/ImageJ [142]. The smFISH channel from each image was processed by background subtraction. Putative transcript puncta were identified using the *Find Maxima* function with a fixed prominence threshold of 40 for all images analyzed. Quantification was performed within rectangular regions of interest (ROIs) scaled to correspond to the same anatomical area across images. Guided by tissue morphology, these regions generally correspond to GSC/early spermatogonia (ES), late spermatogonia (LS), late spermatocytes (LSp), or spermatids (S). The number of detected puncta within each ROI was recorded and used for downstream comparisons of smFISH signal across strains and developmental stages.

Genomic features

We examined several functional and genomic features in the context of expression divergence between *D. melanogaster* and *D. simulans*. Chromosomal locations and sequence coordinates were retrieved from FlyBase [143]. For expression breadth, we obtained gene expression values across 31 tissues and body parts from adult individuals of both sexes as in FlyAtlas2 [72], and calculated the tissue-specific index tau (τ) [71]. The tau index ranges from 0 to 1 with lower and higher values indicating broader and narrower expression, respectively.

We retrieved the phylogenetic age of genes inferred under a parsimonious framework based on syntenic alignments among 20 fly species, including the species *Scaptodrosophila lebanonensis* and *Bactrocera dorsalis*, which are outgroups to the genus *Drosophila* [78]. The five age classes considered correspond to different tree branches in the original publication: class A corresponds to genes originated in branches −1 and −2, *i.e.*, those present before the *Drosophila* radiation; class B corresponds to branch 0, *i.e.*, those that arose prior to the split between the *Drosophila* and *Sophophora* subgenera; class C corresponds to branches 1 and 2, *i.e.*, to those genes originated in the lineages to the *D. willistoni* and *D. pseudoobscura* species groups; class D corresponds to branch 3, *i.e.*, to those genes originated during the radiation of the *D. melanogaster* species group; and class E corresponds to branches 4–6, *i.e.*, to those originated during the evolution of the *D. melanogaster* species subgroup, which includes *D. melanogaster* and *D. simulans*.

Gene sequence evolution was analyzed through a modified version of the McDonald and Kreitman test, the extended McDonald and Kreitman test [144], which controls for low frequency polymorphisms and differentiates neutral from weakly deleterious variants, thus minimizing the problem of underestimating α , *i.e.*, the proportion of substitutions fixed by positive

selection (α) [82]. Synonymous and nonsynonymous nucleotide substitutions were obtained from the comparison of sequences involving *D. simulans*, 197 lines from a *D. melanogaster* African population (Zambia), and 205 lines from a North American *D. melanogaster* population (Raleigh, North Carolina) [145], using the iMKT R package [144]. From the iMKT site, we also downloaded the allele frequency spectrum. To estimate the number of neutral segregating sites in the nonsynonymous category, we applied a minimum 5% frequency threshold to remove slightly deleterious mutations [144]. Alpha was calculated considering the corrected estimates, which allowed us to calculate the rate of protein evolution (ω), the rate of adaptive evolution (ω_a), and the rate of nonadaptive evolution (ω_{na}) [80,81].

Statistical analyses

Two-sample test for equality of proportions with continuity correction, two-tailed Fisher's exact test, chi-squared goodness-of-fit, chi-squared test of independence, analysis of residuals, Kruskal-Wallis rank sum, post-hoc tests, two-way ART (aligned rank transform) ANOVA, regression analysis, linear modeling, and Benjamini-Hochberg correction for multiple testing were conducted using built-in functions in R (version 4.3.2) [146,147].

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this manuscript, the authors used ZotGPT (GPT-4o, GPT-5, and Claude Sonnet 4.5) to improve readability and language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Supporting information

S1 Fig. Pseudotime projected onto the UMAPs of the testis and ovary datasets. UMAPs of the (A) testis and (B) ovary, colored by pseudotime values as inferred using Monocle 3 [135]. In both tissues, pseudotime recapitulates the expected developmental progression across gametogenesis, with early germline or progenitor populations occupying lower pseudotime values and more differentiated cell states distributed along later pseudotime regions of the embedding. The overall distribution of pseudotime is consistent with the annotated developmental hierarchy of testis and ovary cell types.
(TIFF)

S2 Fig. Pseudotime progression across annotated cell types of the testis and ovary inferred from developmental trajectories. (A, B) Boxplots show the distribution of pseudotime values for cells annotated as specific cell types, ordered vertically to reflect the inferred developmental progression for (A) testis and (B) ovary, based on Monocle3 analysis. In both tissues, pseudotime recapitulates known differentiation hierarchies, with early germline or stem cell populations occupying lower pseudotime values and differentiated somatic or germline cells positioned later.
(TIFF)

S3 Fig. Regression analysis of cell type composition across pairwise comparisons between strains. (A, B) Best-fit regression lines for the cell type proportions detected with snRNA-seq between A4 and two other strains (ISO1, blue; and w^{501} , orange) are shown with their corresponding equations, coefficients of determination, and statistical significance of the latter for the testis and ovary, respectively. ISO1 and A4, *D. melanogaster*; w^{501} , *D. simulans*. Cell clusters that were unannotated or did not fulfill the requirement of at least 50 nuclei per strain and 350 across the three strains (*i.e.*, terminal corpus luteum cells (TCLC) and stretch cells (SC) in the ovary), were not considered in the analysis. The diameter of each datapoint denotes the coefficient of variation (CV) for the proportion of a particular cell type across the three strains considered. The 95% confidence intervals are shaded.
(TIFF)

S4 Fig. Conserved marker gene expression in testis. Dot plots showing the average gene expression of marker genes in different testis cell types: **(A)** mitotic cells – spermatogonia (ES and LS); **(B)** somatic cells (HC, CC, and EC); **(C)** meiotic cells – spermatocytes (ESp, MSp, LSp, and MPSP); and **(D)** spermatids (S). The size of each dot represents the proportion of cells in which the gene is expressed. Gene expression values are normalized and scaled, with darker solid colors indicating overexpression relative to the average and lighter gray shades indicating underexpression. EC, epithelial cells; HC, hub cells; CC, cyst cells; ES, germline stem cells and early spermatogonia; LS, late spermatogonia; ESp, early spermatocytes; MSp, mid spermatocytes; LSp, late spermatocytes; MPSP, maturing primary spermatocytes; and S, spermatids.

(TIFF)

S5 Fig. Conserved marker gene expression in ovary. Dot plots showing the average gene expression of marker genes in different ovary cell types: **(A)** germline cells (G1-2a, G2a-2b and G2b-3); **(B)** germarium somatic cells (SPC, FSC/preFCs, EFC, MFC and PMFC); **(C)** epithelium somatic cells – main body follicle cells (V7, V8, V9-10A, C12 and C14); and **(D)** ovarian sheath muscle (OSM) and oviduct (O). The size of each dot represents the proportion of cells in which the gene is expressed. Gene expression values are normalized and scaled, with darker solid colors indicating overexpression and lighter gray shades indicating underexpression. Cell types: GSC/G1-2a, germline stem cells and germarium region 1 and 2a cells; G2a-2b, germarium region 2a and 2b cells; G2b-3, germarium region 2b and 3 cells; SPC, stalk and polar cells; FSC/preFCs, follicle stem cells and pre-follicle cells; EFC, early follicle cells; MFC, mitotic follicle cells stage 1–5; PMFC, post-mitotic follicle cells stage 6; V7, vitellogenic main-body follicle cells (MBFCs) stage 7; V8, vitellogenic MBFCs stage 8; V9-10A, vitellogenic MBFCs stage 9-10A; C12, choriogenic MBFCs stage 12; C14, choriogenic MBFCs stage 14; TCLC, terminal corpus luteum cells; SC, stretch cells; OSM, ovarian sheath muscle; and O, oviduct.

(TIFF)

S6 Fig. Number of genes expressed across cell types and strains in the testis and ovary. **(A, B)** Geyser plots showing the number of expressed genes detected across the cells within particular cell types for the three strains assayed (A4 and ISO1, *D. melanogaster*: w^{501} , *D. simulans*) in the testis and ovary, respectively. Broad categories of cell types for each tissue are indicated on top. Within each plot, the median, the 66% (thick line) and the 95% (thin line) of the data are shown. Testis cell types: EC, epithelial cells; HC, hub cells; CC, cyst cells; ES, germline stem cells and early spermatogonia; LS, late spermatogonia; ESp, early spermatocytes; MSp, mid spermatocytes; LSp, late spermatocytes; MPSP, maturing primary spermatocytes; and S, spermatids. Ovary cell types: GSC/G1-2a, germline stem cells and germarium region 1 and 2a cells; G2a-2b, germarium region 2a and 2b cells; G2b-3, germarium region 2b and 3 cells; SPC, stalk and polar cells; FSC/preFCs, follicle stem cells and pre-follicle cells; EFC, early follicle cells; MFC, mitotic follicle cells stage 1–5; PMFC, post-mitotic follicle cells stage 6; V7, vitellogenic main-body follicle cells (MBFCs) stage 7; V8, vitellogenic MBFCs stage 8; V9-10A, vitellogenic MBFCs stage 9-10A; C12, choriogenic MBFCs stage 12; C14, choriogenic MBFCs stage 14; TCLC, terminal corpus luteum cells; SC, stretch cells; OSM, ovarian sheath muscle; and O, oviduct.

(TIFF)

S7 Fig. Within-strain expression correlations differ significantly among broad cell-type categories in both the testis and ovary. Violin (foreground) and box (background) plots showing the distribution of within-strain expression correlations across contrasts between broad cell-type categories in the testis and ovary. Boxes represent the interquartile range (IQR) around the median (horizontal line), and whiskers extend to 1.5 times the IQR. Median correlation values, top; number of correlation values considered (n), bottom. The outcomes of the post hoc tests can be found in [S8 Table](#). Testis broad cell-type categories: Me, meiotic; Mi, mitotic; PMe, post-meiotic; and S, somatic. Ovary broad cell-type categories: EpS, epithelium somatic; G, germline; and GS, germarium somatic.

(TIFF)

S8 Fig. Within-strain expression correlations between pairs of cell types are lower in the testis than in the ovary.

Violin (foreground) and box (background) plots showing the distribution of within-strain expression correlation values between pairs of cell types in the testis and ovary across the three strains assayed. Boxes represent the interquartile range (IQR) around the median (horizontal line), and whiskers extend to 1.5 times the IQR. Median correlation values, top; number of correlation values considered (n), bottom. The outcomes of the post hoc tests can be found in [S9 Table](#). Strains: *D. melanogaster* (A4, ISO1); *D. simulans* (w^{501}). (TIFF)

S9 Fig. Between-strain expression correlations across equivalent cell types are higher within species than between species and are higher in the ovary than in the testis. Violin (foreground) and box (background) plots showing the distribution of between-strain expression correlation values for equivalent cell types in the testis and ovary across the three pairwise strain contrasts. Boxes represent the interquartile range (IQR) around the median (horizontal line), and whiskers extend to 1.5 times the IQR. Median correlation values, top; number of correlation values considered (n), bottom. The outcomes of the post hoc tests can be found in [S10 Table](#). Strains: *D. melanogaster* (A4, ISO1); *D. simulans* (w^{501}). (TIFF)

S10 Fig. Between-strain expression correlations across equivalent broad cell-type categories in the testis are largely similar within and between species and are lowest in somatic cells. Violin (foreground) and box (background) plots showing the distribution of between-strain expression correlation values across three pairwise strain contrasts and four broad cell-type categories in the testis. Boxes represent the interquartile range (IQR) around the median (horizontal line), and whiskers extend to 1.5 times the IQR. Median correlation values, top; number of correlation values considered (n), bottom. The outcomes of the post hoc tests can be found in [S11 Table](#). Testis broad cell-type categories: Me, meiotic; Mi, mitotic; PMe, post-meiotic; and S, somatic. Strains: *D. melanogaster* (A4, ISO1); *D. simulans* (w^{501}). T, testis. (TIFF)

S11 Fig. Between-strain expression correlations are similar across equivalent broad cell-type categories in the ovary, although higher within than between species. Violin (foreground) and box (background) plots showing the distribution of between-strain expression correlation values across three pairwise strain contrasts and three broad cell-type categories in the ovary. Boxes represent the interquartile range (IQR) around the median (horizontal line), and whiskers extend to 1.5 times the IQR. Median correlation values, top; number of correlation values considered (n), bottom. The outcomes of the post hoc tests can be found in [S12 Table](#). Ovary broad cell-type categories: EpS, epithelium somatic; G, germline; and GS, germarium somatic. Strains: *D. melanogaster* (A4, ISO1); *D. simulans* (w^{501}). O, ovary. (TIFF)

S12 Fig. Shared and unique patterns of differential expression across cell types in *Drosophila* testis and ovary. UpSet plots depicting differential expression patterns at different phylogenetic scales for testis and ovary comparisons. Testis: (A) *D. melanogaster* ISO1 versus A4 (intraspecific); (B) *D. melanogaster* ISO1 versus *D. simulans* w^{501} (interspecific); and (C) *D. melanogaster* A4 versus *D. simulans* w^{501} (interspecific). Ovary: (D) *D. melanogaster* ISO1 versus A4; (E) *D. melanogaster* ISO1 versus *D. simulans* w^{501} ; and (F) *D. melanogaster* A4 versus *D. simulans* w^{501} . A bar graph is plotted alongside each UpSet plot to depict the number of differentially expressed genes detected per cell type. Broad cell type categories are indicated: testis – somatic (S), mitotic (Mi), and meiotic (Me); and ovary – germline (G), germarium somatic (GS), and epithelium somatic (EpS). (TIFF)

S13 Fig. Scope of interspecific differentiation in gene expression across cell types of gonadal tissues. Fraction of differentially expressed genes (DEGs) between species, categorized by the number of different cell types in which both *D. melanogaster* strains (ISO1, A4) differ from *D. simulans* w^{501} . (TIFF)

S14 Fig. The directionality of interspecific expression for different patterns of differential expression and tissues. Up- and downregulation refer to significantly higher and lower expression levels in *D. melanogaster* versus *D. simulans*, respectively. Differentially expressed genes (DEGs) exclusively in testis, or ovary, do not tend to be more upregulated in *D. melanogaster* compared to *D. simulans* relative to DEGs in both tissues, as determined by Fisher's Exact Tests (p -values shown for each contrast). When focusing on DEGs exclusively in one tissue, there is some evidence of a higher tendency of upregulation in *D. melanogaster* relative to *D. simulans* for those detected only in testis versus those detected only in ovary.
(TIFF)

S15 Fig. Distributions of the number of differentially expressed genes across cell types in the testis and ovary for different strain comparisons. Violin (foreground) and box (background) plots showing the distribution of differentially expressed genes (DEGs) at the inter and intraspecific levels. At the intraspecific level, genes differentially expressed at the cell type level between the two strains of *D. melanogaster* (ISO1, A4) were considered. At the interspecific level, only genes consistently found differentially expressed in the same cell type of the two strains of *D. melanogaster* relative to one of *D. simulans* (w^{501}) were considered. Boxes represent the interquartile range (IQR) around the median (horizontal line), and whiskers extend to 1.5 times the IQR. Median DEG counts, top; number of cell types considered (n), bottom. The outcomes of the post hoc tests can be found in [S15 Table](#).
(TIFF)

S16 Fig. Ratio of interspecific to intraspecific differentiation in gene expression across testis and ovary cell types. (A, B) Testis and ovary, respectively. Ratio of the average number of DEGs in interspecific comparisons (*D. melanogaster* A4 versus *D. simulans* w^{501} , *D. melanogaster* ISO1 versus *D. simulans* w^{501}) to the number of DEGs between the two strains of *D. melanogaster* (A4 versus ISO1). Testis cell types: EC, epithelial cells; HC, hub cells; CC, cyst cells; ES, germline stem cells and early spermatogonia; LS, late spermatogonia; ES_p, early spermatocytes; MS_p, mid spermatocytes; LS_p, late spermatocytes; MPS_p, maturing primary spermatocytes; and S, spermatids. Ovary cell types: GSC/G1-2a, germline stem cells and germarium region 1 and 2a cells; G2a-2b, germarium region 2a and 2b cells; G2b-3, germarium region 2b and 3 cells; SPC, stalk and polar cells; FSC/preFCs, follicle stem cells and pre-follicle cells; EFC, early follicle cells; MFC, mitotic follicle cells stage 1–5; PMFC, post-mitotic follicle cells stage 6; V7, vitellogenic main-body follicle cells (MBFCs) stage 7; V8, vitellogenic MBFCs stage 8; V9-10A, vitellogenic MBFCs stage 9-10A; C12, choriogenic MBFCs stage 12; C14, choriogenic MBFCs stage 14; TCLC, terminal corpus luteum cells; SC, stretch cells; OSM, ovarian sheath muscle; and O, oviduct.
(TIFF)

S17 Fig. smFISH validation of *bol* expression levels across cell types and strains. (A) Representative smFISH images of *bol* mRNA expression in the testis from the strains A4, ISO1, and w^{501} . Pink and red dashed boxes mark the standardized regions of interest (ROIs) used for quantification in the GSC/Early Spermatogonia (ES) and Spermatids (S), respectively. Nuclei were marked with DAPI. Scale bars, 100 μ m. **(B)** smFISH puncta counts quantified within fixed-size ROIs in ES and S. Violin plots show the distribution across images, embedded boxplots show the median and interquartile range, and points represent individual values. Sample sizes are shown above each strain. Kruskal–Wallis tests were performed separately for ES and S. *bol* signal intensity did not differ significantly across strains in either cell type. The data underlying this figure are provided in [S5 Data](#).
(TIFF)

S18 Fig. smFISH validation of the expression levels in different cell types of the gene *eIF4G1* across strains. (A) Representative smFISH images of *eIF4G1* mRNA expression in the testis from the strains A4, ISO1, and w^{501} . Pink and yellow dashed boxes mark the standardized regions of interest (ROIs) used for quantification in Late Spermatogonia (LS)

and Late Spermatocytes (LSp), respectively. Nuclei were marked with DAPI. Scale bars, 100 μm . **(B)** smFISH puncta counts quantified within fixed-size ROIs in LS and LSp. Violin plots show the distribution across images, embedded boxplots show the median and interquartile range, and points represent individual values. Sample sizes are shown above each strain. Kruskal–Wallis tests were performed separately for LS and LSp, followed by pairwise two-sided Mann–Whitney tests with Benjamini–Hochberg correction for multiple tests in the case of LSp. *eIF4G1* signal intensity was similar across strains in LS, whereas in LSp it was significantly elevated in *w⁵⁰¹* relative to A4 and ISO1. The data underlying this figure are provided in [S6 Data](#).

(TIFF)

S19 Fig. Examples of expression differentiation among three *Drosophila* strains at the cell type level in ovary. (A)

The gene *orb* shows no differential expression across cell types or strains, suggesting functional constraints operating on mRNA abundance. **(B)** The gene *ovo* shows consistent upregulation in the two *D. melanogaster* strains compared to *D. simulans* in two cell types (V8, V9–10A), with no intraspecific difference. Cell types: GSC/G1–2a, germline stem cells and germarium region 1 and 2a cells; G2a–2b, germarium region 2a and 2b cells; G2b–3, germarium region 2b and 3 cells; SPC, stalk and polar cells; FSC/preFCs, follicle stem cells and pre-follicle cells; EFC, early follicle cells; MFC, mitotic follicle cells stage 1–5; PMFC, post-mitotic follicle cells stage 6; V7, vitellogenic main-body follicle cells (MBFCs) stage 7; V8, vitellogenic MBFCs stage 8; V9–10A, vitellogenic MBFCs stage 9–10A; C12, choriogenic MBFCs stage 12; C14, choriogenic MBFCs stage 14; TCLC, terminal corpus luteum cells; SC, stretch cells; OSM, ovarian sheath muscle; and O, oviduct.

(TIFF)

S20 Fig. Salient features of coexpression modules in *Drosophila* gonads. (A) Number of genes present across different coexpression modules. **(B)** Number of conserved and divergent coexpression modules between *D. melanogaster* and *D. simulans* per cell type. Coexpression modules were delineated with hdWGCNA [\[54\]](#).

(TIFF)

S21 Fig. The number of DEGs scales up with the number of divergent coexpression modules in the testis but not in ovary. (A, B) Testis and ovary. Divergent coexpression modules between *D. melanogaster* and *D. simulans* were delineated with hdWGCNA [\[54\]](#), while differential expression on a gene basis was determined independently at a 1% FDR and a $\log_2$ fold-change $\geq |1|$ for both interspecific contrasts, i.e., A4 versus *w⁵⁰¹* and ISO1 versus *w⁵⁰¹*. The Spearman's rho and its statistical significance are provided for each linear relationship.

(TIFF)

S22 Fig. *D. melanogaster*-biased coexpression module and gene expression divergence in late spermatocytes.

(A) Gene coexpression network structure of the top 160 genes in Module M1 in late spermatocytes (LSp; medium confidence threshold = 0.4; FDR = 5%). Green nodes represent genes that are significantly upregulated in *D. melanogaster* relative to *D. simulans* according to our differential expression analysis. Network edges denote the level of confidence and disconnected nodes are omitted. **(B)** Gene Ontology enrichment results (Biological Process category) for the top 500 genes in Module M1. **(C)** Single-nucleus RNA-seq expression patterns for one of the genes in this coexpression module, *janB*, across the meiotic testis cell types (also circled in (A)). This gene is upregulated in *D. melanogaster* relative to *D. simulans* in several cell types, including late spermatocytes (LSp). Asterisks denote statistically significant differences in mRNA levels at 1% FDR and a $\log_2$ fold-change $\geq |1|$. Testis cell types (c): ESp, early spermatocytes; MSp, mid spermatocytes; LSp, late spermatocytes; MPSP, maturing primary spermatocytes; S, spermatids.

(TIFF)

S23 Fig. smFISH validation of the expression levels in different cell types of the gene *Klp10A* across strains.

(A) Representative smFISH images of *Klp10A* mRNA expression in the testis from the strains A4, ISO1, and *w⁵⁰¹*. Pink

and red dashed boxes mark the standardized regions of interest (ROIs) used for quantification in Late Spermatogonia (LS) and Late Spermatocytes (LSp), respectively. Nuclei were marked with DAPI. Scale bars, 100 μm . **(B)** smFISH puncta counts quantified within fixed-size ROIs in LS and LSp. Violin plots show the distribution across images, embedded boxplots show the median and interquartile range, and points represent individual values. Sample sizes are shown above each strain. Kruskal–Wallis tests were performed separately for LS and LSp, followed by pairwise two-sided Mann–Whitney tests with Benjamini–Hochberg correction for LSp. *Klp10A* signal intensity was similar across strains in LS, whereas in LSp it was significantly elevated in *w⁵⁰¹* relative to A4 and ISO1. The data underlying this figure are provided in [S9 Data](#).

(TIFF)

S24 Fig. Gene coexpression network expression divergence between *D. melanogaster* and *D. simulans* in germarium region 2a–2b germline cells of the ovary. **(A)** Differential module eigengene (DME) analysis showing $\log_2$ (fold change) of coexpression module activity between species. Significant interspecific differences affect Modules M1 and M2 (≥ 100 genes). **(B, C)** Coexpression network of the top 160 genes in Modules M1 and M2 (medium confidence = 0.4, 5%FDR). Red nodes indicate downregulation in *D. melanogaster* (M1), while green nodes indicate upregulation in *D. melanogaster* (M2). Network edges indicate the level of confidence, and disconnected nodes are hidden. **(D, E)** GO enrichment analysis (Biological Process) for the top 500 genes in Modules M1 and M2. **(F, G)** Expression levels of the genes *rhi* (M1) and *mei-W68* (M2) across early germline and follicular ovary cell types; the genes are also indicated with blue and turquoise circles in (B) and (C), respectively. The cell type of interest, Germline-Germarium 2a–2b cells, is highlighted. Cell types: GSC/G1-2a, germline stem cells and germarium region 1 and 2a cells; G2a-2b, germarium region 2a and 2b cells; G2b-3, germarium region 2b and 3 cells; EFC, early follicle cells; and MFC, mitotic follicle cells stage 1–5.

(TIFF)

S25 Fig. Intra-tissue expression specificity across the testis and ovary cell types. **(A, B)** Cell-type tau distributions (τ), a measure of intra-tissue expression specificity, across testis and ovary, respectively, for all expressed genes by cell type. Higher τ values indicate more restricted expression across cell types within a given tissue. Only cell types directly involved in gametogenesis and with at least 100 cells per strain and 350 across all strains were included. Genes expressed in at least 1% of each cell type and with a minimum average expression of 0.01 are considered. Boxes represent the interquartile range around the median (black horizontal line) and whiskers extend to 1.5 times the interquartile range. Each point represents the cell-type tau value of a particular gene, and values are grouped by strain and cell type. A4 and ISO1, *D. melanogaster* strains; *w⁵⁰¹*, *D. simulans* strain.

(TIFF)

S26 Fig. Phylogenetic age distribution of genes expressed across testis cell types. Bar plots show the observed and expected (darker and lighter bars, respectively) percentages of expressed genes in each phylogenetic age class **(A–E)** across testis cell types for each strain (ISO1 and A4, *D. melanogaster*; *w⁵⁰¹*, *D. simulans*). Class A includes the oldest genes and Class E the youngest. Vertical arrows denote statistically significant deviations ($\uparrow$, enrichment; $\downarrow$, depletion) between observed and expected values based on a permutation test and corrected for multiple tests [139]. Numbers below each cell type indicate the observed counts of expressed genes per age class. The five phylogenetic age classes as defined by Dong and colleagues [78]: class A, genes present before the *Drosophila* radiation; class B, genes originated in the genus *Drosophila*; class C, genes originated in the *Sophophora* subgenus; class D, genes originated in the *melanogaster* species group; class E, genes originated in the *D. melanogaster* species subgroup comprising the *simulans* species complex and *D. melanogaster*. Cell types: EC, epithelial cells; HC, hub cells; CC, cyst cells; ES, germline stem cells and early spermatogonia; LS, late spermatogonia; ESp, early spermatocytes; MSp, mid spermatocytes; LSp, late spermatocytes; MPSP, maturing primary spermatocytes; and S, spermatids.

(TIFF)

S27 Fig. Phylogenetic age distribution of genes expressed across ovary cell types. Bar plots show the observed and expected (darker and lighter bars, respectively) percentages of expressed genes in each phylogenetic age class (A–E) across ovary cell types for each strain (ISO1 and A4, *D. melanogaster*; *w⁵⁰¹*, *D. simulans*). Class A includes the oldest genes and Class E the youngest. Vertical arrows denote statistically significant deviations ($\uparrow$, enrichment; $\downarrow$, depletion) between observed and expected values based on a permutation test and corrected for multiple tests [139]. Numbers below each cell type indicate the observed counts of expressed genes per age class. The five phylogenetic age classes as defined by Dong and colleagues [78]: class A, genes present before the *Drosophila* radiation; class B, genes originated in the genus *Drosophila*; class C, genes originated in the *Sophophora* subgenus; class D, genes originated in the *melanogaster* species group; class E, genes originated in the *D. melanogaster* species subgroup comprising the *simulans* species complex and *D. melanogaster*. Cell types: GSC/G1-2a, germline stem cells and germarium region 1 and 2a cells; G2a-2b, germarium region 2a and 2b cells; G2b-3, germarium region 2b and 3 cells; SPC, stalk and polar cells; FSC/preFCs, follicle stem cells and pre-follicle cells; EFC, early follicle cells; MFC, mitotic follicle cells stage 1–5; PMFC, post-mitotic follicle cells stage 6; V7, vitellogenic main-body follicle cells (MBFCs) stage 7; V8, vitellogenic MBFCs stage 8; V9-10A, vitellogenic MBFCs stage 9-10A; C12, choriogenic MBFCs stage 12; C14, choriogenic MBFCs stage 14; TLC, terminal corpus luteum cells; SC, stretch cells; OSM, ovarian sheath muscle; and O, oviduct.
(TIFF)

S28 Fig. Selection pressure on gene expression divergence at the cell type level. Box plots displaying three metrics of evolutionary change at sequence level of the genes expressed across cell types and strains (A4 and ISO1, *D. melanogaster*; *w⁵⁰¹*, *D. simulans*). The plots show: (A, B) the ratio of nonsynonymous to synonymous substitutions (ω , left); (C, D) the adaptive component of the ratio of protein evolution (ω_a , center); and (E, F) the nonadaptive component of the rate of protein evolution (ω_{na} , right). Testis, top; ovary, bottom. Boxes represent the interquartile range around the median (black horizontal line) and whiskers extend to 1.5 times the IQR. Each point represents the value of a particular gene, and values are grouped by strain and cell type. Gene counts for each combination (n) are shown above each boxplot. Testis cell types: EC, epithelial cells; HC, hub cells; CC, cyst cells; ES, germline stem cells and early spermatogonia; LS, late spermatogonia; ESp, early spermatocytes; MSp, mid spermatocytes; LSp, late spermatocytes; MPSp, maturing primary spermatocytes; and S, spermatids. Ovary cell types: GSC/G1-2a, germline stem cells and germarium region 1 and 2a cells; G2a-2b, germarium region 2a and 2b cells; G2b-3, germarium region 2b and 3 cells; SPC, stalk and polar cells; FSC/preFCs, follicle stem cells and pre-follicle cells; EFC, early follicle cells; MFC, mitotic follicle cells stage 1–5; PMFC, post-mitotic follicle cells stage 6; V7, vitellogenic main-body follicle cells (MBFCs) stage 7; V8, vitellogenic MBFCs stage 8; V9-10A, vitellogenic MBFCs stage 9-10A; C12, choriogenic MBFCs stage 12; C14, choriogenic MBFCs stage 14; TLC, terminal corpus luteum cells; SC, stretch cells; OSM, ovarian sheath muscle; and O, oviduct.
(TIFF)

S29 Fig. Adaptive component of selection pressure on gene expression divergence at the cell type level for X-linked and autosomal genes. Box plots displaying the adaptive component of the ratio of protein evolution (ω_a) at sequence level for the genes expressed across cell types and strains (A4 and ISO1, *D. melanogaster*; *w⁵⁰¹*, *D. simulans*) split by (A, B) X-linked and (C, D) autosomal genes in the testis (top) and ovary (bottom). Boxes represent the interquartile range around the median (black horizontal line) and whiskers extend to 1.5 times the IQR. Each point represents the ω_a value of a particular gene, and values are grouped by strain and cell type. Testis cell types: EC, epithelial cells; HC, hub cells; CC, cyst cells; ES, germline stem cells and early spermatogonia; LS, late spermatogonia; ESp, early spermatocytes; MSp, mid spermatocytes; LSp, late spermatocytes; MPSp, maturing primary spermatocytes; and S, spermatids. Ovary cell types: GSC/G1-2a, germline stem cells and germarium region 1 and 2a cells; G2a-2b, germarium region 2a and 2b cells; G2b-3, germarium region 2b and 3 cells; SPC, stalk and polar cells; FSC/preFCs, follicle stem cells and pre-follicle

cells; EFC, early follicle cells; MFC, mitotic follicle cells stage 1–5; PMFC, post-mitotic follicle cells stage 6; V7, vitellogenic main-body follicle cells (MBFCs) stage 7; V8, vitellogenic MBFCs stage 8; V9-10A, vitellogenic MBFCs stage 9-10A; C12, choriogenic MBFCs stage 12; C14, choriogenic MBFCs stage 14; TCLC, terminal corpus luteum cells; SC, stretch cells; OSM, ovarian sheath muscle; and O, oviduct.

(TIFF)

S30 Fig. Genomic distribution of interspecifically diverged genes between *D. melanogaster* and *D. simulans*. (A)

Chromosome distribution of DEGs between species. Only genes that showed consistent expression differences across the two interspecific comparisons (*i.e.*, A4 versus w^{501} and ISO1 versus w^{501}) for at least one same cell type and not located on the dot-like chromosome 4, are considered. DEGs only in testis, $n=842$, only in ovary, $n=431$, in both tissues, $n=155$. The karyotype was generated using the R package karyoploteR [148]. (B) Percentage of expected versus observed DEGs for each category (testis-exclusive, ovary-exclusive, shared by both tissues) across different chromosomes. The probability of finding a difference between both distributions was determined by a goodness-of-fit chi-squared test (in parenthesis). Post-hoc tests revealed which chromosome contributes to the global statistical difference found. P -values were adjusted for multiple tests correction [139].

(TIFF)

S31 Fig. Interstrain expression divergence and intersex expression correlation across chromosomes and tissues at the pseudo-bulk level. (A)

Pairwise Spearman's Rho correlations of gene expression across chromosomes are shown for three strain comparisons and two tissues, testis and ovary. In the ovary but not in the testis, there is evidence of faster-X evolution in gene expression (*i.e.*, lower correlation) in the two interspecific contrasts. For each graph and chromosome, the dot corresponds to the correlation value and the error bar is the 95% CI. In each graph, the genome-wide correlation is shown as a red dashed line while the red dotted lines correspond to the genome-wide 95% CI. Only the four rod-like autosomes (the left and right arms -L and R- of chromosomes 2 and 3) are considered. (B) Spearman's Rho correlations of gene expression across chromosomes between expression levels in testis and ovaries are shown for three strains. Gene expression on the X chromosome is not subject to more relaxed functional constraints compared to the autosomes as shown by the highest correlation values of the X versus the autosomes. Strains: A4 and ISO1, *D. melanogaster*; w^{501} , *D. simulans*.

(TIFF)

S32 Fig. Cell cluster relationships across clustering resolutions in the testis and ovary. Clustree plots show the relationships among clusters across increasing clustering resolutions (0–2.0, in increments of 0.2) for (A) testis and (B) ovary. Each node represents a cluster identified at a given resolution, and edges connect clusters across adjacent resolutions based on shared cells. These plots were used to assess cluster stability and subdivision across resolutions and to guide selection of an appropriate resolution for downstream annotation and analysis.

(TIFF)

S1 Text. Analysis of cell type proportions across strains.

(DOCX)

S1 Table. Salient stats of the snRNA-seq experiments performed in *Drosophila* gonadal tissues.

(XLSX)

S2 Table. Comparison across scRNA-seq studies in the testis.

(XLSX)

S3 Table. Comparison across scRNA-seq studies in the ovary.

(XLSX)

S4 Table. List of marker genes used to annotate testis and ovary cell types.

(XLSX)

S5 Table. Results of linear models to test for differences in cell type composition across pairwise comparisons of *Drosophila* strains.

(XLSX)

S6 Table. Correlation coefficients for the percentage of different cell types between any two *Drosophila* strains surveyed.

(XLSX)

S7 Table. Test for differences in the number of expressed genes between cell types across the strains and tissues considered.

(XLSX)

S8 Table. Test for differences in within-strain expression correlations across broad-cell type categories in the testis and ovary.

(XLSX)

S9 Table. Test for differences in within-strain expression correlations between different broad cell type categories across strains and tissues.

(XLSX)

S10 Table. Test for differences in between-strain expression correlation values within and between species across tissues.

(XLSX)

S11 Table. Test for differences across broad cell type categories in the testis and across between-strain contrasts.

(XLSX)

S12 Table. Test for differences across broad cell type categories in the ovary and across between-strain contrasts.

(XLSX)

S13 Table. Linear relationship metrics between the number of DEGs and different cell types across different pairwise comparisons.

(XLSX)

S14 Table. Test of differences in the proportion of differentially expressed genes between the strains of *D. melanogaster* and *D. simulans* that are shared at the cell type level across cell types within a given tissue.

(XLSX)

S15 Table. Test for differences in number of DEGs detected at the cell type level between tissues and type of contrast.

(XLSX)

S16 Table. Salient features of coexpression modules in testis and ovary.

(XLSX)

S17 Table. Gene ontology (GO) enrichment analysis of top 10 hub genes across coexpression modules of the testis and ovary.

(XLSX)

S18 Table. Gene ontology (GO) enrichment analysis of species-divergent coexpression modules of the testis.

(XLSX)

S19 Table. Gene ontology (GO) enrichment analysis of species-divergent coexpression modules of the ovary.

(XLSX)

S20 Table. Tests for differences in tissue specificity for the genes expressed between cell types and across strains and tissues.

(XLSX)

S21 Table. Tests for differences in cell-type specificity for the genes expressed between cell types across strains and tissues.

(XLSX)

S22 Table. Tests for differences in sequence evolution metrics of the genes expressed between testis cell types across the strains considered.

(XLSX)

S23 Table. Tests for differences in sequence evolution metrics of the genes expressed between ovary cell types across the strains.

(XLSX)

S24 Table. Test for deviations in the distribution of DEG genes between *D. melanogaster* and *D. simulans* across chromosomes.

(XLSX)

S25 Table. Test for differences in interstrains expression correlations per chromosome across testis and ovary cell types.

(XLSX)

S1 Data. Average expression of genes per cell type and strain in the testis.

(XLSX)

S2 Data. Average expression of genes per cell type and strain in the ovary.

(XLSX)

S3 Data. Differential expression of genes per cell type and strain in the testis.

(XLSX)

S4 Data. Differential expression of genes per cell type and strain in the ovary.

(XLSX)

S5 Data. Data underlying [S17 Fig](#).

(XLSX)

S6 Data. Data underlying [S18 Fig](#).

(XLSX)

S7 Data. Coexpression network analysis across testis cell types and strains.
(XLSX)

S8 Data. Coexpression network analysis across ovary cell types and strains.
(XLSX)

S9 Data. Data underlying [S23 Fig](#).
(XLSX)

Acknowledgments

We thank Alberto Civetta, Brandon Gaut, Ana Llopart, Anthony Long, Colin Meiklejohn, and Kyoichi Sawamura for comments on early versions of this manuscript. We also thank the Genomic High Throughput Facility and the Research Cyberinfrastructure Center at the University of California Irvine for facilitating our sequencing experiments and computational analyses, respectively.

Author contributions

Conceptualization: Imtiyaz E. Hariyani, José M. Ranz.

Data curation: Imtiyaz E. Hariyani, Sudeshna Das.

Formal analysis: Imtiyaz E. Hariyani, José M. Ranz.

Funding acquisition: Vivek Swarup, José M. Ranz.

Investigation: Imtiyaz E. Hariyani, Sudeshna Das, Emma M. Le, Carmen Gamero-Castano, Tina Soroudi, Spring Momeni, Justin C. Kim, José M. Ranz.

Methodology: Imtiyaz E. Hariyani.

Project administration: José M. Ranz.

Resources: Imtiyaz E. Hariyani, Joshua Choi, Vivek Swarup, José M. Ranz.

Software: Imtiyaz E. Hariyani, Rongying Lu.

Supervision: Imtiyaz E. Hariyani, Vivek Swarup, José M. Ranz.

Validation: Imtiyaz E. Hariyani, José M. Ranz.

Visualization: Imtiyaz E. Hariyani, José M. Ranz.

Writing – original draft: Imtiyaz E. Hariyani, José M. Ranz.

Writing – review & editing: Imtiyaz E. Hariyani, José M. Ranz.

References

1. Gompel N, Prud'homme B, Wittkopp PJ, Kassner VA, Carroll SB. Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature*. 2005;433(7025):481–7. <https://doi.org/10.1038/nature03235> PMID: [15690032](#)
2. Kratochwil CF, Liang Y, Gerwin J, Woltering JM, Urban S, Henning F, et al. Agouti-related peptide 2 facilitates convergent evolution of stripe patterns across cichlid fish radiations. *Science*. 2018;362(6413):457–60. <https://doi.org/10.1126/science.aao6809> PMID: [30361373](#)
3. Wang J, Ding J, Tan B, Robinson KM, Michelson IH, Johansson A, et al. A major locus controls local adaptation and adaptive life history variation in a perennial plant. *Genome Biol*. 2018;19(1):72. <https://doi.org/10.1186/s13059-018-1444-y> PMID: [29866176](#)
4. King MC, Wilson AC. Evolution at two levels in humans and chimpanzees. *Science*. 1975;188(4184):107–16. <https://doi.org/10.1126/science.1090005> PMID: [1090005](#)
5. Hunnicutt KE, Good JM, Larson EL. Unraveling patterns of disrupted gene expression across a complex tissue. *Evolution*. 2022;76(2):275–91. <https://doi.org/10.1111/evo.14420> PMID: [34882778](#)

6. Price PD, Palmer Drogue DH, Taylor JA, Kim DW, Place ES, Rogers TF, et al. Detecting signatures of selection on gene expression. *Nat Ecol Evol.* 2022;6(7):1035–45. <https://doi.org/10.1038/s41559-022-01761-8> PMID: 35551249
7. Li H, Janssens J, De Waegeneer M, Kolluru SS, Davie K, Gardeux V, et al. Fly Cell Atlas: a single-nucleus transcriptomic atlas of the adult fruit fly. *Science.* 2022;375(6584):eabk2432. <https://doi.org/10.1126/science.abk2432> PMID: 35239393
8. Han X, Wang R, Zhou Y, Fei L, Sun H, Lai S, et al. Mapping the mouse cell atlas by microwell-seq. *Cell.* 2018;172(5):1091–1107.e17. <https://doi.org/10.1016/j.cell.2018.02.001> PMID: 29474909
9. Murat F, Mbengue N, Winge SB, Trefzer T, Leushkin E, Sepp M, et al. The molecular evolution of spermatogenesis across mammals. *Nature.* 2023;613(7943):308–16. <https://doi.org/10.1038/s41586-022-05547-7> PMID: 36544022
10. Zemke NR, Armand EJ, Wang W, Lee S, Zhou J, Li YE, et al. Conserved and divergent gene regulatory programs of the mammalian neocortex. *Nature.* 2023;624(7991):390–402. <https://doi.org/10.1038/s41586-023-06819-6> PMID: 38092918
11. Suresh H, Crow M, Jorstad N, Hodge R, Lein E, Dobin A, et al. Comparative single-cell transcriptomic analysis of primate brains highlights human-specific regulatory evolution. *Nat Ecol Evol.* 2023;7(11):1930–43. <https://doi.org/10.1038/s41559-023-02186-7> PMID: 37667001
12. Ellegren H, Parsch J. The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev Genet.* 2007;8(9):689–98. <https://doi.org/10.1038/nrg2167> PMID: 17680007
13. Ranz JM, Go AC, González PM, Clifton BD, Gomes S, Jaberyzadeh A, et al. Gene expression differentiation in the reproductive tissues of *Drosophila willistoni* subspecies and their hybrids. *Mol Ecol.* 2023;32(13):3605–23. <https://doi.org/10.1111/mec.16941> PMID: 37000122
14. Papoli Yazdi H, Ravinet M, Rowe M, Saetre G-P, Guldvog CØ, Eroukhanoff F, et al. Extensive transgressive gene expression in testis but not ovary in the homoploid hybrid Italian sparrow. *Mol Ecol.* 2022;31(15):4067–77. <https://doi.org/10.1111/mec.16572> PMID: 35726533
15. Sylvestre E-L, Robert C, Pennetier S, Labrecque R, Gilbert I, Dufort I, et al. Evolutionary conservation of the oocyte transcriptome among vertebrates and its implications for understanding human reproductive function. *Mol Hum Reprod.* 2013;19(6):369–79. <https://doi.org/10.1093/molehr/gat006> PMID: 23340479
16. Zhang Y, Sturgill D, Parisi M, Kumar S, Oliver B. Constraint and turnover in sex-biased gene expression in the genus *Drosophila*. *Nature.* 2007;450(7167):233–7. <https://doi.org/10.1038/nature06323> PMID: 17994089
17. Assis R, Zhou Q, Bachtrög D. Sex-biased transcriptome evolution in *Drosophila*. *Genome Biol Evol.* 2012;4(11):1189–200. <https://doi.org/10.1093/gbe/evs093> PMID: 23097318
18. Müller L, Hutter S, Stamboliyska R, Saminadin-Peter SS, Stephan W, Parsch J. Population transcriptomics of *Drosophila melanogaster* females. *BMC Genomics.* 2011;12:81. <https://doi.org/10.1186/1471-2164-12-81> PMID: 21276238
19. Meisel RP, Malone JH, Clark AG. Faster-X evolution of gene expression in *Drosophila*. *PLoS Genet.* 2012;8(10):e1003013. <https://doi.org/10.1371/journal.pgen.1003013> PMID: 23071459
20. Ranz JM, Castillo-Davis CI, Meiklejohn CD, Hartl DL. Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. *Science.* 2003;300(5626):1742–5. <https://doi.org/10.1126/science.1085881> PMID: 12805547
21. Jiang Z-F, Machado CA. Evolution of sex-dependent gene expression in three recently diverged species of *Drosophila*. *Genetics.* 2009;183(3):1175–85. <https://doi.org/10.1534/genetics.109.105775> PMID: 19720861
22. Witt E, Benjamin S, Svetec N, Zhao L. Testis single-cell RNA-seq reveals the dynamics of de novo gene transcription and germline mutational bias in *Drosophila*. *Elife.* 2019. <https://doi.org/10.7554/eLife.47138> PMID: 31418408
23. Mahadevaraju S, Fear JM, Akeju M, Galletta BJ, Pinheiro MMLS, Avelino CC, et al. Dynamic sex chromosome expression in *Drosophila* male germ cells. *Nat Commun.* 2021;12(1):892. <https://doi.org/10.1038/s41467-021-20897-y> PMID: 33563972
24. Raz AA, Vida GS, Stern SR, Mahadevaraju S, Fingerhut JM, Viveiros JM, et al. Emergent dynamics of adult stem cell lineages from single nucleus and single cell RNA-Seq of *Drosophila* testes. *Elife.* 2023;12:e82201. <https://doi.org/10.7554/eLife.82201> PMID: 36795469
25. Jevitt A, Chatterjee D, Xie G, Wang X-F, Otwell T, Huang Y-C, et al. A single-cell atlas of adult *Drosophila* ovary identifies transcriptional programs and somatic cell lineage regulating oogenesis. *PLoS Biol.* 2020;18(4):e3000538. <https://doi.org/10.1371/journal.pbio.3000538> PMID: 32339165
26. Rust K, Byrnes LE, Yu KS, Park JS, Sneddon JB, Tward AD, et al. A single-cell atlas and lineage analysis of the adult *Drosophila* ovary. *Nat Commun.* 2020;11(1):5628. <https://doi.org/10.1038/s41467-020-19361-0> PMID: 33159074
27. Slaidina M, Gupta S, Banisch TU, Lehmann R. A single-cell atlas reveals unanticipated cell type complexity in *Drosophila* ovaries. *Genome Res.* 2021;31(10):1938–51. <https://doi.org/10.1101/gr.274340.120> PMID: 34389661
28. Slaidina M, Banisch TU, Gupta S, Lehmann R. A single-cell atlas of the developing *Drosophila* ovary identifies follicle stem cell progenitors. *Genes Dev.* 2020;34(3–4):239–49. <https://doi.org/10.1101/gad.330464.119> PMID: 31919193
29. Lee U, Li C, Langer CB, Svetec N, Zhao L. Comparative single-cell analysis of transcriptional bursting reveals the role of genome organization in de novo transcript origination. *Proc Natl Acad Sci U S A.* 2025;122(18):e2425618122. <https://doi.org/10.1073/pnas.2425618122> PMID: 40305051
30. Wei KH-C, Chatla K, Bachtrög D. Single-cell RNA-seq of *Drosophila* miranda testis reveals the evolution and trajectory of germline sex chromosome regulation. *PLoS Biol.* 2024;22(4):e3002605. <https://doi.org/10.1371/journal.pbio.3002605> PMID: 38687805
31. Chakraborty M, Emerson JJ, Macdonald SJ, Long AD. Structural variants exhibit widespread allelic heterogeneity and shape variation in complex traits. *Nat Commun.* 2019;10(1):4872. <https://doi.org/10.1038/s41467-019-12884-1> PMID: 31653862

32. Chakraborty M, VanKuren NW, Zhao R, Zhang X, Kalsow S, Emerson JJ. Hidden genetic variation shapes the structure of functional elements in *Drosophila*. *Nat Genet*. 2018;50(1):20–5. <https://doi.org/10.1038/s41588-017-0010-y> PMID: 29255259
33. Grenier JK, Arguello JR, Moreira MC, Gottipati S, Mohammed J, Hackett SR, et al. Global diversity lines - a five-continent reference panel of sequenced *Drosophila melanogaster* strains. *G3 (Bethesda)*. 2015;5(4):593–603. <https://doi.org/10.1534/g3.114.015883> PMID: 25673134
34. Nielsen R. Molecular signatures of natural selection. *Annu Rev Genet*. 2005;39:197–218. <https://doi.org/10.1146/annurev.genet.39.073003.112420> PMID: 16285858
35. Haerty W, Ponting CP. Mutations within lncRNAs are effectively selected against in fruitfly but not in human. *Genome Biol*. 2013;14(5):R49. <https://doi.org/10.1186/gb-2013-14-5-r49> PMID: 23710818
36. Hackstein JH, Hochstenbach R, Pearson PL. Towards an understanding of the genetics of human male infertility: lessons from flies. *Trends Genet*. 2000;16(12):565–72. [https://doi.org/10.1016/s0168-9525\(00\)02140-5](https://doi.org/10.1016/s0168-9525(00)02140-5) PMID: 11102707
37. White-Cooper H, Bausek N. Evolution and spermatogenesis. *Philos Trans R Soc Lond B Biol Sci*. 2010;365(1546):1465–80. <https://doi.org/10.1098/rstb.2009.0323> PMID: 20403864
38. Spradling AC, Niu W, Yin Q, Pathak M, Maurya B. Conservation of oocyte development in germline cysts from *Drosophila* to mouse. *Elife*. 2022;11:e83230. <https://doi.org/10.7554/eLife.83230> PMID: 36445738
39. Kopania EEK, Larson EL, Callahan C, Keeble S, Good JM. Molecular Evolution across mouse spermatogenesis. *Mol Biol Evol*. 2022;39(2):msac023. <https://doi.org/10.1093/molbev/msac023> PMID: 35099536
40. Lee M-H, Navarro RE, Han SM. Editorial: germline development: from germline stem cells to gametes. *Front Cell Dev Biol*. 2020;8:650. <https://doi.org/10.3389/fcell.2020.00650> PMID: 32850795
41. Zhang Q, Sun F, Zhang R, Zhao D, Zhu R, Cheng X, et al. The evolution of ovarian somatic cells characterized by transcriptome and chromatin accessibility across rodents, monkeys, and humans. *Life Med*. 2024;3(5):lnae028. <https://doi.org/10.1093/lifemedi/lnae028> PMID: 39872443
42. Obbard DJ, Maclennan J, Kim K-W, Rambaut A, O'Grady PM, Jiggins FM. Estimating divergence dates and substitution rates in the *Drosophila* phylogeny. *Mol Biol Evol*. 2012;29(11):3459–73. <https://doi.org/10.1093/molbev/mss150> PMID: 22683811
43. Rosenberg AB, Roco CM, Muscat RA, Kuchina A, Sample P, Yao Z, et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science*. 2018;360(6385):176–82. <https://doi.org/10.1126/science.aam8999> PMID: 29545511
44. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods*. 2019;16(12):1289–96. <https://doi.org/10.1038/s41592-019-0619-0> PMID: 31740819
45. Zheng Q, Chen X, Qiao C, Wang M, Chen W, Luan X, et al. Somatic CG6015 mediates cyst stem cell maintenance and germline stem cell differentiation via EGFR signaling in *Drosophila* testes. *Cell Death Discov*. 2021;7(1):68. <https://doi.org/10.1038/s41420-021-00452-w> PMID: 33824283
46. Decotto E, Spradling AC. The *Drosophila* ovarian and testis stem cell niches: similar somatic stem cells and signals. *Dev Cell*. 2005;9(4):501–10. <https://doi.org/10.1016/j.devcel.2005.08.012> PMID: 16198292
47. Janzer B, Steinmann-Zwicky M. Cell-autonomous and somatic signals control sex-specific gene expression in XY germ cells of *Drosophila*. *Mech Dev*. 2001;100(1):3–13. [https://doi.org/10.1016/s0925-4773\(00\)00529-3](https://doi.org/10.1016/s0925-4773(00)00529-3) PMID: 11118879
48. Waterhouse RM, Tegenfeldt F, Li J, Zdobnov EM, Kriventseva EV. OrthoDB: a hierarchical catalog of animal, fungal and bacterial orthologs. *Nucleic Acids Res*. 2013;41(Database issue):D358–65. <https://doi.org/10.1093/nar/gks1116> PMID: 23180791
49. Cheng MH, Maines JZ, Wasserman SA. Biphasic subcellular localization of the DAZL-related protein boule in *Drosophila* spermatogenesis. *Dev Biol*. 1998;204(2):567–76. <https://doi.org/10.1006/dbio.1998.9098> PMID: 9882490
50. Eberhart CG, Maines JZ, Wasserman SA. Meiotic cell cycle requirement for a fly homologue of human Deleted in Azoospermia. *Nature*. 1996;381(6585):783–5. <https://doi.org/10.1038/381783a0> PMID: 8657280
51. Ghosh S, Lasko P. Loss-of-function analysis reveals distinct requirements of the translation initiation factors eIF4E, eIF4E-3, eIF4G and eIF4G2 in *Drosophila* spermatogenesis. *PLoS One*. 2015;10(4):e0122519. <https://doi.org/10.1371/journal.pone.0122519> PMID: 25849588
52. Benner L, Muron S, Gomez JG, Oliver B. OVO positively regulates essential maternal pathways by binding near the transcriptional start sites in the *Drosophila* female germline. *Elife*. 2024;13:RP94631. <https://doi.org/10.7554/eLife.94631> PMID: 39291827
53. Lantz V, Chang JS, Horabin JI, Bopp D, Schedl P. The *Drosophila* orb RNA-binding protein is required for the formation of the egg chamber and establishment of polarity. *Genes Dev*. 1994;8(5):598–613. <https://doi.org/10.1101/gad.8.5.598> PMID: 7523244
54. Morabito S, Reese F, Rahimzadeh N, Miyoshi E, Swarup V. hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. *Cell Rep Methods*. 2023;3(6):100498. <https://doi.org/10.1016/j.crmeth.2023.100498> PMID: 37426759
55. Carmena M, Riparbelli MG, Minestrini G, Tavares AM, Adams R, Callaini G, et al. *Drosophila* polo kinase is required for cytokinesis. *J Cell Biol*. 1998;143(3):659–71. <https://doi.org/10.1083/jcb.143.3.659> PMID: 9813088
56. Ables ET, Drummond-Barbosa D. Cyclin E controls *Drosophila* female germline stem cell maintenance independently of its role in proliferation by modulating responsiveness to niche signals. *Development*. 2013;140(3):530–40. <https://doi.org/10.1242/dev.088583> PMID: 23293285
57. Li Y, Maines JZ, Tastan OY, McKearin DM, Buszczak M. Mei-P26 regulates the maintenance of ovarian germline stem cells by promoting BMP signaling. *Development*. 2012;139(9):1547–56. <https://doi.org/10.1242/dev.077412> PMID: 22438571
58. Horvath S. Evaluating whether a module is preserved in another network. 2011. https://doi.org/10.1007/978-1-4419-8819-5_9

59. Santel A, Blümer N, Kämpfer M, Renkawitz-Pohl R. Flagellar mitochondrial association of the male-specific Don Juan protein in *Drosophila* spermatozoa. *J Cell Sci*. 1998;111 (Pt 22):3299–309. <https://doi.org/10.1242/jcs.111.22.3299> PMID: [9788872](#)
60. Laurinyecz B, Vedelek V, Kovács AL, Szilasi K, Lipinski Z, Slezák C, et al. Sperm-Leucylaminopeptidases are required for male fertility as structural components of mitochondrial paracrystalline material in *Drosophila melanogaster* sperm. *PLoS Genet*. 2019;15(2):e1007987. <https://doi.org/10.1371/journal.pgen.1007987> PMID: [30802236](#)
61. Brattig-Correia R, Almeida JM, Wyrwoll MJ, Julca I, Sobral D, Misra CS, et al. The conserved genetic program of male germ cells uncovers ancient regulators of human spermatogenesis. *Elife*. 2024;13:RP95774. <https://doi.org/10.7554/eLife.95774> PMID: [39388236](#)
62. Delgehyr N, Rangone H, Fu J, Mao G, Tom B, Riparbelli MG, et al. Klp10A, a microtubule-depolymerizing kinesin-13, cooperates with CP110 to control *Drosophila* centriole length. *Curr Biol*. 2012;22(6):502–9. <https://doi.org/10.1016/j.cub.2012.01.046> PMID: [22365849](#)
63. Parsch J, Meiklejohn CD, Hauschteck-Jungen E, Hunziker P, Hartl DL. Molecular evolution of the ocnus and janus genes in the *Drosophila melanogaster* species subgroup. *Mol Biol Evol*. 2001;18(5):801–11. <https://doi.org/10.1093/oxfordjournals.molbev.a003862> PMID: [11319264](#)
64. Klattenhoff C, Xi H, Li C, Lee S, Xu J, Khurana JS, et al. The *Drosophila* HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell*. 2009;138(6):1137–49. <https://doi.org/10.1016/j.cell.2009.07.014> PMID: [19732946](#)
65. Parhad SS, Tu S, Weng Z, Theurkauf WE. Adaptive evolution leads to cross-species incompatibility in the piRNA transposon silencing machinery. *Dev Cell*. 2017;43(1):60–70.e5. <https://doi.org/10.1016/j.devcel.2017.08.012> PMID: [28919205](#)
66. Yu B, Lin YA, Parhad SS, Jin Z, Ma J, Theurkauf WE, et al. Structural insights into rhino-deadlock complex for germline piRNA cluster specification. *EMBO Rep*. 2018;19(7):e45418. <https://doi.org/10.15252/embr.201745418> PMID: [29858487](#)
67. McKim KS, Hayashi-Hagihara A. mei-W68 in *Drosophila melanogaster* encodes a Spo11 homolog: evidence that the mechanism for initiating meiotic recombination is conserved. *Genes Dev*. 1998;12(18):2932–42. <https://doi.org/10.1101/gad.12.18.2932> PMID: [9744869](#)
68. Khaitovich P, Hellmann I, Enard W, Nowick K, Leinweber M, Franz H, et al. Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science*. 2005;309(5742):1850–4. <https://doi.org/10.1126/science.1108296> PMID: [16141373](#)
69. Meisel RP. Towards a more nuanced understanding of the relationship between sex-biased gene expression and rates of protein-coding sequence evolution. *Mol Biol Evol*. 2011;28(6):1893–900. <https://doi.org/10.1093/molbev/msr010> PMID: [21239389](#)
70. Lemos B, Meiklejohn CD, Hartl DL. Regulatory evolution across the protein interaction network. *Nat Genet*. 2004;36(10):1059–60. <https://doi.org/10.1038/ng1427> PMID: [15378060](#)
71. Yanai I, Benjamin H, Shmoini M, Chalifa-Caspi V, Shklar M, Ophir R, et al. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics*. 2005;21(5):650–9. <https://doi.org/10.1093/bioinformatics/bti042> PMID: [15388519](#)
72. Leader DP, Krause SA, Pandit A, Davies SA, Dow JAT. FlyAtlas 2: a new version of the *Drosophila melanogaster* expression atlas with RNA-Seq, miRNA-Seq and sex-specific data. *Nucleic Acids Res*. 2018;46(D1):D809–15. <https://doi.org/10.1093/nar/gkx976> PMID: [29069479](#)
73. Larracuente AM, Sackton TB, Greenberg AJ, Wong A, Singh ND, Sturgill D, et al. Evolution of protein-coding genes in *Drosophila*. *Trends Genet*. 2008;24(3):114–23. <https://doi.org/10.1016/j.tig.2007.12.001> PMID: [18249460](#)
74. Soumillon M, Necsulea A, Weier M, Brawand D, Zhang X, Gu H, et al. Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. *Cell Rep*. 2013;3(6):2179–90. <https://doi.org/10.1016/j.celrep.2013.05.031> PMID: [23791531](#)
75. Witt E, Svetec N, Benjamin S, Zhao L. Transcription factors drive opposite relationships between gene age and tissue specificity in male and female *Drosophila* gonads. *Mol Biol Evol*. 2021;38(5):2104–15. <https://doi.org/10.1093/molbev/msab011> PMID: [33481021](#)
76. Kaessmann H. Origins, evolution, and phenotypic impact of new genes. *Genome Res*. 2010;20(10):1313–26. PMID: [20651121](#)
77. Hao L, Ge X, Wan H, Hu S, Lercher MJ, Yu J, et al. Human functional genetic studies are biased against the medically most relevant primate-specific genes. *BMC Evol Biol*. 2010;10:316. <https://doi.org/10.1186/1471-2148-10-316> PMID: [20961448](#)
78. Dong C, Xia S, Zhang L, Arsala D, Fang C, Tan S, et al. Subcellular enrichment patterns of new genes in *Drosophila* evolution. *Mol Biol Evol*. 2025;42(2). <https://doi.org/10.1093/molbev/msaf038> PMID: [39920336](#)
79. Yang Z, Bielawski J. Statistical methods for detecting molecular adaptation. *Trends Ecol Evol*. 2000;15(12):496–503. [https://doi.org/10.1016/s0169-5347\(00\)01994-7](https://doi.org/10.1016/s0169-5347(00)01994-7) PMID: [11114436](#)
80. Booker TR, Jackson BC, Keightley PD. Detecting positive selection in the genome. *BMC Biol*. 2017;15(1):98. <https://doi.org/10.1186/s12915-017-0434-y> PMID: [29084517](#)
81. Galtier N. Adaptive protein evolution in animals and the effective population size hypothesis. *PLoS Genet*. 2016;12(1):e1005774. <https://doi.org/10.1371/journal.pgen.1005774> PMID: [26752180](#)
82. Smith NGC, Eyre-Walker A. Adaptive protein evolution in *Drosophila*. *Nature*. 2002;415(6875):1022–4. <https://doi.org/10.1038/4151022a> PMID: [11875568](#)
83. Charlesworth B, Campos JL, Jackson BC. Faster-X evolution: theory and evidence from *Drosophila*. *Mol Ecol*. 2018;27(19):3753–71. <https://doi.org/10.1111/mec.14534> PMID: [29431881](#)
84. Brawand D, Soumillon M, Necsulea A, Julien P, Csárdi G, Harrigan P, et al. The evolution of gene expression levels in mammalian organs. *Nature*. 2011;478(7369):343–8. <https://doi.org/10.1038/nature10532> PMID: [22012392](#)
85. Llopart A. The rapid evolution of X-linked male-biased gene expression and the large-X effect in *Drosophila yakuba*, *D. santomea*, and their hybrids. *Mol Biol Evol*. 2012;29(12):3873–86. <https://doi.org/10.1093/molbev/mss190> PMID: [22844069](#)

86. Kayserili MA, Gerrard DT, Tomancak P, Kalinka AT. An excess of gene expression divergence on the X chromosome in *Drosophila* embryos: implications for the faster-X hypothesis. *PLoS Genet*. 2012;8(12):e1003200. <https://doi.org/10.1371/journal.pgen.1003200> PMID: 23300473
87. Vicoso B, Charlesworth B. Effective population size and the faster-X effect: an extended model. *Evolution*. 2009;63(9):2413–26. <https://doi.org/10.1111/j.1558-5646.2009.00719.x> PMID: 19473388
88. Warnefors M, Kaessmann H. Evolution of the correlation between expression divergence and protein divergence in mammals. *Genome Biol Evol*. 2013;5(7):1324–35. <https://doi.org/10.1093/gbe/evt093> PMID: 23781097
89. Hense W, Baines JF, Parsch J. X chromosome inactivation during *Drosophila* spermatogenesis. *PLoS Biol*. 2007;5(10):e273. <https://doi.org/10.1371/journal.pbio.0050273> PMID: 17927450
90. Charlesworth B, Coyne JA, Barton NH. The relative rates of evolution of sex chromosomes and autosomes. *Am Nat*. 1987;130(1):113–46.
91. Wolfenbarger LL, Wilkinson GS. Sex-linked expression of a sexually selected trait in the stalk-eyed fly, *Cyrtodopsis dalmanni*. *Evolution*. 2001;55(1):103–10. <https://doi.org/10.1111/j.0014-3820.2001.tb01276.x> PMID: 11263731
92. Whittle CA, Extavour CG. Selection shapes turnover and magnitude of sex-biased expression in *Drosophila* gonads. *BMC Evol Biol*. 2019;19(1):60. <https://doi.org/10.1186/s12862-019-1377-4> PMID: 30786879
93. Meisel RP, Malone JH, Clark AG. Disentangling the relationship between sex-biased gene expression and X-linkage. *Genome Res*. 2012;22(7):1255–65. <https://doi.org/10.1101/gr.132100.111> PMID: 22499666
94. Karpova N, Bobinac Y, Fouix S, Huitorel P, Debec A. Jupiter, a new *Drosophila* protein associated with microtubules. *Cell Motil Cytoskeleton*. 2006;63(5):301–12. <https://doi.org/10.1002/cm.20124> PMID: 16518797
95. Baker CC, Fuller MT. Translational control of meiotic cell cycle progression and spermatid differentiation in male germ cells by a novel eIF4G homolog. *Development*. 2007;134(15):2863–9. <https://doi.org/10.1242/dev.003764> PMID: 17611220
96. Danev N, Li G, Duan JE, Van de Walle GR. Comparative transcriptomic analysis of bovine mesenchymal stromal cells reveals tissue-source and species-specific differences. *iScience*. 2024;27(2):108886. <https://doi.org/10.1016/j.isci.2024.108886> PMID: 38318381
97. Lee D, Shahandeh MP, Abuin L, Benton R. Comparative single-cell transcriptomic atlases of drosophilid brains suggest glial evolution during ecological adaptation. *PLoS Biol*. 2025;23(4):e3003120. <https://doi.org/10.1371/journal.pbio.3003120> PMID: 40299832
98. Wang B, Starr AL, Fraser HB. Cell-type-specific cis-regulatory divergence in gene expression and chromatin accessibility revealed by human-chimpanzee hybrid cells. *Elife*. 2024;12:RP89594. <https://doi.org/10.7554/eLife.89594> PMID: 38358392
99. Shami AN, Zheng X, Munyoki SK, Ma Q, Manske GL, Green CD, et al. Single-cell RNA sequencing of human, macaque, and mouse testes uncovers conserved and divergent features of mammalian spermatogenesis. *Dev Cell*. 2020;54(4):529–547.e12. <https://doi.org/10.1016/j.dev-cel.2020.05.010> PMID: 32504559
100. Andrews J, Garcia-Estefania D, Delon I, Lü J, Mével-Ninio M, Spierer A, et al. OVO transcription factors function antagonistically in the *Drosophila* female germline. *Development*. 2000;127(4):881–92. <https://doi.org/10.1242/dev.127.4.881> PMID: 10648246
101. Huang Y, Shang R, Lu G-A, Zeng W, Huang C, Zou C, et al. Spatiotemporal regulation of a single adaptively evolving trans-regulatory element contributes to spermatogenetic expression divergence in *Drosophila*. *Mol Biol Evol*. 2022;39(7):msac127. <https://doi.org/10.1093/molbev/msac127> PMID: 35687719
102. Good JM, Nachman MW. Rates of protein evolution are positively correlated with developmental timing of expression during mouse spermatogenesis. *Mol Biol Evol*. 2005;22(4):1044–52. <https://doi.org/10.1093/molbev/msi087> PMID: 15647515
103. Ramm SA, Schärer L, Ehmcke J, Wistuba J. Sperm competition and the evolution of spermatogenesis. *Mol Hum Reprod*. 2014;20(12):1169–79. <https://doi.org/10.1093/molehr/gau070> PMID: 25323971
104. Vedelek V, Jankovics F, Zádori J, Sinka R. Mitochondrial differentiation during spermatogenesis: lessons from *Drosophila melanogaster*. *Int J Mol Sci*. 2024;25(7):3980. <https://doi.org/10.3390/ijms25073980> PMID: 38612789
105. Riparbelli MG, Persico V, Callaini G. The microtubule cytoskeleton during the early *Drosophila* spermiogenesis. *Cells*. 2020;9(12):2684. <https://doi.org/10.3390/cells9122684>
106. Civetta A, Ranz JM. Genetic factors influencing sperm competition. *Front Genet*. 2019;10(820):820. <https://doi.org/10.3389/fgene.2019.00820> PMID: 31572439
107. Joly D, Korol A, Nevo E. Sperm size evolution in *Drosophila*: inter- and intraspecific analysis. *Genetica*. 2004;120(1–3):233–44. <https://doi.org/10.1023/b:gene.0000017644.63389.57> PMID: 15088661
108. Spradling A. Developmental genetics of oogenesis. In: Bate A, Martinez-Arias A, editors. *The development of Drosophila melanogaster*. New York: CSHL; 1993. p. 1–70.
109. Carpenter AT. Electron microscopy of meiosis in *Drosophila melanogaster* females. I. Structure, arrangement, and temporal change of the synaptonemal complex in wild-type. *Chromosoma*. 1975;51(2):157–82. <https://doi.org/10.1007/BF00319833> PMID: 806439
110. Zhao Z-H, Meng T-G, Zhang H-Y, Hou Y, Schatten H, Wang Z-B, et al. Single-cell RNA sequencing reveals species-specific time spans of cell cycle transitions in early oogenesis. *Hum Mol Genet*. 2021;30(7):525–35. <https://doi.org/10.1093/hmg/ddab048> PMID: 33575778
111. Trost N, Fallahshahroudi A, Sarropoulos I, Schneider C, Schmidt J, Mbengue N. The evolution of gene regulatory programs controlling gonadal development in primates. *bioRxiv*. 2025. <https://doi.org/10.1101/2025.06.17.659946>

112. Vallés AM, Rubin T, Macaisne N, Dal Toe L, Molla-Herman A, Antoniewski C, et al. Transcriptomic analysis of meiotic genes during the mitosis-to-meiosis transition in *Drosophila* females. *Genetics*. 2024;228(2):iyae130. <https://doi.org/10.1093/genetics/iyae130> PMID: 39225982
113. Parker GA. Selection on non-random fusion of gametes during the evolution of anisogamy. *J Theor Biol*. 1978;73(1):1–28. [https://doi.org/10.1016/0022-5193\(78\)90177-7](https://doi.org/10.1016/0022-5193(78)90177-7) PMID: 567721
114. Pizzari T, Parker GA. Sperm competition and sperm phenotype. In: *Sperm Biology*. Elsevier; 2009. p. 207–45. <https://doi.org/10.1016/b978-0-12-372568-4.00006-9>
115. King EG, Merkes CM, McNeil CL, Hofer SR, Sen S, Broman KW, et al. Genetic dissection of a model complex trait using the *Drosophila* synthetic population resource. *Genome Res*. 2012;22(8):1558–66. <https://doi.org/10.1101/gr.134031.111> PMID: 22496517
116. Brizuela BJ, Elfring L, Ballard J, Tamkun JW, Kennison JA. Genetic analysis of the brahma gene of *Drosophila melanogaster* and polytene chromosome subdivisions 72AB. *Genetics*. 1994;137(3):803–13. <https://doi.org/10.1093/genetics/137.3.803> PMID: 7916308
117. Drosophila 12 Genomes Consortium, Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, et al. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature*. 2007;450(7167):203–18. <https://doi.org/10.1038/nature06341> PMID: 17994087
118. Markow TA, O'Grady PM. *Drosophila*: a guide to species identification and use. Netherlands: Academic Press; 2006. p. 189–99.
119. Andrews S. FastQC: a quality control tool for high throughput sequence data; 2010.
120. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, et al. The genome sequence of *Drosophila melanogaster*. *Science*. 2000;287(5461):2185–95. <https://doi.org/10.1126/science.287.5461.2185> PMID: 10731132
121. Hu TT, Eisen MB, Thornton KR, Andolfatto P. A second-generation assembly of the *Drosophila simulans* genome provides new insights into patterns of lineage-specific divergence. *Genome Res*. 2013;23(1):89–98. <https://doi.org/10.1101/gr.141689.112> PMID: 22936249
122. Shumate A, Salzberg SL. Liftoff: accurate mapping of gene annotations. *Bioinformatics*. 2021;37(12):1639–43. <https://doi.org/10.1093/bioinformatics/btaa1016> PMID: 33320174
123. Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. *Nat Commun*. 2017;8:14049. <https://doi.org/10.1038/ncomms14049> PMID: 28091601
124. Tran V, Papalexi E, Schroeder S, Kim G, Sapre A, Pangallo J, et al. High sensitivity single cell RNA sequencing with split pool barcoding. *bioRxiv*. 2022. <https://doi.org/10.1101/2022.08.27.505512>
125. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol*. 2024;42(2):293–304. <https://doi.org/10.1038/s41587-023-01767-y> PMID: 37231261
126. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. *Cell Syst*. 2019;8(4):329–337.e4. <https://doi.org/10.1016/j.cels.2019.03.003> PMID: 30954475
127. Liu J, Gao C, Sodico J, Kozareva V, Macosko EZ, Welch JD. Jointly defining cell types from multiple single-cell datasets using LIGER. *Nat Protoc*. 2020;15(11):3632–62. <https://doi.org/10.1038/s41596-020-0391-8> PMID: 33046898
128. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018;36(5):411–20. <https://doi.org/10.1038/nbt.4096> PMID: 29608179
129. Lopez R, Regier J, Cole MB, Jordan MI, Yosef N. Deep generative modeling for single-cell transcriptomics. *Nat Methods*. 2018;15(12):1053–8. <https://doi.org/10.1038/s41592-018-0229-2> PMID: 30504886
130. Song Y, Miao Z, Brazma A, Papatheodorou I. Benchmarking strategies for cross-species integration of single-cell RNA sequencing data. *Nat Commun*. 2023;14(1):6495. <https://doi.org/10.1038/s41467-023-41855-w> PMID: 37838716
131. Zappia L, Oshlack A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. *Gigascience*. 2018;7(7):giy083. <https://doi.org/10.1093/gigascience/giy083> PMID: 30010766
132. Blanco-Carmona E. Generating publication ready visualizations for single cell transcriptomics using SCpubr. *bioRxiv*. 2022:2022.02.28.482303. <https://doi.org/10.1101/2022.02.28.482303>
133. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol*. 2014;32(4):381–6. <https://doi.org/10.1038/nbt.2859> PMID: 24658644
134. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods*. 2017;14(10):979–82. <https://doi.org/10.1038/nmeth.4402> PMID: 28825705
135. Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature*. 2019;566(7745):496–502. <https://doi.org/10.1038/s41586-019-0969-x> PMID: 30787437
136. Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol*. 2015;16:278. <https://doi.org/10.1186/s13059-015-0844-5> PMID: 26653891
137. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z. ClusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)*. 2021;2(3):100141. <https://doi.org/10.1016/j.xinn.2021.100141> PMID: 34557778
138. Carlson M. Org.Dm.eg.db: Genome wide annotation for Fly; 2023.

139. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Stat Methodol.* 1995;57(1):289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
140. Fingerhut JM, Yamashita YM. Analysis of gene expression patterns and RNA localization by fluorescence in situ hybridization in whole mount *Drosophila* testes. *Methods Mol Biol.* 2023;2666:15–28. https://doi.org/10.1007/978-1-0716-3191-1_2 PMID: 37166654
141. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997;25(17):3389–402. <https://doi.org/10.1093/nar/25.17.3389> PMID: 9254694
142. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods.* 2012;9(7):676–82. <https://doi.org/10.1038/nmeth.2019> PMID: 22743772
143. Gramates LS, Agapite J, Attrill H, Calvi BR, Crosby MA, Dos Santos G, et al. FlyBase: a guided tour of highlighted features. *Genetics.* 2022;220(4):iyac035. <https://doi.org/10.1093/genetics/iyac035> PMID: 35266522
144. Murga-Moreno J, Coronado-Zamora M, Hervas S, Casillas S, Barbadilla A. iMKT: the integrative McDonald and Kreitman test. *Nucleic Acids Res.* 2019;47(W1):W283–8. <https://doi.org/10.1093/nar/gkz372> PMID: 31081014
145. Mackay TF, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu D. The *Drosophila melanogaster* genetic reference panel. *Nature.* 2012;482(7384):173–8. <https://doi.org/10.1038/nature10811> PMID: 22318601
146. R Development Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2016.
147. Hariyani I, Ranz JM. Ranz-Lab/Drosophila-snrRNAseq: v1.0.1. 1.0.1 ed. Geneva, Switzerland: Zenodo; 2025.
148. Gel B, Serra E. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics.* 2017;33(19):3088–90. <https://doi.org/10.1093/bioinformatics/btx346> PMID: 28575171