RESEARCH ARTICLE SUMMARY

GENETICS

# Fly Cell Atlas: A single-nucleus transcriptomic atlas of the adult fruit fly

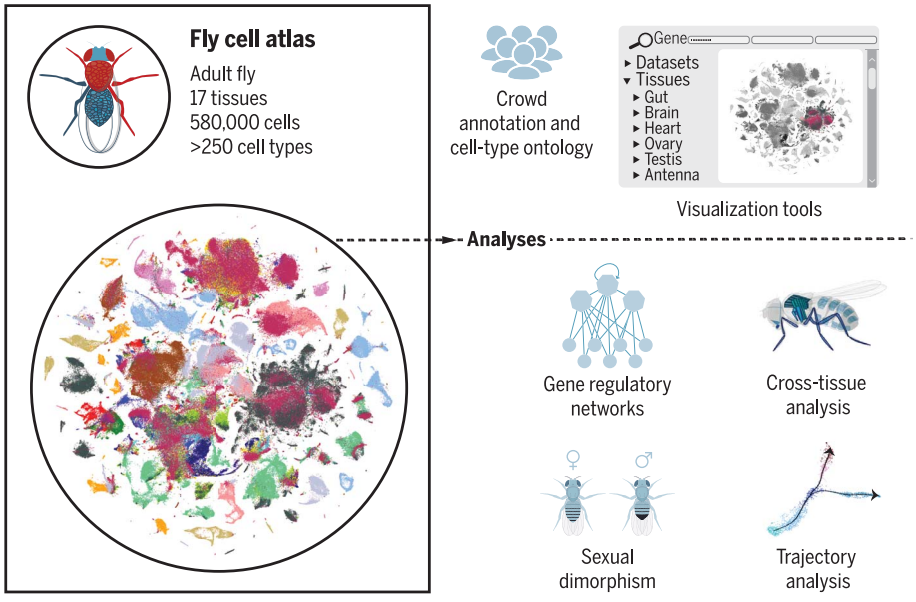
**Hongjie Li†, Jasper Janssens†, Maxime De Waegeneer, Sai Saroja Kolluru, Kristofer Davie, Vincent Gardeux, Wouter Saelens, Fabrice P. A. David, Maria Brbić, Katina Spanier, Jure Leskovec, Colleen N. McLaughlin, Qijing Xie, Robert C. Jones, Katja Brueckner, Jiwon Shim, Sudhir Gopal Tattikota, Frank Schnorrer, Katja Rust, Todd G. Nystul, Zita Carvalho-Santos, Carlos Ribeiro, Soumitra Pal, Sharvani Mahadevaraju, Teresa M. Przytycka, Aaron M. Allen, Stephen F. Goodwin, Cameron W. Berry, Margaret T. Fuller, Helen White-Cooper, Erika L. Matunis, Stephen DiNardo, Anthony Galenza, Lucy Erin O'Brien, Julian A. T. Dow, FCA Consortium, Heinrich Jasper, Brian Oliver, Norbert Perrimon*, Bart Deplancke*, Stephen R. Quake*, Liqun Luo*, Stein Aerts***

**INTRODUCTION:** *Drosophila melanogaster* has had a fruitful history in biological research because it has contributed to many key discoveries in genetics, development, and neurobiology. The fruit fly genome contains ~14,000 protein-coding genes, ~63% of which have human orthologs. Single-cell RNA-sequencing has recently been applied to multiple *Drosophila* tissues and developmental stages. However, these data have been generated by different laboratories on different genetic backgrounds with different dissociation protocols and sequencing platforms, which has hindered the systematic comparison of gene expression across cells and tissues.

**RATIONALE:** We aimed to establish a cell atlas for the entire adult *Drosophila* with the same genetic background, dissociation protocol, and sequencing platform to (i) obtain a comprehensive categorization of cell types, (ii) integrate single-cell transcriptome data with existing knowledge about gene expression and cell types, (iii) systematically compare gene expression across the entire organism and between males and females, and (iv) identify cell type–specific markers across the entire organism. We chose single-nucleus RNA-sequencing (snRNA-seq) to circumvent the difficulties of dissociating cells that are embedded in the cuticle (e.g., sensory neurons) or that are multinucleated (e.g., muscle cells). We took two complementary strategies: sequencing nuclei from dissected tissues to know the identity of the tissue source and sequencing nuclei from the entire head and body to ensure that all cells are sampled. Experts from 40 laboratories participated in crowd annotation to assign transcriptomic cell types with the best knowledge available.

**RESULTS:** We sequenced 570,000 cells using droplet-based 10x Genomics from 15 dissected tissues as well as whole heads and bodies, separately in females and males. We also sequenced 10,000 cells from dissected tissues using the plate-based Smart-seq2 platform, providing deeper coverage per cell. We developed reproducible analysis pipelines using NextFlow and implemented a distributed cell-type annotation system with controlled vocabularies in SCope. Crowd-based annotations of transcriptomes from dissected tissues identified 17 main cell categories and 251 detailed cell types linked to FlyBase ontologies. Many of these cell types are characterized for the first time, either because they emerged only after increasing cell coverage or because they reside in tissues that had not been previously subjected to scRNA-seq. The excellent correspondence of transcriptomic clusters from whole body and dissected tissues allowed us to transfer annotations and identify a few cuticular cell types not detected in individual tissues. Cross-tissue analysis revealed location-specific subdivisions of muscle cells and heterogeneity within blood cells. We then determined cell type–specific marker genes and transcription factors with different specificity levels, enabling the construction of gene regulatory networks. Finally, we explored sexual dimorphism, finding a link between sex-biased expression and the presence of *doublesex*, and investigated tissue dynamics through trajectory analyses.

**CONCLUSION:** Our Fly Cell Atlas (FCA) constitutes a valuable resource for the *Drosophila* community as a reference for studies of gene function at single-cell resolution. All the FCA data are freely available for further analysis through multiple portals and can be downloaded for custom analyses using other single-cell tools. The ability to annotate cell types by sequencing the entire head and body will facilitate the use of *Drosophila* in the study of biological processes and in modeling human diseases at a whole-organism level with cell-type resolution. All data with annotations can be accessed from www.flycellatlas.org, which provides links to SCope, ASAP, and cellxgene portals. ■




**Tabula *Drosophilae*.** In this single-cell atlas of the adult fruit fly, 580,000 cells were sequenced and >250 cell types were annotated. They are from 15 individually dissected sexed tissues as well as the entire head and body. All data are freely available for visualization and download, with featured analyses shown at the bottom right.

The list of author affiliations is available in the full article online.
***Corresponding author. Email: perrimon@genetics.med.harvard.edu (N.P.); bart.deplancke@epfl.ch (B.D.); steve@quake-lab.org (S.R.Q.); lluo@stanford.edu (L.L.); stein.aerts@kuleuven.be (S.A.)**
†These authors contributed equally to this work.
**Cite this article as H. Li *et al.*, *Science* 375, eabk2432 (2022). DOI: 10.1126/science.abk2432**

**READ THE FULL ARTICLE AT**
https://doi.org/10.1126/science.abk2432

RESEARCH ARTICLE

GENETICS

# Fly Cell Atlas: A single-nucleus transcriptomic atlas of the adult fruit fly

Hongjie Li[1,2,3]†, Jasper Janssens[4,5]†, Maxime De Waegeneer[4,5], Sai Saroja Kolluru[6,7,8], Kristofer Davie[4], Vincent Gardeux[9,10], Wouter Saelens[9,10], Fabrice P. A. David[9,10,11], Maria Brbić[12,8], Katina Spanier[4,5], Jure Leskovec[12,8], Colleen N. McLaughlin[1], Qijing Xie[1], Robert C. Jones[6,7,8], Katja Brueckner[13]‡, Jiwon Shim[14], Sudhir Gopal Tattikota[15,16], Frank Schnorrer[17], Katja Rust[18,19], Todd G. Nystul[19], Zita Carvalho-Santos[20], Carlos Ribeiro[20], Soumitra Pal[21], Sharvani Mahadevaraju[22], Teresa M. Przytycka[21], Aaron M. Allen[23], Stephen F. Goodwin[23], Cameron W. Berry[24], Margaret T. Fuller[24], Helen White-Cooper[25], Erika L. Matunis[26], Stephen DiNardo[27,28], Anthony Galenza[29], Lucy Erin O'Brien[29], Julian A. T. Dow[30], FCA Consortium§, Heinrich Jasper[31], Brian Oliver[22], Norbert Perrimon[15,16]*, Bart Deplancke[9,10]*, Stephen R. Quake[6,7,8]*, Liqun Luo[1]*, Stein Aerts[4,5]*

For more than 100 years, the fruit fly *Drosophila melanogaster* has been one of the most studied model organisms. Here, we present a single-cell atlas of the adult fly, Tabula *Drosophilae*, that includes 580,000 nuclei from 15 individually dissected sexed tissues as well as the entire head and body, annotated to >250 distinct cell types. We provide an in-depth analysis of cell type–related gene signatures and transcription factor markers, as well as sexual dimorphism, across the whole animal. Analysis of common cell types between tissues, such as blood and muscle cells, reveals rare cell types and tissue-specific subtypes. This atlas provides a valuable resource for the *Drosophila* community and serves as a reference to study genetic perturbations and disease models at single-cell resolution.

*Drosophila melanogaster* has had a fruitful history in biological research, dating back to the experiments of Thomas Hunt Morgan more than a century ago (*1*), and has been at the basis of many key biological discoveries. The highly collaborative nature of the *Drosophila* community contributed to many of these successes and led to the development of essential research resources, including a high-quality genome (*2*), a large collection of genetic and molecular tools, and important databases such as Flybase (*3*), FlyMine (*4*), FlyLight (*5*), VirtualFlyBrain (*6*), and ModERN (*7*). The fly genome contains about 17,000 genes, including 13,968 protein-coding genes of which ~63% have human orthologs. Studies such as ModENCODE (*8*) and FlyAtlas (*9*) explored expression patterns in different tissues but lacked cell-type resolution. Recent advances in single-cell technologies have enabled the transcriptomic profiling of thousands of cells at once, facilitating the creation of tissue-wide atlases. Several studies have already applied single-cell RNA sequencing (scRNA-seq) to multiple *Drosophila* tissues and developmental stages (*10*). However, these data were generated by different laboratories on different genetic backgrounds with different dissociation protocols and sequencing platforms, which has hindered the systematic comparison of gene expression across cells and tissues.

Here, we present a single-cell transcriptomic atlas of the entire adult *Drosophila*, with male and female samples separately analyzed, using a uniform genotype and a unified single-nucleus RNA-seq (snRNA-seq) platform (*11*) with two sequencing strategies: droplet-based 10x Genomics (*12*) and plate-based Smart-seq2 (*13*). The resulting Tabula *Drosophilae*, the first dataset within the Fly Cell Atlas (FCA) Consortium, contains more than 580,000 cells, resulting in >250 distinct cell types annotated by >100 experts from 40 laboratories. This atlas reports cellular signatures for each tissue, providing the entire *Drosophila* community a reference for studies that probe the effects of genetic perturbations and disease models at single-cell resolution. All data and annotations can be accessed through multiple visualization and analysis portals from https://flycellatlas.org (figs. S1 to S3).

### Sampling single cells across the entire adult fly

We used a unified snRNA-seq platform for all samples because it is difficult to isolate intact cells from many adult *Drosophila* tissues, especially cuticular ones (e.g., antenna, wing) and adipocyte-enriched ones (e.g., fat body). In addition, snRNA-seq can be applied to large multinucleated cells (e.g., muscle) and facilitates (frozen) tissue collection from different laboratories. Finally, 70 to 90% of transcriptomic information is preserved from snRNA-seq compared with scRNA-seq of the same fly cell types (*11*).

To achieve a comprehensive sampling, we used two complementary strategies. First, we dissected 12 individual tissues from both males and females as well as three sex-specific tissues (Fig. 1A). For tissues that are localized across the body (fat body, oenocytes, and trachea) and cannot be directly dissected, we used specific GAL4 lines driving nuclear–green fluorescent protein (GFP) to label and collect nuclei using fluorescence-activated cell sorting (FACS). In addition, two rare cell types were sequenced only with Smart-seq2: insulin-producing cells and corpora cardiaca cells. Second, we sorted and profiled nuclei from the entire head and body, aiming to detect cell types not covered by the selected tissues. In total, we obtained 580,000 high-quality nuclei: 570,000 from 10x Genomics and 10,000 from Smart-seq2 (Fig. 1A).



[1]Howard Hughes Medical Institute, Department of Biology, Stanford University, Stanford, CA 94305, USA. [2]Huffington Center on Aging, Baylor College of Medicine, Houston, TX 77030, USA. [3]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA. [4]VIB-KU Leuven Center for Brain and Disease Research, KU Leuven, 3000 Leuven, Belgium. [5]Laboratory of Computational Biology, Department of Human Genetics, KU Leuven, Leuven 3000, Belgium. [6]Department of Bioengineering, Stanford University, Stanford, CA, USA. [7]Department of Applied Physics, Stanford University, Stanford, CA, USA. [8]Chan Zuckerberg Biohub, San Francisco CA, USA. [9]Laboratory of Systems Biology and Genetics, Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland. [10]Swiss Institute of Bioinformatics, CH-1015 Lausanne, Switzerland. [11]Bioinformatics Competence Center, EPFL, Switzerland. [12]Department of Computer Science, Stanford University, Stanford, CA 94305, USA. [13]Department of Cell and Tissue Biology, University of California, San Francisco, CA 94143, USA. [14]Department of Life Science, College of Natural Science, Hanyang University, 04763 Seoul, Republic of Korea. [15]Department of Genetics, Blavatnik Institute, Harvard Medical School, Boston, MA 02115, USA. [16]Howard Hughes Medical Institute, Harvard Medical School, Boston, MA, USA. [17]Aix-Marseille University, CNRS, IBDM (UMR 7288), Turing Centre for Living Systems, 13009 Marseille, France. [18]Institute of Physiology and Pathophysiology, Department of Molecular Cell Physiology, Philipps-University, Marburg, Germany. [19]Department of Anatomy, University of California, San Francisco, CA 94143, USA. [20]Behavior and Metabolism Laboratory, Champalimaud Research, Champalimaud Centre for the Unknown, Lisbon, Portugal. [21]National Center of Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA. [22]Laboratory of Cellular and Developmental Biology, National Institute of Diabetes and Kidney and Digestive Diseases, National Institutes of Health, Bethesda, MD 20892, USA. [23]Centre for Neural Circuits and Behaviour, University of Oxford, Oxford OX1 3SR, UK. [24]Department of Developmental Biology and Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA. [25]Molecular Biosciences Division, Cardiff University, Cardiff CF10 3AX, UK. [26]Department of Cell Biology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. [27]Perelman School of Medicine, The University of Pennsylvania, Philadelphia, PA 19104, USA. [28]The Penn Institute for Regenerative Medicine, Philadelphia, PA 19104, USA. [29]Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, CA 94305, USA. [30]Institute of Molecular, Cell and Systems Biology, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK. [31]Immunology Discovery, Genentech, Inc., South San Francisco, CA 94080, USA.

***Corresponding author. Email: perrimon@genetics.med.harvard.edu (N.P.); bart.deplancke@epfl.ch (B.D.); steve@quake-lab.org (S.R.Q.); lluo@stanford.edu (L.L.); stein.aerts@kuleuven.be (S.A.)**

†These authors contributed equally to this work. ‡Deceased. §FCA Consortium authors and affiliations are listed at the end of this paper.

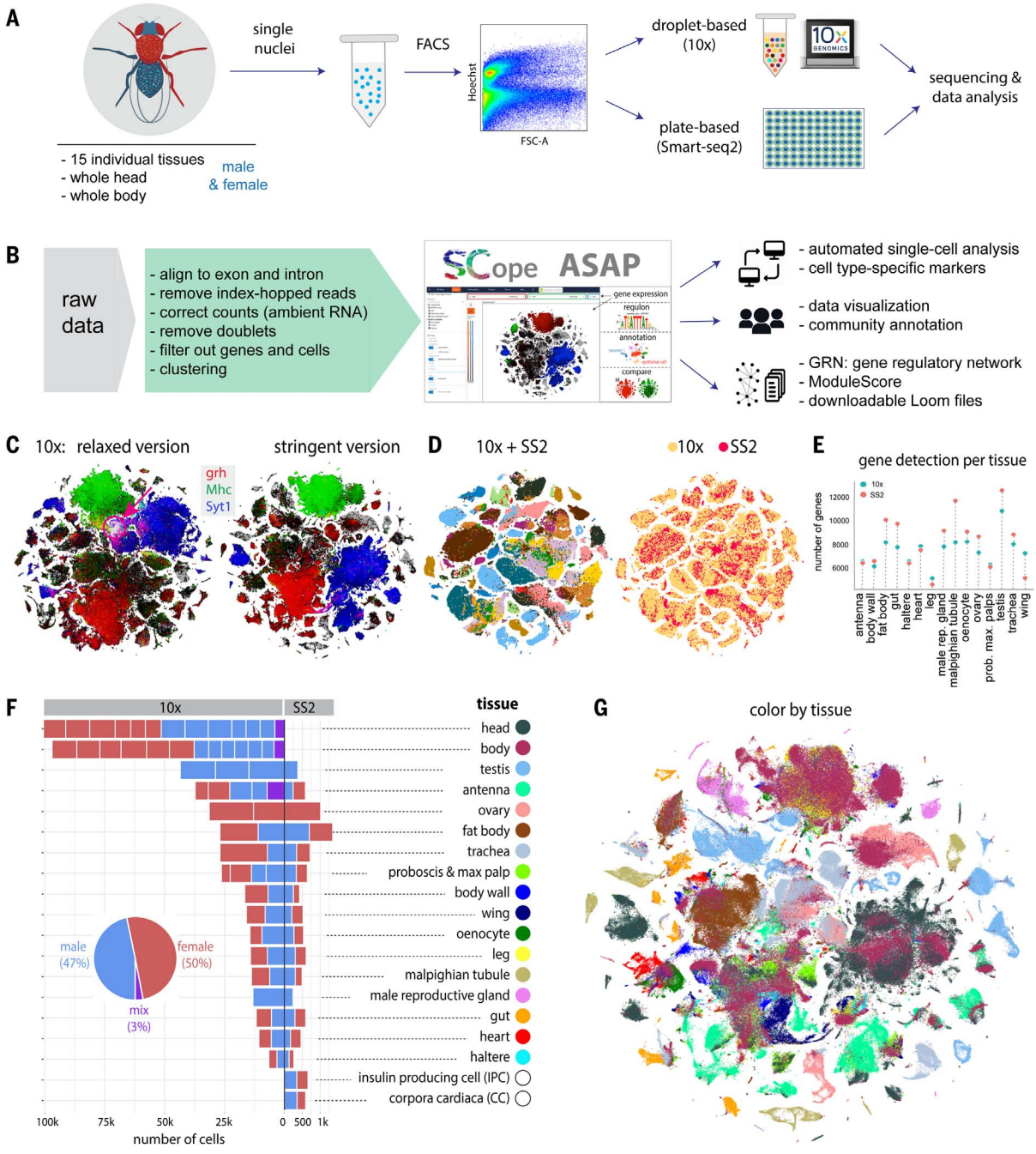



**Fig. 1. Overview of the FCA.** (**A**) Experimental platform of snRNA-seq using 10x Genomics and Smart-seq2. (**B**) Data analysis pipeline and data visualization using SCope (*17*) and ASAP (*18*). (**C**) Two versions of 10x datasets: Relaxed and Stringent. *t*-distributed stochastic neighbor embedding (tSNE) colors are based on gene expression: *grh* (epithelia, red), *Mhc* (muscle, green), and *Syt1* (neuron, blue). The red arrow denotes an artifactual cluster with coexpression of all three markers in the Relaxed dataset. (**D**) tSNE visualization of cells from the Stringent 10x dataset and Smart-seq2 (SS2) cells. 10x cells are from individual tissues. Integrated data are colored by tissue (left) and platform (right). (**E**) Tissue-level comparison of the number of detected genes between 10x and Smart-seq2 platforms. (**F**) Number of cells for each tissue by 10x and Smart-seq2. Male and female cells are indicated. Mixed cells are from pilot experiments where flies were not sexed. Different batches are separated by vertical white lines. (**G**) All 10x cells from the Stringent dataset clustered together; cells are colored by tissue type. Tissue names and colors are indexed as in (F).

To analyze the 10x Genomics data in a reproducible manner, we used the automated VSN pipeline (*14*) (methods and table S1), which takes the raw sequencing data as input and performs preprocessing (e.g., normalization, doublet removal, batch-effect correction) to produce LoomX-formatted files with expression data, embeddings, and clusterings (Fig. 1B and fig. S4). A presumed artifactual cluster showed expression of nearly all genes, so we added an additional preprocessing step that models and subtracts ambient RNA signals (*15*) to remove this cluster, resulting in a Stringent dataset of 510,000 cells (see methods and Fig. 1C). However, because adjusting the gene expression values per cell can introduce other biases (e.g., overcorrection, removal of nondoublet cells), we also retained the original

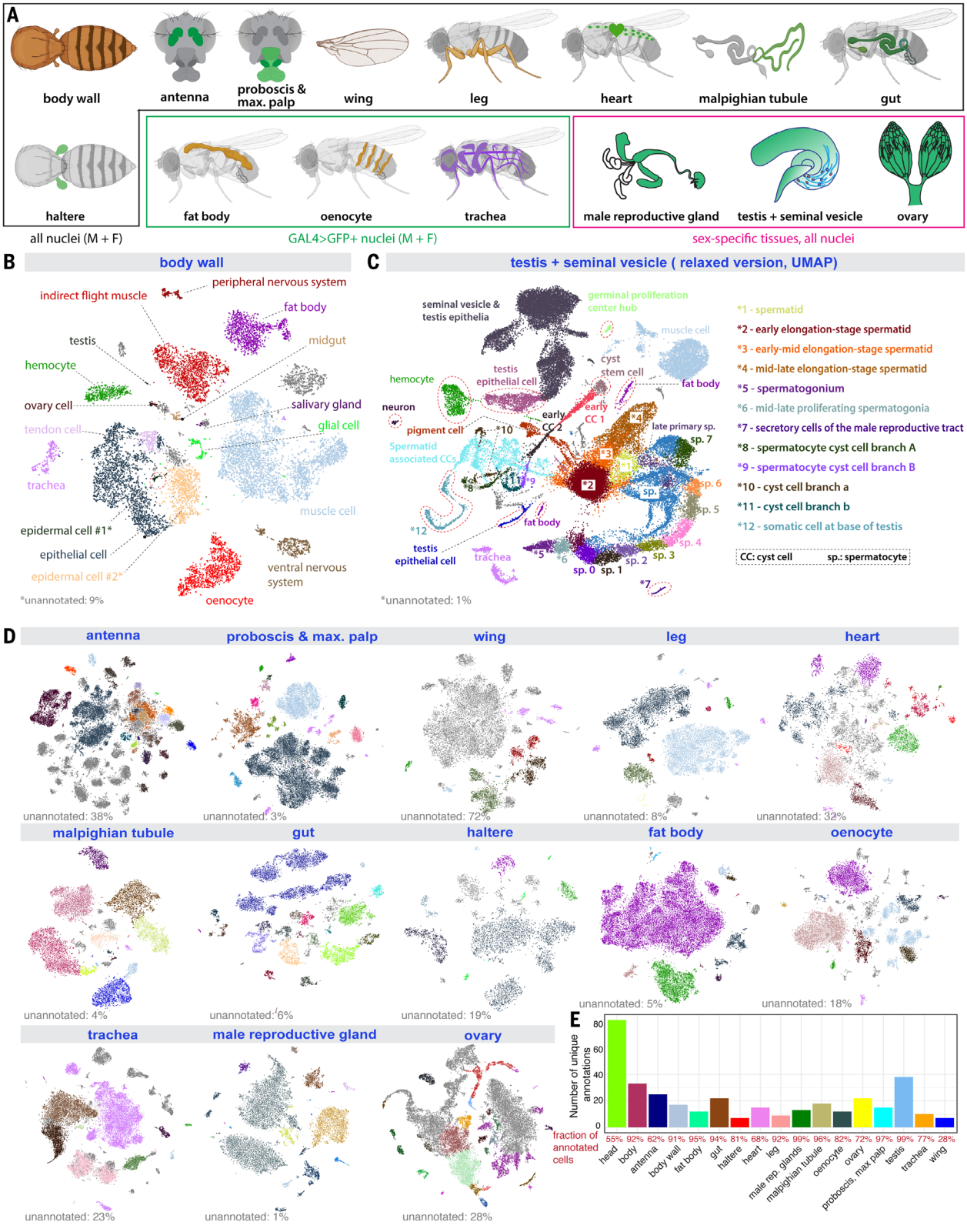



**Fig. 2. Cell-type annotation for dissected tissues.** (**A**) Illustration of 15 individual tissues: Twelve are sequenced separately from males and females and three are sex-specific. Fat body, oenocyte, and tracheal nuclei were labeled using a tissue-specific GAL4 driving UAS–nuclear-GFP. (**B**) tSNE plot with annotations for the body wall from the Stringent 10x dataset. *1, epidermal cells of the abdominal posterior compartment; *2, epidermal cells specialized in antimicrobial response. (**C**) Uniform manifold approximation and projection (UMAP) plot with annotations for the testis from the Relaxed 10x dataset. (**D**) tSNE plots of the other 13 tissues from the Stringent 10x dataset. Detailed annotations are in figs. S6 to S18. (**E**) Number of unique annotations for each tissue. Fractions of annotated cells over all analyzed cells from the Relaxed dataset are indicated in red.

Relaxed dataset of 570,000 cells. In the analyses below, unless mentioned otherwise (e.g., Fig. 2C), the Stringent dataset was used.

Cells from 10x Genomics and Smart-seq2 were well integrated after batch correction using Harmony (*16*) (Fig. 1D). Smart-seq2 yielded a higher number of detected genes for most tissues (Fig. 1E) because cells were sequenced to a higher depth. We analyzed each tissue separately, combining the male and female runs, which yielded between 6500 (haltere) and 100,000 (head) cells and a median of 16,500 cells per tissue for 10x and between 263 (male reproductive gland) and 1349 (fat body) cells and a median of 534 cells per tissue for Smart-seq2 (Fig. 1F). We obtained similar numbers of male and female cells for non–sex-specific tissues with, on average, 1895 unique molecular identifiers (UMIs) and 828 genes per cell (fig. S5). Next, all cells were combined in a meta-analysis, showing tissue-specific clusters like the germline cells of the testis and ovary and shared clusters of common cell types (Fig. 1G and figs. S24 and S25).

### Crowd-based cell-type annotation by tissue experts

Experts from 40 laboratories collaborated on cell-type annotation for 15 individual tissues, including 12 tissues for both sexes (antenna, body wall, fat body, haltere, heart, gut, leg, Malpighian tubule, oenocyte, proboscis with maxillary palp, trachea, and wing) and three sex-specific tissues (male reproductive gland, testis, and ovary) (Fig. 2A). We developed a consensus-voting strategy within the SCope web application (https://flycellatlas.org/scope) (*17*), where curators annotated clusters at multiple resolutions (ranging from 0.8 to 8; fig. S6A), with additional analysis performed in ASAP (https://flycellatlas.org/asap) (*18*). To ensure that cell-type annotations are consistent with previous literature and databases and to allow posteriori computational analyses at different anatomical resolutions, we used Flybase anatomy ontology terms (*19*).

Because some cell types are annotated at low resolutions and others at high resolutions, we collapsed all annotations across resolutions and retained the annotation with the highest number of up votes. All initial annotations were performed on the Relaxed dataset and were then exported to the Stringent dataset, where field experts verified the accuracy of the annotation transfer (Fig. 2, A to E, and figs. S6 to S18). Overall, we annotated 251 cell types in the Stringent dataset (262 cell types if combining Relaxed and Stringent datasets; table S2), with a median of 15 cell types per tissue.

Our dataset provides a single-cell transcriptomic profile for several adult tissues not profiled previously, including the haltere, heart, leg, Malpighian tubule, proboscis, maxillary palp, trachea, and wing (figs. S6 to S18). In these tissues, all major expected cell types were identified. In the proboscis and maxillary palp (fig. S7, A and B), we could annotate gustatory and olfactory receptor neurons, mechanosensory neurons, and several glial clusters. All seven olfactory receptors expressed in the maxillary palp were detected. In the wing (fig. S8), we could identify four different neuronal types—gustatory receptor neurons, pheromone-sensing neurons, nociceptive neurons, and mechanosensory neurons—as well as three glial clusters. In the leg (fig. S9), we could distinguish gustatory receptor neurons from two clusters of mechanosensory neurons. In the heart (fig. S10), we found a large proportion of resident hemocytes and muscle cells, with cardiac cells marked by the genes *Hand* and *tinman* constituting a small proportion. In the Malpighian tubule (fig. S11), 15 cell types were identified, including the different principal cells of the stellate and main segments. In the haltere (fig. S13), we identified two clusters of neurons, three clusters of glial cells, and a large population of epithelial cells. In some tissues, cell types formed a big cluster instead of being split into distinct populations. In these cases, we identified genes or pathways that showed a gradient or compartmentalized expression. For example, in the fat body (figs. S14 and S19), the main fat body cells formed one big cluster, but our metabolic pathway enrichment analysis performed through ASAP (*18*) revealed that fatty acid biosynthesis and degradation are in fact compartmentalized, highlighting possible fat body cell heterogeneity in metabolic capacities.

Our crowd annotations with tissue experts also revealed cell types that had not been profiled previously, such as multinucleated muscle cells (Fig. 2B) and two distinct types of nuclei among the main cells in the male accessory gland (fig. S17), a cell type that was previously thought to be uniform. The high number of nuclei analyzed allowed identification of rare cell types. For example, in the testis (Fig. 2C), we identified 25 distinct cell types, covering all expected cell types, including very rare cells, such as germinal proliferation center hub cells (79 nuclei in the Relaxed version, out of 44,621 total testis nuclei).

Next, we compared the distribution of cells between 10x and Smart-seq2 and found a good match based on a coclustering analysis (figs. S20 and S21). Because Smart-seq2 cells only account for a small fraction, our previous annotations focused on 10x cells. The cell-matched coclustering analysis allowed us to transfer annotations from 10x to Smart-seq2 datasets (fig. S20E), using cluster-specific markers as validation (fig. S20F). We also identified genes that were specifically detected using Smart-seq2 thanks to its higher gene detection rate (Fig. 1E and fig. S20G). In summary, the high-throughput 10x datasets form the basis for identifying cell types, whereas the Smart-seq2 datasets facilitate the detection of lowly expressed genes and enable future exploration of cell-specific isoform information.

### Correspondence between dissected tissues and whole head and body

To generate a complete atlas of the fly, we next performed snRNA-seq experiments on whole-head and whole-body samples. Whole-body single-cell experiments were previously performed on less complex animals (*20*, *21*). Full head and body sequencing provides a practical means to assess the impact of mutations or to track disease mechanisms, without having to focus on specific tissues. In addition, it could yield cell types that are not covered by any of the targeted tissue dissections.

In the head, we annotated 81 mostly neuronal cell types (Fig. 3A and fig. S22). In the body, we annotated the top 33 most abundant cell classes, including epithelia, muscle, and ventral nerve cord and peripheral neurons, followed by fat cells, oenocytes, germ line cells, glia, and tracheal cells (Fig. 3B and fig. S23). Many of these cell classes can be further divided into cell types for further annotation (see Fig. 2 and figs. S6 to S18).

Next, we examined how well the head and body samples covered the cell types from the dissected tissues. We analyzed head, body, and tissue samples together, with most of the selected tissues clustering together with the body. We also detected head- and body-enriched clusters (Fig. 3C). One body-specific cluster contained cuticle cells, likely from connective tissue (Fig. 3D). Others were relatively rare cell types in their respective tissues, such as adult stem cells. Conversely, most tissue clusters contained body cells, with only a small number being completely specific to dissected tissues. Because tissue-specific clusters were mostly observed in tissues with high cell coverage, such as the testis and Malpighian tubule, we anticipated that these clusters would also be identified in the body upon sampling a larger number of cells.

For the head, the antenna and proboscis with maxillary palp were dissected for tissue sequencing. Cell types from those two tissues largely overlapped with head cells. Many other cell types, such as central brain cells, including Kenyon cells (*ey*, *prt*) and lamina glia (*repo*, *Optix*), were only detected in the head sample.

To compare our data with existing datasets, we integrated our head snRNA-seq dataset ("head" hereafter) with published brain scRNA-seq data ("brain" hereafter) (*17*, *22*–*24*) (Fig. 3E). Head-specific clusters made up 20% of the cells, including the antennae, photoreceptors, muscle, cone cells, and cuticular cell types, whereas the other 80% were present in clusters containing both head- and brain-derived

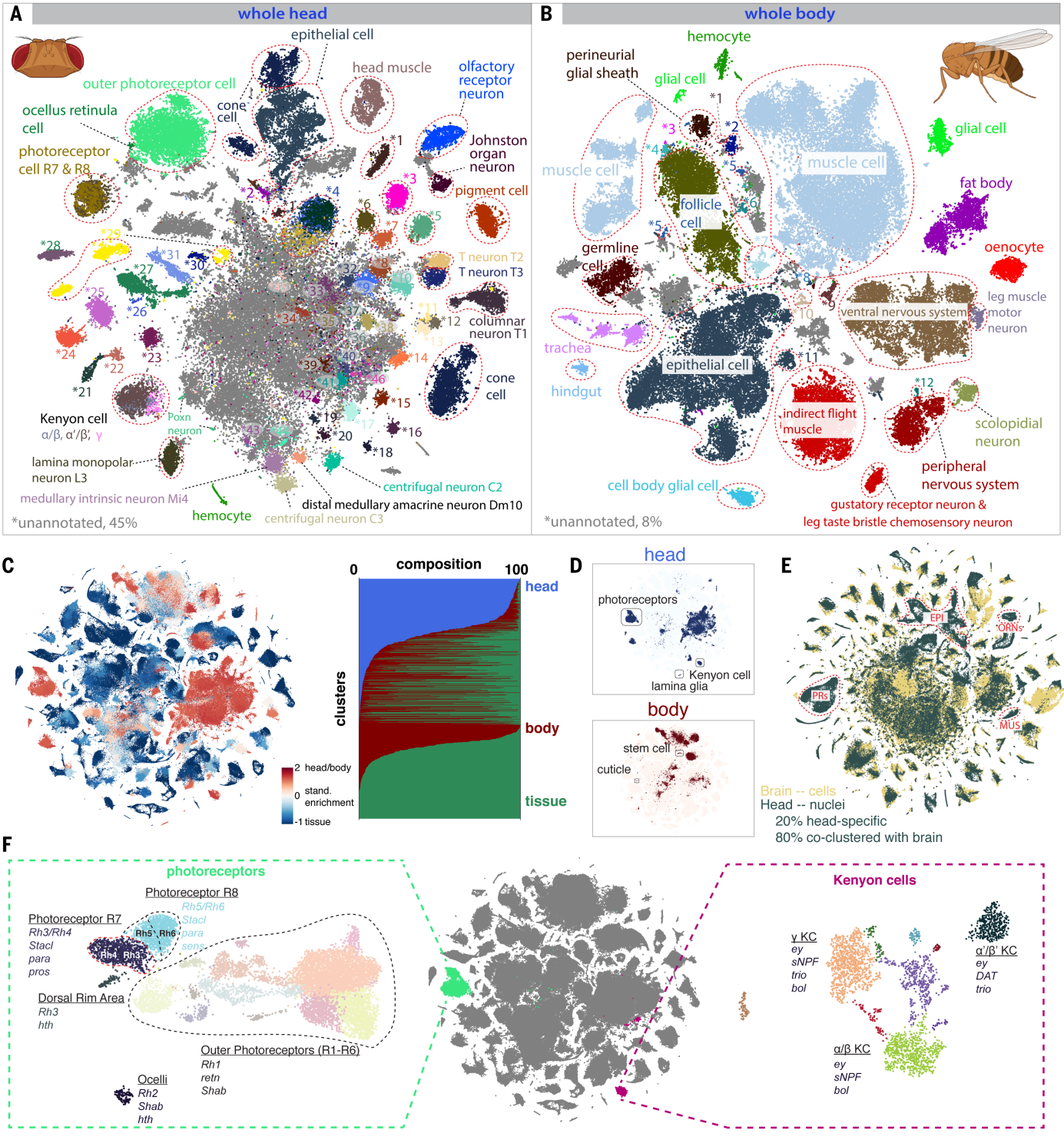



**Fig. 3. Whole-head and whole-body sequencing leads to full coverage of the entire fly.** (**A**) tSNE of the whole-head sample with 81 annotated clusters. See fig. S22 for full cell types. Many cells in the middle (gray) are unannotated, most of which are central brain neurons. (**B**) tSNE of the whole-body sample with 33 annotated clusters, many of which can be further divided into subclusters. Cells in gray are unannotated. See fig. S23 for full cell types. (**C**) tSNE of the entire dataset colored by standardized tissue enrichment, leading to the identification of head- and body-specific clusters, is shown on the left. Stacked bar plot showing tissue composition (head, body, or dissected tissues) for different clusters at Leiden resolution 50 are shown on the right. (**D**) Examples of head- and body-specific clusters. (**E**) Integration of a brain scRNA-seq dataset with the head snRNA-seq for label transfer. Outlined are example clusters revealed by the head snRNA-seq dataset but not by the brain scRNA-seq datasets, including epithelial cells (EPI), photoreceptors (PRs), olfactory receptor neurons (ORNs), and muscle cells (MUS). (**F**) Subclustering analysis reveals types of photoreceptors, including inner and outer photoreceptors, with the inner photoreceptors further splitting into R7 and R8 types, and mushroom body Kenyon cells comprising three distinct types: α/β, α′/β′, and γ.

cells covering the neuronal and glial cell types of the brain. This coclustering across genotypes and protocols underscores the quality and utility of our snRNA-seq data compared with that of scRNA-seq data. Next, we used machine-learning models to predict annotations per cluster, followed by manual curation (*22*). Given the high number of neuron types, additional subclustering was performed on each cluster, identifying subtypes of peptidergic neurons (*dimm*, *Pdf*) and olfactory projection neurons based on *oaz*, *c15*, and *kn*. Finally, we identified many cell types in the optic lobe, including lamina (e.g., L1 to L5), medulla (e.g., Mi1, Mi15), lobula (e.g., LC), and lobula plate (e.g., LPLC). Using *acj6* and *SoxN*, we identified the T4/T5 neurons of the optic lobe that split in T4/T5a-b and T4/T5c-d subtypes by subclustering. A big clump of neurons remained unannotated (Fig. 3A), indicating that our dataset cannot resolve the complexity of the central brain, which may contain hundreds to thousands of neuron types.

Subclustering in the combined dataset separated inner and outer photoreceptors from the dorsal rim area and ocellar photoreceptors, with the inner photoreceptors further splitting into R7 and R8 types, each with *pale* and *yellow* types based on *rhodopsin* expression (Fig. 3F). Additionally, Kenyon cells were split into three types: α/β, α′/β′, and γ (*17*). These cases highlight the resolution in our dataset and the potential of using subclustering to discover rare cell types.

### Cross-tissue analyses allow comparison of cell types by location

Using the whole-body and whole-head sequencing data, we assigned cells to major cell classes (e.g., epithelial cells, neurons, muscle cells, hemocytes), which allowed us to compare common classes across tissues (Fig. 4, A to C, and figs. S24 and S25). First, we compared blood cells across tissues by selecting all *Hml*-positive cells, a known marker for hemocytes (Fig. 4D). Combining hemocytes across tissues revealed a major group of plasmatocytes, the most common hemocyte type (~56%), crystal cells (1.5%; *PPO1*, *PPO2*), and several unknown types (fig. S26, A and B). Looking deeper into the plasmatocytes, we uncovered gradients based on the expression of *Pxn*, *LysX*, *Tep4*, *trol*, and *Nplp2* that can be linked to maturation and plasticity, with *Pxn*-positive cells showing the highest *Hml* expression, whereas *Tep4*, *trol*, and *Nplp2* are prohemocyte markers (*25*). Furthermore, different antimicrobial peptide families such as the Attacins and Cecropins were expressed in different subgroups, indicating specialization. Finally, expression of acetylcholine receptors was specific for a subset of hemocytes, relating to the cholinergic anti-inflammatory pathway as described in humans and mice (*26*). Lamellocytes were not observed in adults as previously suggested (*27*). On the contrary, an unknown hemocyte type expressed *Antp* and *kn* (43 cells, 0.5%) reminiscent of the posterior signaling center in the lymph gland, an organization center previously thought to be absent in the adult (*28*, *29*) (fig. S26B). These findings highlight the value of performing a whole organism–level single-cell analysis and constitute a foundation for investigating the fly immune system in greater detail.

Second, we compared the muscle cells of the different tissues (Fig. 4E and fig. S26, C and D). Muscle cells are syncytia—individual cells containing many nuclei—and to our knowledge have not been profiled by single-cell sequencing before our study. With snRNA-seq, we recovered all known muscle cell types, with specific enrichment in the body, body wall, and leg. This comprehensive view of the fly muscular system highlights a separation of visceral, skeletal, and indirect flight muscle based on the expression of different troponins. Specifically, we discovered gradients of *dysf* and *fln* in the indirect flight muscle, which may indicate regional differences in these very large cells (>1000 nuclei) (fig. S26E). We identified four types of visceral muscle in the gut based on expression of the *AstC*, *Ms*, *Dh31*, and *CCAP* neuropeptide receptors, indicating potential modulators for muscle contraction (*30*). *Ms* and *Dh31* have been described to function in spatially restricted domains (*30*–*32*), suggesting similar domains for *AstC* and *CCAP*. All visceral muscle cells are enriched for the receptor of *Pdf*, a neuropeptide involved in circadian rhythms, pointing toward a function in muscle contraction as well (*33*).

### Transcription factors and cell-type specificity

Our data allow the comparison of gene expression across the entire fly. Clustering cell types showed the germline cells as the most distinct group, followed by neurons (figs. S27 to S32). We calculated marker genes for every cell type using the whole FCA data as background, with 14,240 genes found as a marker for at least one cell type and a median of 638 markers per cell type [minimum: visceral muscle (94); maximum: spermatocyte (7736)]. Notably, markers specific for cell types in a tissue were not always specific in the whole body (fig. S33).

Next, we calculated the tau score of tissue specificity (*34*) for all predicted transcription factors (TFs) (*3*) and identified 500 TFs with a score >0.85, indicating a high specificity for one or very few cell types (Fig. 5A and table S3). Of these TFs, 127 were "CGs" (computed genes), indicating that their functions are poorly studied. We found that the male germline stands out in showing expression of a great number of cell type–specific TFs. This may be related to the broad activation of many genes in late spermatocytes, as discussed below.

Similar analysis across broad cell types (Fig. 5, B and C) identified 156 TFs with high tau scores, for example, the known regulators *grh* for epithelial cells and *repo* for glia, as well as 24 uncharacterized genes. Network visualization shows the grouping of central nervous system (CNS) neurons and sensory organ cells, including many sensory neurons, with shared panneuronal factors such as *onecut* and *scrt* but with each cluster having a distinct set of TFs, such as *ey*, *scro*, and *dati* for CNS neurons and *lz* and *glass* (*gl*) for sensory neurons.

In addition to the specificity of TF expression, we predicted gene regulatory networks based on coexpression and motif enrichment using SCENIC (*31*). Because of the stochasticity of this network inference method, we ran SCENIC 100 times, ranking predicted target genes by their recurrence. This approach selected 6112 "regulons" for 583 specific TFs across all tissues, whereby each regulon consists of the TF, its enriched motif, and the set of target genes that are predicted in at least 5 out of 100 runs. In fat cells, our analysis predicted a regulon for *sugarbabe* (*sug*), a sugar-sensitive TF necessary for the induction of lipogenesis (*32*). In photoreceptors, the analysis identified a *gl* regulon, with key photoreceptor markers such as *Arr1*, *eya*, and multiple rhodopsins as predicted target genes (Fig. 5, D and E) (*33*). The SCENIC predictions for all cell types are available through SCope (https://flycellatlas.org/scope).

A comparative analysis of genes across broad cell types or tissues (Fig. 5F and fig. S34) identified common genes and specifically expressed genes, such as a shared set of 555 housekeeping genes that are expressed in all tissues. The testis has the highest number of specifically expressed genes consistent with previous reports (*34*), followed by the Malpighian tubule and male reproductive glands (fig. S34). These tissue-specific genes seemed to be evolutionarily "younger" based on GenTree age compared with the set of commonly expressed genes that are all present in the common ancestor. This suggests that natural selection works on the tissue specialization level, with the strongest selection on testis, male reproductive tract, and Malpighian tubules (*35*). In addition, this analysis allowed an estimation of transcriptomic similarity or difference measured by the number of shared distinct genes. For example, the two flight appendages, the haltere and wing, share a set of 16 specifically expressed genes, reflecting the evolutionary origin of halteres as a modified wing (*36*) (fig. S34).

### Analysis of sex-biased expression and sex-specialized tissues

To study sex-related differences, we compared male- and female-derived nuclei for all common tissues (fig. S35) and found *roX1/2* and *Yp1/2/3* as the top male- and female-specific genes, respectively. Notably, a large fraction

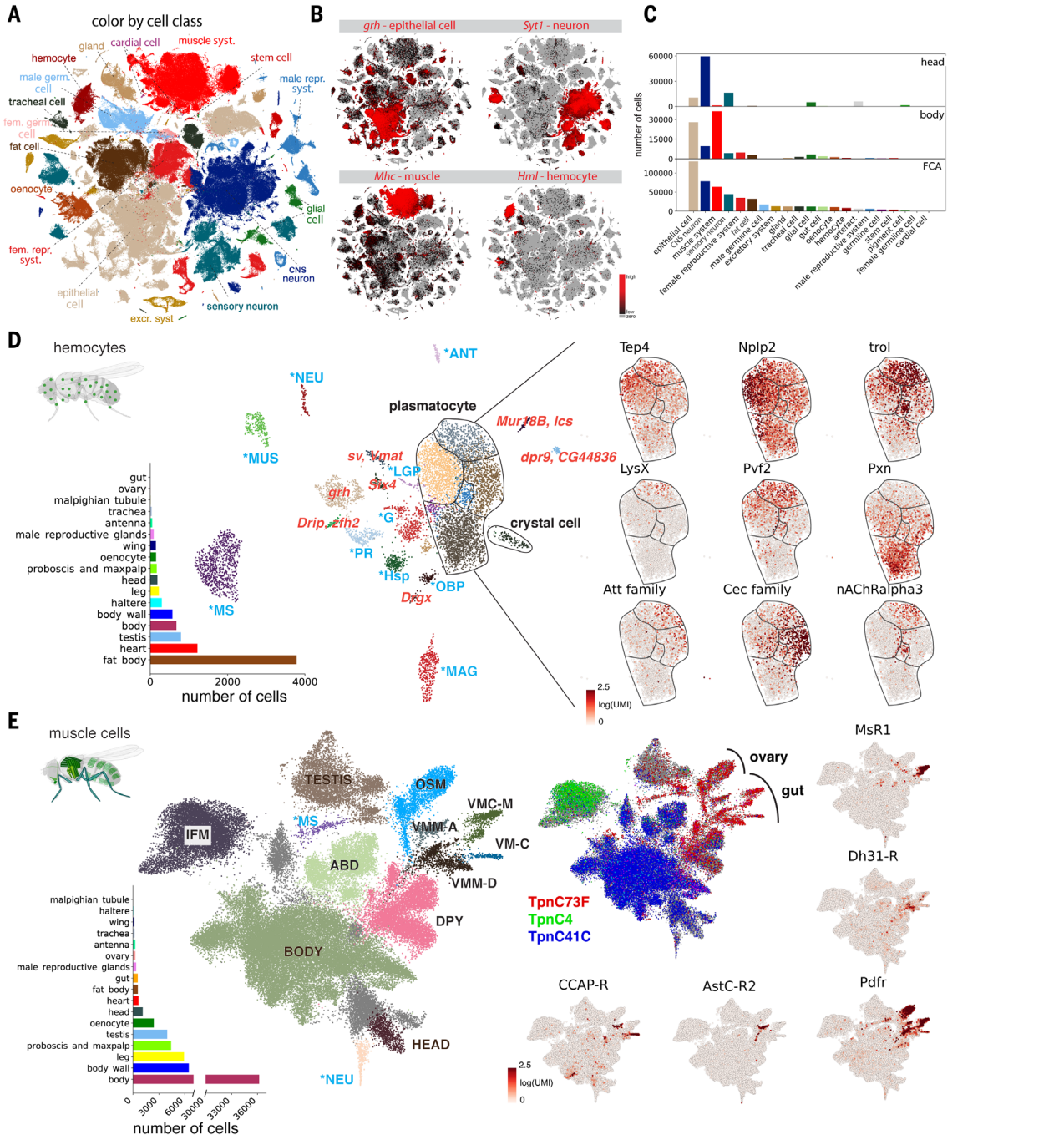



**Fig. 4. Cross-tissue analyses of common cell classes.** (**A**) Overview of main cell classes identified throughout the fly cell atlas. Male repr. syst. and fem. repr. syst., male and female reproductive system; male germ. cell and fem. germ. cell, male and female germline cells. (**B**) tSNE plots showing expression of four markers in four common cell classes. (**C**) Composition of whole-head and whole-body samples, showing a shift from neurons to epithelial and muscle cells. Composition of the entire FCA shows enrichment for rarer cell classes compared with the whole-body sample. (**D**) Cross-tissue analysis of hemocytes reveals different cell states of plasmatocytes. Annotations marked in blue are hemocytes containing markers of different cell types, including lymph gland posterior signaling center (LGP), muscle (MUS), antenna (ANT), neurons (NEU), photoreceptor (PR), male accessory glands (MAG), glia (G), male testis and spermatocyte (MS), odorant-binding proteins (OBP), and heat-shock proteins (Hsp). Other abbreviations show top marker gene(s) in red. Plasmatocytes and crystal cells are indicated. On the right are genes showing compartmentalized expression patterns within the plasmatocyte cluster. (**E**) Cross-tissue analysis of muscle cells reveals subdivision of the visceral muscle cells based on neuropeptide receptors. Annotations marked in blue are muscle cells containing markers of different cell types, including neuron (NEU) and male testis and spermatocyte (MS). Muscle cells from three body parts are indicated: head muscle (HEAD), body muscle (BODY), and testis muscle (TESTIS). Other annotated muscle types include indirect flight muscle (IFM), ovarian sheath muscle (OSM), abdominal visceral muscle (ABD), *dpy* expressing muscle (DPY), visceral muscle of the midgut *AstC-R2* (VMM-A), visceral muscle of the crop *MsR1* (VMC-M), visceral muscle of the midgut *Dh31-R* (VMM-D), and visceral muscle *CCAP-R* (VM-C). *Pdfr* is expressed in all visceral muscle cells, including the ovarian sheath muscle; the other four receptor genes (*AstC-R2*, *MsR1*, *Dh31-R*, *CCAP-R*) are expressed in different gut visceral muscle types.

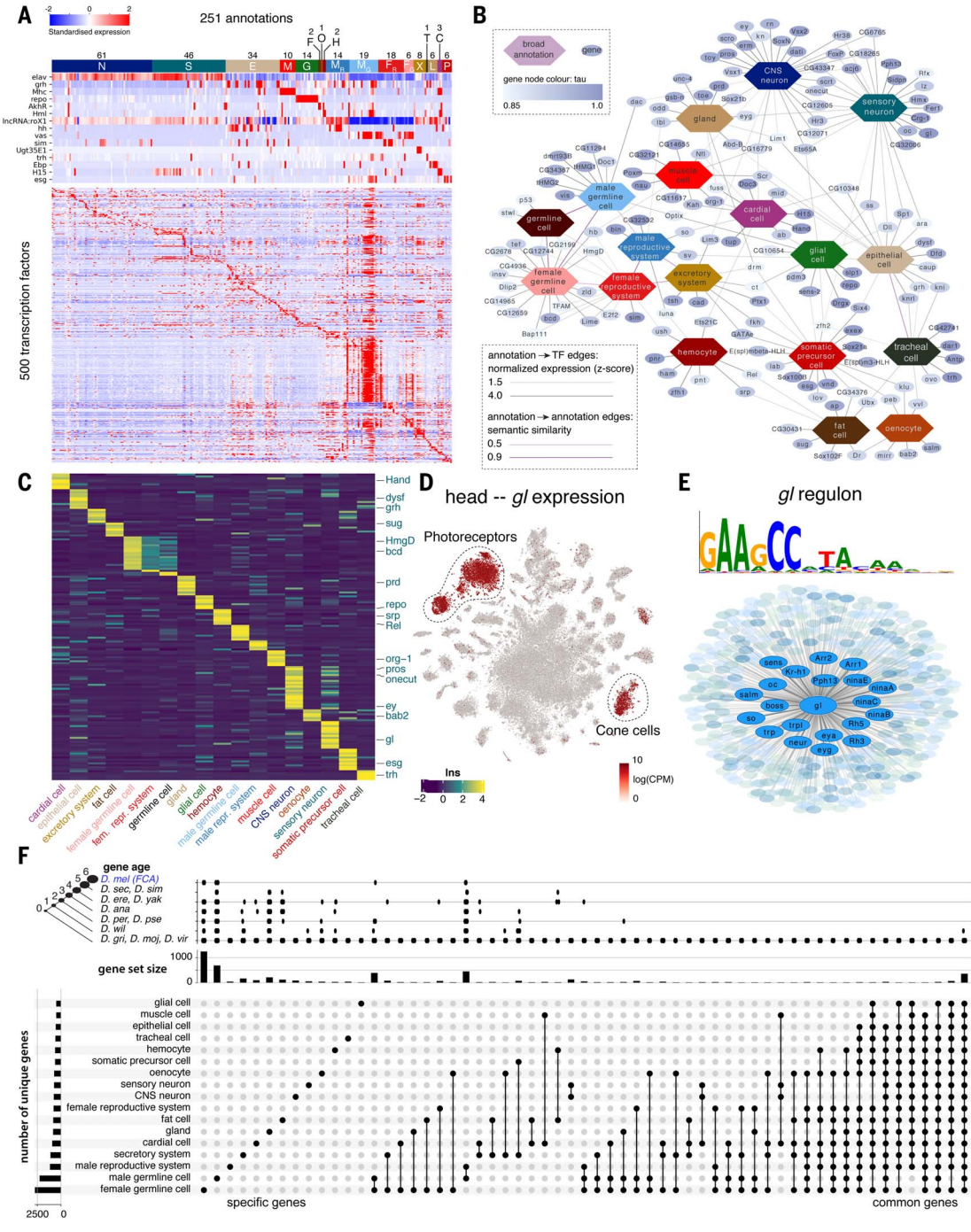



**Fig. 5. TF pleiotropy versus cell-type specificity.**
(**A**) Heatmap showing the expression of key marker genes and distinctive TF profiles for each of the annotated cell types. TFs were selected based on tau score. Cell types were grouped based on hierarchical terms: CNS neurons (N), sensory organ cells (S), epithelial cells (E), muscle cells (M), glia (G), fat cells (F), oenocytes (O), hemocytes (H), (fe)male reproductive system and germline (MR, MG, FR, FG), excretory system (X), tracheal cell (T), gland (L), cardiac cell (C), and somatic precursor cell (P). (**B**) A network analysis of TFs and cell classes based on similarity of ontology terms, reveals specific and shared TFs across the individual tissues. (**C**) Heatmap showing the expression of specific TFs per cell class. Factors from the literature are highlighted. (**D**) *Glass* is specifically expressed in photoreceptors and cone cells in the head. (**E**) Overview of the *Glass* regulon of 444 target genes, highlighting known photoreceptor marker genes. (**F**) Gene expression comparison across broad cell types. Only sets with more than 10 genes are shown. The left bar graph shows the number of uniquely expressed genes for each tissue. The top bar graph shows the gene age in branches, ranging from the common ancestor to *D. melanogaster*–specific genes (http://gentree.ioz.ac.cn). See fig. S34 for a tissue-based comparison.

of genes with male-enriched expression were uncharacterized (*37*). The primary sex-determination pathway in somatic cells leads to sex-specific splicing of *doublesex* (*dsx*) to encode female- or male-specific TFs (*38*) (Fig. 6A). Consistent with this, we found *dsx* expression in a largely non–sex-specific pattern, whereas many other genes showed sex-biased expression (Fig. 6B).

Next, we performed differential expression between sexes for all cell types. Notably, cell types tended to show either high female or male bias, but not both (Fig. 6, B and C). We found strong female bias in the excretory system, including the principal and stellate cells of the Malpighian tubule and in the pericardial nephrocytes (Fig. 6C). Female-biased genes (i.e., *Ics* and *whe*) were differentially expressed under high-salt conditions, suggesting sex-bias in nephric ion transport. Across cell types, sex-biased expression strongly correlated with *dsx* expression (Fig. 6D) (*39*), consistent with the role of Dsx as a key regulator.

Among all tissues in the adult fly, those best characterized that have ongoing cellular differentiation are the gut, ovaries, and testis. Trajectory analysis has been performed on the gut and ovary stem cell lineages in previous studies (*40*–*42*), and our FCA data on gut and ovary accurately coclustered with these published datasets (figs. S36 and S37). Therefore, we focused on the testis plus seminal vesicle as a case study. The testis has two populations of stem cells, the somatic cyst stem cells (CySCs) that produce cell types with supporting roles essential to spermatogenesis, and the germline stem cells (GSCs) that produce haploid

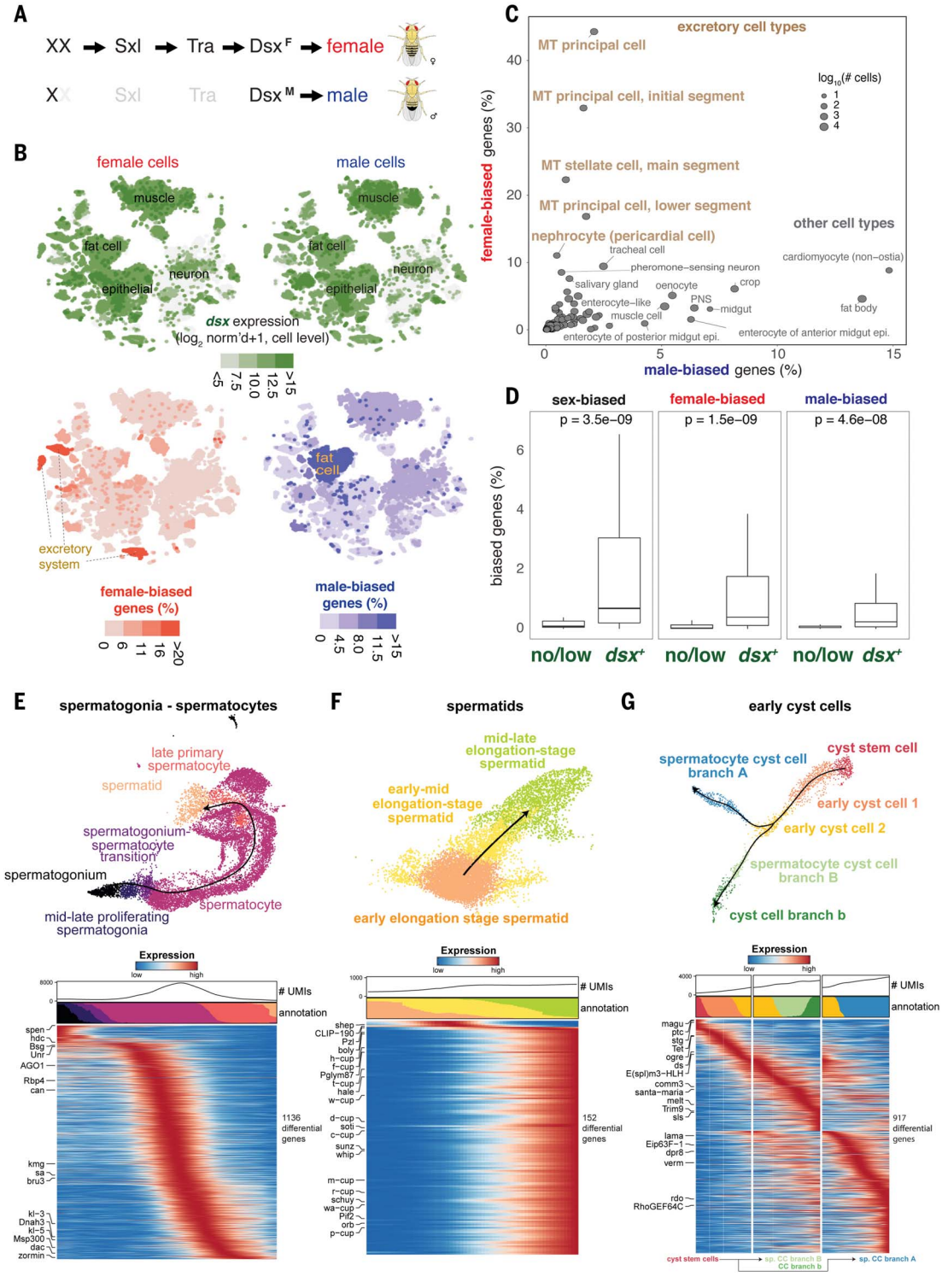



**Fig. 6. Sex-biased expression and trajectory analysis of testis cell lineages.** (**A**) Simplified sex-determination pathway. Sex-chromosome karyotype (XX) activates Sex-lethal (Sxl), which regulates transformer (Tra), resulting in a female Dsx isoform ($Dsx^F$). In XY (or X0) flies, Sxl and Tra are inactive (light gray) and the male-specific $Dsx^M$ is produced. (**B**) *dsx* expression and female- and male-biased expression projected onto tSNE plots of all female (top left) and male (top right) cells except reproductive tissue cells (tables S4 and S5). Female-biased (bottom left) and male-biased (bottom right) expression measured as the percentage of genes in the cluster showing biased expression in favor of the respective sex (table S6) are also shown. The percentage values were computed for each annotated cluster, and those cluster-level values were projected onto the individual cells in the corresponding clusters. For all four tSNE plots, values outside the scale in the heatmap key are represented by the closest extreme color (> and < signs in the scale). (**C**) Scatter plot of female- and male-bias values across non–reproductive cell clusters defined as percentage of sex-biased genes (at least twofold change with false discovery rate <0.05 based on the Wilcoxon test and Benjamini-Hochberg correction) in the cluster (table S6). Data point size indicates cell numbers per cluster (key). Selected clusters are labeled, with those from excretory cells highlighted (brown). MT, Malpighian tubule. (**D**) Box plots showing the relationship between *dsx* gene expression and sex-biased expression (table S5). Clusters (B) were partitioned into the set of clusters with *dsx* expression ($dsx^+$) or not (no/low) using *dsx* expression in germ cells as an expression cutoff. Each box shows hinges at first and third quartiles and median in the middle. The upper whisker extends from the upper hinge to the largest value no further than 1.5 times the interquartile range (IQR) from the hinge (where the IQR is the distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value, at most 1.5 times the IQR of the hinge. Outliers are not shown. *p* values are based on the Wilcoxon test. (**E** to **G**) Trajectory of testis subsets. We used slingshot to infer a possibly branching trajectory for spermatogonia-spermatocytes (E), spermatids (F), and early cyst cells (G). Shown are the trajectories on a UMAP (top) and the expression patterns of the strongest differentially expressed genes, together with the smoothed proportions of annotated cells and average number of UMIs along the trajectory (bottom).

sperm (Fig. 2C). The main testis analysis (Fig. 2C) revealed transitions from GSCs and proliferating spermatogonia to spermatocytes, then to maturing spermatids, and finally to late elongation stage spermatids.

We further performed trajectory inference on spermatocytes and spermatids separately (Fig. 6, E and F). As expected, the spermatocyte stage featured a continuous increase in the number of genes being transcribed (Fig. 6E), with many of the strongly up-regulated genes (*kmg*, *Rbp4*, *fzo*, *can*, *sa*, and, for later spermatocytes, Y-linked fertility factors *kl-3* and *kl-5*) not substantially expressed in any other cell type. Late spermatocytes, however, showed expression of marker genes from many other cell types like somatic cells (*Upd1*, *eya*), epithelial cells (*grh*), muscle (*Mhc*), or hemocytes (*Hml*) (Fig. 5A), although their expression level was lower than in their marked cell type. Early spermatids are in transcriptional quiescence, as can be seen by a very low number of nuclear transcripts (Fig. 6F; low UMI), followed by a burst of new transcription in elongating spermatids, including many *cup* genes. In the somatic cyst cell lineage, we found CySCs expressing the cell cycle marker *string* that were transitioning into postmitotic (no *string* expression) early cyst cells and branching into two related clusters of cyst cells likely associated with spermatocytes (Fig. 6G).

## Discussion

Recent technological developments have enabled single-cell transcriptomic atlases of *Caenorhabditis elegans* (*21*) and selected tissues in mice and humans (*43*–*46*). Here, we provide a single-cell transcriptomic map of the entire adult *D. melanogaster*, a premier model organism for studies of fundamental and evolutionarily conserved biological mechanisms. The FCA provides a resource for the *Drosophila* community as a reference for studies of gene function at single-cell resolution.

A key challenge in large-scale cell atlas projects is the definition of cell types. We addressed this using a consensus-based voting system across multiple resolutions. An FCA cell type is thus defined as a transcriptomic cluster detected at any clustering resolution that could be separated by the expression of known marker genes from other clusters. Further, all annotations were manually curated by tissue experts, leading to a high-confidence dataset with more than 250 annotated cell types. We note differences in annotation depth for different cell groups, with some cell types only linked to broad classes (e.g., epithelial cell), in contrast to other, more detailed cell types (e.g., different olfactory receptor neurons). We also note that although many marker genes are useful in identifying cell types, some marker-gene expression was not congruent with cluster expression. This can be caused by discrepancies between mRNA and expression or by mistakes that were made in the literature. These examples highlight the need for and the opportunities presented by Tabula *Drosophilae* to serve as the basis for future validation.

We have generated lists of marker genes per cell type with different levels of specificity, ranging from the tissue-wide to the animal-wide level. This distinctive level of precision presents a blueprint for future integration with other data modalities such as single-cell assay for transposase-accessible chromatin (ATAC)–seq (*47*) and spatial omics and for generating cell-type reporter lines to study new cellular functions. Furthermore, the large number of uncharacterized genes that show cell-type specific, sex-biased, or trajectory-dependent expression provides the foundation for many follow-up studies. Our analysis also presents several technical novelties, including the use of reproducible Nextflow pipelines (VSN, https://github.com/vib-singlecell-nf), the availability of raw and processed datasets for users to explore, and the development of a crowd-annotation platform with voting, comments, and references through SCope (https://flycellatlas.org/scope), linked to an online analysis platform in ASAP (https://asap.epfl.ch/fca). These elements may inspire future atlas projects. Given the work in other model organisms, we also envision a use for the FCA data in cross-species studies. Furthermore, Tabula *Drosophilae* is fully linked to existing *Drosophila* databases by a common vocabulary, benefitting its use and integration in future projects. Finally, all FCA data are freely available for further analysis through multiple portals and can be downloaded for custom analysis using other single-cell tools (fig. S1; links available on www.flycellatlas.org).

## Materials and methods summary

For most samples, 5-day-old adult $w^{1118}$ flies were used for both male and female tissues except sex-specific tissues. We estimated the required tissue number based on three factors: total cell number in each tissue, targeted cell number, and recovery rate. Fly tissues were dissected by different dissection labs, flash-frozen using liquid nitrogen, stored at −80°C, and then processed using the same platform. The snRNA-seq was largely adapted from our recently published protocol (*11*). All libraries were sequenced using Illumina NovaSeq 6000.

Before read alignment, the raw FASTQ files from 10x Genomics were processed with the index-hopping-filter software. A Cell Ranger (version 3.1.0) index was built from a pre-mRNA GTF from the Flybase version r6.31. To ensure reproducibility of the 10x Genomics data processing, all the analyses from raw counts to final processed files were performed using the Nextflow VSN-Pipelines. Two versions of the processed data were generated: Relaxed and Stringent. For most analyses, we focused on the Stringent dataset, which should be used as a default for new users. Leiden clustering was performed for a wide range of resolutions, and large clusters were subclustered. Crowd annotation by tissue experts was performed across all cluster resolutions in SCope, using terms from the FBbt ontology. ASAP was used to perform more specific analyses. Loom and H5AD files are available for download, visualization in SCope and cellxgene, and detailed analyses in ASAP.

10x Genomics and Smart-seq2 data were integrated using Harmony. For brain-head data integration, annotations were added using computational approaches, and all annotations were then manually curated in jamborees. For common cell analyses, hemocytes and muscle cells were extracted from different tissues, and harmony was used to remove batch effects. Cell type–specific TFs were identified using the tau factor, and TF regulons were predicted using SCENIC. For the sex-bias analysis, sex-specific cells were removed, and about 270,000 cells from 176 annotated clusters were used to calculate male- and female-bias genes for each cluster. Trajectory analyses of the testes were performed using slingshot. The strongest differentially expressed genes along the trajectories were calculated and shown using heatmaps.

Detailed descriptions of all experimental protocols and analyses are provided in the supplementary materials.

## REFERENCES AND NOTES


1. T. H. Morgan, Sex limited inheritance in *Drosophila*. *Science* **32**, 120–122 (1910). doi: 10.1126/science.32.812.120; pmid: 17759620
2. M. D. Adams *et al*., The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185–2195 (2000). doi: 10.1126/science.287.5461.2185; pmid: 10731132
3. A. Larkin *et al*., FlyBase: Updates to the *Drosophila melanogaster* knowledge base. *Nucleic Acids Res.* **49**, D899–D907 (2021). doi: 10.1093/nar/gkaa1026; pmid: 33219682
4. R. Lyne *et al*., FlyMine: An integrated database for *Drosophila* and *Anopheles* genomics. *Genome Biol.* **8**, R129 (2007). doi: 10.1186/gb-2007-8-7-r129; pmid: 17615057
5. A. Jenett *et al*., A GAL4-driver line resource for *Drosophila* neurobiology. *Cell Rep.* **2**, 991–1001 (2012). doi: 10.1016/j.celrep.2012.09.011; pmid: 23063364
6. N. Milyaev *et al*., The Virtual Fly Brain browser and query interface. *Bioinformatics* **28**, 411–415 (2012). doi: 10.1093/bioinformatics/btr677; pmid: 22180411
7. M. M. Kudron *et al*., The ModERN Resource: Genome-wide binding profiles for hundreds of *Drosophila* and *Caenorhabditis elegans* transcription factors. *Genetics* **208**, 937–949 (2018). doi: 10.1534/genetics.117.300657; pmid: 29284660
8. S. Roy *et al*., Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science* **330**, 1787–1797 (2010). doi: 10.1126/science.1198374; pmid: 21177974
9. V. R. Chintapalli, J. Wang, J. A. T. Dow, Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. *Nat. Genet.* **39**, 715–720 (2007). doi: 10.1038/ng2049; pmid: 17534367
10. H. Li, Single-cell RNA sequencing in *Drosophila*: Technologies and applications. *Wiley Interdiscip. Rev. Dev. Biol.* **10**, e396 (2021). doi: 10.1002/wdev.396; pmid: 32940008
11. C. N. McLaughlin *et al*., Single-cell transcriptomes of developing and adult olfactory receptor neurons in *Drosophila*. *eLife* **10**, e63856 (2021). doi: 10.7554/eLife.63856; pmid: 33555999
12. G. X. Y. Zheng *et al*., Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017). doi: 10.1038/ncomms14049; pmid: 28091601

13. S. Picelli *et al*., Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10**, 1096–1098 (2013). doi: 10.1038/nmeth.2639; pmid: 24056875
14. M. De Waegeneer, C. C. Flerin, K. Davie, G. Hulselmans, vib-singlecell-nf/vsn-pipelines: v0.26.0. *Zenodo* (2021); https://doi.org/10.5281/zenodo.5055627.
15. S. Yang *et al*., Decontamination of ambient RNA in single-cell RNA-seq with DecontX. *Genome Biol.* **21**, 57 (2020). doi: 10.1186/s13059-020-1950-6; pmid: 32138770
16. I. Korsunsky *et al*., Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019). doi: 10.1038/s41592-019-0619-0; pmid: 31740819
17. K. Davie *et al*., A single-cell transcriptome atlas of the aging *Drosophila* brain. *Cell* **174**, 982–998.e20 (2018). doi: 10.1016/j.cell.2018.05.057; pmid: 29909982
18. F. P. A. David, M. Litovchenko, B. Deplancke, V. Gardeux, ASAP 2020 update: An open, scalable and interactive web-based portal for (single-cell) omics analyses. *Nucleic Acids Res.* **48**, W403–W414 (2020). doi: 10.1093/nar/gkaa412; pmid: 32449934
19. M. Costa, S. Reeve, G. Grumbling, D. Osumi-Sutherland, The *Drosophila* anatomy ontology. *J. Biomed. Semantics* **4**, 32 (2013). doi: 10.1186/2041-1480-4-32; pmid: 24139062
20. S. Levy *et al*., A stony coral cell atlas illuminates the molecular and cellular basis of coral symbiosis, calcification, and immunity. *Cell* **184**, 2973–2987.e18 (2021). doi: 10.1016/j.cell.2021.04.005; pmid: 33945788
21. J. Cao *et al*., Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* **357**, 661–667 (2017). doi: 10.1126/science.aam8940; pmid: 28818938
22. M. N. Özel *et al*., Neuronal diversity and convergence in a visual system developmental atlas. *Nature* **589**, 88–95 (2021). doi: 10.1038/s41586-020-2879-3; pmid: 33149298
23. H. Li *et al*., Classifying *Drosophila* olfactory projection neuron subtypes by single-cell RNA sequencing. *Cell* **171**, 1206–1220.e22 (2017). doi: 10.1016/j.cell.2017.10.019; pmid: 29149607
24. Y. Z. Kurmangaliyev, J. Yoo, J. Valdes-Aleman, P. Sanfilippo, S. L. Zipursky, Transcriptional programs of circuit assembly in the drosophila visual system. *Neuron* **108**, 1045–1057.e6 (2020). doi: 10.1016/j.neuron.2020.10.006; pmid: 33125872
25. B. Cho *et al*., Single-cell transcriptome maps of myeloid blood cell lineages in *Drosophila*. *Nat. Commun.* **11**, 4483 (2020). doi: 10.1038/s41467-020-18135-y; pmid: 32900993
26. V. A. Pavlov, K. J. Tracey, The cholinergic anti-inflammatory pathway. *Brain Behav. Immun.* **19**, 493–499 (2005). doi: 10.1016/j.bbi.2005.03.015; pmid: 15922555
27. P. Sanchez Bosch *et al*., Adult *Drosophila* lack hematopoiesis but rely on a blood cell reservoir at the respiratory epithelia to relay infection signals to surrounding tissues. *Dev. Cell* **51**, 787–803.e5 (2019). doi: 10.1016/j.devcel.2019.10.017; pmid: 31735669
28. J. Krzemień *et al*., Control of blood cell homeostasis in *Drosophila* larvae by the posterior signalling centre. *Nature* **446**, 325–328 (2007). doi: 10.1038/nature05650; pmid: 17361184
29. L. Mandal, J. A. Martinez-Agosto, C. J. Evans, V. Hartenstein, U. Banerjee, A Hedgehog- and Antennapedia-dependent niche maintains *Drosophila* haematopoietic precursors. *Nature* **446**, 320–324 (2007). doi: 10.1038/nature05585; pmid: 17361183
30. R. J. Siviter *et al*., Expression and functional characterization of a *Drosophila* neuropeptide precursor with homology to mammalian preprotachykinin A. *J. Biol. Chem.* **275**, 23273–23280 (2000). doi: 10.1074/jbc.M002875200; pmid: 10801863
31. S. Aibar *et al*., SCENIC: Single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017). doi: 10.1038/nmeth.4463; pmid: 28991892
32. J. Mattila, V. Hietakangas, Regulation of carbohydrate energy metabolism in *Drosophila melanogaster*. *Genetics* **207**, 1231–1253 (2017). pmid: 29203701
33. K. Moses, M. C. Ellis, G. M. Rubin, The glass gene encodes a zinc-finger protein required by *Drosophila* photoreceptor cells. *Nature* **340**, 531–536 (1989). doi: 10.1038/340531a0; pmid: 2770860
34. H. Kaessmann, Origins, evolution, and phenotypic impact of new genes. *Genome Res.* **20**, 1313–1326 (2010). doi: 10.1101/gr.101386.109; pmid: 20651121
35. Y. Shao *et al*., GenTree, an integrated resource for analyzing the evolution and function of primate-specific coding genes. *Genome Res.* **29**, 682–696 (2019). doi: 10.1101/gr.238733.118; pmid: 30862647
36. E. B. Lewis, A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565–570 (1978). doi: 10.1038/276565a0; pmid: 103000
37. J. Andrews *et al*., Gene discovery using computational and microarray analysis of transcription in the *Drosophila melanogaster* testis. *Genome Res.* **10**, 2030–2043 (2000). doi: 10.1101/gr.159800; pmid: 11116097
38. H. K. Salz, J. W. Erickson, Sex determination in *Drosophila*: The view from the top. *Fly* **4**, 60–70 (2010). doi: 10.4161/fly.4.1.11277; pmid: 20160499
39. E. Clough *et al*., Sex- and tissue-specific functions of *Drosophila* doublesex transcription factor target genes. *Dev. Cell* **31**, 761–773 (2014). doi: 10.1016/j.devcel.2014.11.021; pmid: 25535918
40. R.-J. Hung *et al*., A cell atlas of the adult *Drosophila* midgut. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 1514–1523 (2020). doi: 10.1073/pnas.1916820117; pmid: 31915294
41. K. Rust *et al*., A single-cell atlas and lineage analysis of the adult *Drosophila* ovary. *Nat. Commun.* **11**, 5628 (2020). doi: 10.1038/s41467-020-19361-0; pmid: 33159074
42. A. Jevitt *et al*., A single-cell atlas of adult Drosophila ovary identifies transcriptional programs and somatic cell lineage regulating oogenesis. *PLOS Biol.* **18**, e3000538 (2020). doi: 10.1371/journal.pbio.3000538; pmid: 32339165
43. Tabula Muris Consortium, Single-cell transcriptomics of 20 mouse organs creates a *Tabula Muris*. *Nature* **562**, 367–372 (2018). doi: 10.1038/s41586-018-0590-4; pmid: 30283141
44. X. Han *et al*., Mapping the mouse cell atlas by microwell-seq. *Cell* **173**, 1307 (2018). doi: 10.1016/j.cell.2018.05.012; pmid: 29775597
45. J. Cao *et al*., A human cell atlas of fetal gene expression. *Science* **370**, eaba7721 (2020). doi: 10.1126/science.aba7721; pmid: 33184181
46. X. Han *et al*., Construction of a human cell landscape at single-cell level. *Nature* **581**, 303–309 (2020). doi: 10.1038/s41586-020-2157-4; pmid: 32214235
47. J. Janssens *et al*., Decoding gene regulation in the fly brain. *Nature* **601**, 630–636 (2022). doi: 10.1038/s41586-021-04262-z; pmid: 34987221


## ACKNOWLEDGMENTS

We thank the fly community for the enthusiastic support for this project and W. Burkholder, C. Murphy, and K. Vogelaers for coordinating FCA and all Jamboree meetings. **Funding:** The sequencing was supported by the Chan Zuckerberg Biohub (S.R.Q.), Genentech Inc. (H.J.), National Institutes of Health intramural 1ZIADK015600 (B.O.), national funds through the FCT in the framework of the financing of the Norma Transitória DL 57/2016 (Z.C.-S.), Wu Tsai Neurosciences Institute at Stanford (S.R.Q. and L.L.), and the Howard Hughes Medical Institute and a National Institutes of Health grant (L.L.). Computational work was supported by the KU Leuven and the Flemish Supercomputer Center (VSC) (S.A.) and EPFL (B.D.). FCA Consortium funding is provided in the supplementary materials. **Author contributions:** All byline authors are members of the FCA Consortium. See FCA Consortium contribution list in the supplementary materials. **Competing interests:** H.J., N. S. Katheder, and X. T. Cai are employees of Genentech, Inc. The other authors declare no competing interests. **Data and materials availability:** All data are available for user-friendly querying at https://flycellatlas.org/scope and for custom analyses at https://flycellatlas.org/asap. For each tissue, a cellxgene portal is also available (www.flycellatlas.org). Raw data and count matrices can be downloaded from ArrayExpress (accession number E-MTAB-10519 for 10x, and E-MTAB-10628 for Smart-seq2; the same accession numbers are available at EBI Single Cell Expression Atlas at www.ebi.ac.uk/gxa/sc). Files with expression data, clustering, embeddings, and annotation can be downloaded for each tissue, or all data combined, in h5ad and LoomX formats from www.flycellatlas.org. Three supplementary figures describe how to access and explore FCA data: fig. S1 for summary of data availability and figs. S2 and S3 for how to use SCope and ASAP. We also include a video tutorial for using Scope (www.youtube.com/watch?v=yNETQVaSJYM&t=349s). Analysis codes are at Github (https://github.com/flycellatlas). Dataset access is as follows: GSE107451 (scRNA-seq adult fly brain), GSE120537 (scRNA-seq adult fly gut), and GSE136162, GSE146040, and GSE131971 (scRNA-seq adult ovary). The neural network is from (*22*) (see appendix 1).

**FCA Consortium authors (alphabetically by last name)**
Stein Aerts[4,5], Devika Agarwal[32], Yasir Ahmed-Braimah[33], Aaron M. Allen[23], Michelle Arbeitman[34], Majd M. Ariss[35], Jordan Augsburger[13], Kumar Ayush[12], Catherine C. Baker[36], Torsten Banisch[37], Cameron W. Berry[24], Katja Birker[38], Rolf Bodmer[38], Benjamin Bolival[36], Susanna E. Brantley[36], Maria Brbić[12,8], Julie A. Brill[39,76], Nora C. Brown[40], Katja Brueckner[13]†, Norene A. Buehner[40], Xiaoyu Tracy Cai[31], Rita Cardoso-Figueiredo[20], Zita Carvalho-Santos[20], Fernando Casares[41], Amy Chang[42], Thomas R. Clandinin[43], Sheela Crasta[6,7,8], Fabrice P. A. David[9,10], Kristofer Davie[4], Bart Deplancke[9,10], Claude Desplan[44], Angela M. Detweiler[8], Darshan B. Dhakan[20], Stephen DiNardo[27,28], Erika Donà[45], Julian A. T. Dow[30], Stefanie Engert[42], Swann Floc'hlay[4,5], Margaret T. Fuller[24], Anthony Galenza[29], Vincent Gardeux[9,10], Nancy George[46], Amanda J. González-Segarra[42], Stephen F. Goodwin[23], Andrew K. Groves[47,48], Samantha Gumbin[29], Yanmeng Guo[49,50], Devon E. Harris[36], Yael Heifetz[51], Stephen L. Holtz[52], Felix Horns[53], Bruno Hudry[54], Ruei-Jiun Hung[15], Yuh Nung Jan[49,50], Jasper Janssens[4,5], Heinrich Jasper[31], Jacob S. Jaszczak[49,50], Gregory S. X. E. Jefferis[45], Robert C. Jones[6,7,8], Jim Karkanias[8], Timothy L. Karr[55], Nadja Sandra Katheder[31], James Kezos[38], Anna A. Kim[29,56,57], Seung K. Kim[24,58], Lutz Kockel[24], Sai Saroja Kolluru[6,7,8], Nikolaos Konstantinides[59], Thomas B. Kornberg[60], Henry M. Krause[61], Andrew Thomas Labott[29], Meghan Laturney[42], Ruth Lehmann[62]‡, Sarah Leinwand[42], Jure Leskovec[12,8], Hongjie Li[1,2,3], Jiefu Li[1], Joshua Shing Shun Li[15], Kai Li[49,50], Ke Li[49,50], Liying Li[49,50], Tun Li[49,50], Maria Litovchenko[9,10], Han-Hsuan Liu[49,50], Yifang Liu[15], Tzu-Chiao Lu[2], Liqun Luo[1], Sharvani Mahadevaraju[22], Jonathan Manning[63], Anjeli Mase[13], Mikaela Matera-Vatnick[40], Neuza Reis Matias[36], Erika L. Matunis[26], Caitlin E. McDonough-Goldstein[64,77], Aaron McGeever[8], Alex D. McLachlan[65], Colleen N. McLaughlin[1], Paola Moreno-Roman[29], Norma Neff[8], Megan Neville[23], Sang Ngo[29], Tanja Nielsen[38], Todd G. Nystul[19], Caitlin E. O'Brien[49,50], Lucy Erin O'Brien[29], Brian Oliver[22], David Osumi-Sutherland[66], Mehmet Neset Özel[67], Soumitra Pal[21], Irene Papatheodorou[46], Norbert Perrimon[15,16], Maja Petkovic[49,50], Clare Pilgrim[68], Angela Oliveira Pisco[8], Teresa M. Przytycka[21], Stephen R. Quake[6,7,8], Carolina Reisenman[42], Carlos Ribeiro[20], Katja Rust[18,19], Wouter Saelens[9,10], Erin Nicole Sanders[29], Gilberto dos Santos[69], Frank Schnorrer[17], Kristin Scott[42], Aparna Sherlekar[29], Jiwon Shim[14], Philip Shiu[42], David Sims[32], Rene V. Sit[8], Maija Slaidina[70]§, Harold E. Smith[71], Katina Spanier[4,5], Gabriella Sterne[42], Yu-Han Su[29], Daniel Sutton[72], Marco Tamayo[38], Michelle Tan[8], Ibrahim Tastekin[20], Sudhir Gopal Tattikota[15,16], Christoph Treiber[73], David Vacek[1], Georg Vogler[38], Scott Waddell[73], Maxime De Waegeneer[4,5], Wanpeng Wang[60], Helen White-Cooper[25], Rachel I. Wilson[52], Mariana F. Wolfner[40], Yiu-Cheung E. Wong[36], Anthony Xie[1], Qijing Xie[1], Jun Xu[15], Shinya Yamamoto[47,74], Jia Yan[8], Zepeng Yao[42], Kazuki Yoda[29], Ruijun Zhu[49,50], Robert P. Zinzen[75].

[1]Howard Hughes Medical Institute, Department of Biology, Stanford University, Stanford, CA 94305, USA. [2]Huffington Center on Aging, Baylor College of Medicine, Houston, TX 77030, USA. [3]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA. [4]VIB-KU Leuven Center for Brain and Disease Research, KU Leuven, 3000 Leuven, Belgium. [5]Laboratory of Computational Biology, Department of Human Genetics, KU Leuven, 3000 Leuven, Belgium. [6]Department of Bioengineering, Stanford University, Stanford, CA, USA. [7]Department of Applied Physics, Stanford University, Stanford, CA, USA. [8]Chan Zuckerberg Biohub, San Francisco CA, USA. [9]Laboratory of Systems Biology and Genetics, Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland. [10]Swiss Institute of Bioinformatics, CH-1015 Lausanne, Switzerland. [11]Bioinformatics Competence Center, EPFL, Switzerland. [12]Department of Computer Science, Stanford University, Stanford, CA 94305, USA. [13]Department of Cell and Tissue Biology, University of California, San Francisco, CA 94143, USA. [14]Department of Life Science, College of Natural Science, Hanyang University, 04763 Seoul, Republic of Korea. [15]Department of Genetics, Blavatnik Institute, Harvard Medical School, Boston, MA 02115, USA. [16]Howard Hughes Medical Institute, Harvard Medical School, Boston, MA, USA. [17]Aix-Marseille University, CNRS, IBDM (UMR 7288), Turing Centre for Living Systems, 13009 Marseille, France. [18]Institute of Physiology and Pathophysiology, Department of Molecular Cell Physiology, Philipps-University, Marburg, Germany. [19]Department of Anatomy, University of California, San Francisco, CA 94143, USA. [20]Behavior and Metabolism Laboratory, Champalimaud Research, Champalimaud Centre for the Unknown, Lisbon, Portugal. [21]National Center of Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA. [22]Laboratory of Cellular and Developmental Biology, National Institute of Diabetes and Kidney and Digestive Diseases, National Institutes of Health, Bethesda, MD 20892, USA. [23]Centre for Neural Circuits and Behaviour, University of Oxford, Oxford OX1 3SR, UK. [24]Department of Developmental Biology and Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA. [25]Molecular Biosciences Division, Cardiff University, Cardiff CF10 3AX, UK. [26]Department of Cell Biology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. [27]Perelman School of Medicine, The University of

Pennsylvania, Philadelphia, PA 19104, USA. [28]The Penn Institute for Regenerative Medicine, Philadelphia, PA 19104, USA. [29]Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, CA 94305, USA. [30]Institute of Molecular, Cell and Systems Biology, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK. [31]Immunology Discovery, Genentech, Inc., South San Francisco, CA 94080, USA. [32]MRC Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford OX3 9DS, UK. [33]Department of Biology, Syracuse University, Syracuse, NY 13244 USA. [34]Biomedical Sciences Department, Florida State University, Tallahassee, FL, USA. [35]Department of Genetics, Harvard Medical School, Boston, MA, USA. [36]Department of Developmental Biology, Stanford University School of Medicine, Stanford, CA 94305, USA. [37]Skirball Institute and HHMI, New York University Langone Medical Center, New York City, NY 10016, USA. [38]Development, Aging and Regeneration Program, Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA 92037, USA. [39]Cell Biology Program, The Hospital for Sick Children (SickKids), Toronto, ON M5G 0A4, Canada. [40]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA. [41]CABD (Andalusian Centre for Developmental Biology), CSIC-UPO-JA. Seville 41013, Spain, [42]Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720, USA. [43]Department of Neurobiology, Stanford University, Stanford, CA 94305, USA. [44]Department of Biology, New York University, New York, New York 10003, USA. [45]Neurobiology Division, MRC Laboratory of Molecular Biology, Cambridge CB2 0QH, UK. [46]European Molecular Biology Laboratory, European Bioinformatics Institute, EMBL-EBI, Wellcome Trust Genome Campus, Hinxton CB10 1SD, UK. [47]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA. [48]Department of Neuroscience, Baylor College of Medicine, Houston, TX 77030, USA. [49]Department of Physiology, Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, CA, USA. [50]Howard Hughes Medical Institute, San Francisco, CA, USA. [51]The Hebrew University of Jerusalem, Jerusalem, Israel. [52]Department of Neurobiology, Harvard Medical School, Boston, MA 02115, USA. [53]Department of Bioengineering and Biophysics Graduate Program, Stanford University, Stanford, CA 94305, USA. [54]Université Côte d'Azur, CNRS, INSERM, iBV, France. [55]Biodesign Institute, Arizona State University, Tempe, AZ 85281, USA. [56]University of California, Santa Barbara, CA 93106, USA. [57]Uppsala University, Sweden. [58]Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA. [59]Institut Jacques Monod, Centre National de la Recherche Scientifique-UMR 7592, Université Paris Diderot, Paris, France. [60]Cardiovascular Research Institute, University of California, San Francisco, CA 94143, USA. [61]Donnelly Centre for Cellular and Biomolecular Research, Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 3E1, Canada. [62]Skirball Institute, Department of Cell Biology and HHMI, New York University Langone Medical Center, New York City, NY 10016. [63]European Molecular Biology Laboratory, European Bioinformatics Institute, EMBL-EBI, Wellcome Trust Genome Campus, Hinxton CB10 1SD, UK. [64]Department of Biology, Syracuse University, Syracuse, NY, USA. [65]Department of Physiology, Development and Neuroscience, University of Cambridge, Downing Street, Cambridge CB2 3DY, UK. [66]European Bioinformatics Institute (EMBL/EBI), Wellcome Trust Genome Campus, Cambridge, UK. [67]Department of Biology, New York University, New York, NY 10003, USA. [68]Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge CB2 3DY, UK. [69]The Biological Laboratories, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138, USA. [70]Skirball Institute, Faculty of Medicine, New York University, New York, NY 10016. [71]Genomics Core, National Institute of Diabetes and Digestive and Kidney Diseases, US National Institutes of Health, Bethesda, MD, USA. [72]Graduate Program in Genetics and Genomics, Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030 USA. [73]Centre for Neural Circuits and Behaviour, University of Oxford, Tinsley Building, Mansfield Road, Oxford OX1 3TA, UK. [74]Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, TX 77030, USA. [75]Laboratory for Systems Biology of Neural Tissue Differentiation, Berlin Institute for Medical Systems Biology (BIMSB), Max Delbrueck Centre for Molecular Medicine (MDC) in the Helmholtz Association, Robert-Roessle-Strasse 12, 13125 Berlin, Germany. [76]Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada. [77]Department of Evolutionary Biology, University of Vienna, Vienna, Austria.

†Deceased.
‡Present address: Whitehead Institute and Department of Biology, MIT, Boston, MA, USA.
§Present address: University of Bergen, Faculty of Medicine, Bergen, Norway.

## SUPPLEMENTARY MATERIALS



View/request a protocol for this paper from *Bio-protocol*.

5 July 2021; accepted 19 January 2022
10.1126/science.abk2432

Science

# Fly Cell Atlas: A single-nucleus transcriptomic atlas of the adult fruit fly

Hongjie LiJasper JanssensMaxime De WaegeneerSai Saroja KolluruKristofer DavieVincent GardeuxWouter SaelensFabrice P. A. DavidMaria Brbi#Katina SpanierJure LeskovecColleen N. McLaughlinQijing XieRobert C. JonesKatja BruecknerJiwon ShimSudhir Gopal TattikotaFrank SchnorrerKatja RustTodd G. NystulZita Carvalho-SantosCarlos RibeiroSoumitra PalSharvani MahadevarajuTeresa M. PrzytyckaAaron M. AllenStephen F. GoodwinCameron W. BerryMargaret T. FullerHelen White-CooperErika L. MatunisStephen DiNardoAnthony GalenzaLucy Erin O'BrienJulian A. T. DowHeinrich JasperBrian OliverNorbert PerrimonBart DeplanckeStephen R. QuakeLiqun LuoStein Aerts, and Devika Agarwal, and Yasir Ahmed-Braimah, and Michelle Arbeitman, and Majd M. Ariss, and Jordan Augsburger, and Kumar Ayush, and Catherine C. Baker, and Torsten Banisch, and Katja Birker, and Rolf Bodmer, and Benjamin Bolival, and Susanna E. Brantley, and Julie A. Brill, and Nora C. Brown, and Norene A. Buehner, and Xiaoyu Tracy Cai, and Rita Cardoso-Figueiredo, and Fernando Casares, and Amy Chang, and Thomas R. Clandinin, and Sheela Crasta, and Claude Desplan, and Angela M. Detweiler, and Darshan B. Dhakan, and Erika Donà, and Stefanie Engert, and Swann Floc'hlay, and Nancy George, and Amanda J. González-Segarra, and Andrew K. Groves, and Samantha Gumbin, and Yanmeng Guo, and Devon E. Harris, and Yael Heifetz, and Stephen L. Holtz, and Felix Horns, and Bruno Hudry, and Ruei-Jiun Hung, and Yuh Nung Jan, and Jacob S. Jaszczak, and Gregory S. X. E. Jefferis, and Jim Karkanias, and Timothy L. Karr, and Nadja Sandra Katheder, and James Kezos, and Anna A. Kim, and Seung K. Kim, and Lutz Kockel, and Nikolaos Konstantinides, and Thomas B. Kornberg, and Henry M. Krause, and Andrew Thomas Labott, and Meghan Laturney, and Ruth Lehmann, and Sarah Leinwand, and Jiefu Li, and Joshua Shing Shun Li, and Kai Li, and Ke Li, and Liying Li, and Tun Li, and Maria Litovchenko, and Han-Hsuan Liu, and Yifang Liu, and Tzu-Chiao Lu, and Jonathan Manning, and Anjeli Mase, and Mikaela Matera-Vatnick, and Neuza Reis Matias, and Caitlin E. McDonough-Goldstein, and Aaron McGeever, and Alex D. McLachlan, and Paola Moreno-Roman, and Norma Neff, and Megan Neville, and Sang Ngo, and Tanja Nielsen, and Caitlin E. O'Brien, and David Osumi-Sutherland, and Mehmet Neset Özel, and Irene Papatheodorou, and Maja Petkovic, and Clare Pilgrim, and Angela Oliveira Pisco, and Carolina Reisenman, and Erin Nicole Sanders, and Gilberto dos Santos, and Kristin Scott, and Aparna Sherlekar, and Philip Shiu, and David Sims, and Rene V. Sit, and Maija Slaidina, and Harold E. Smith, and Gabriella Sterne, and Yu-Han Su, and Daniel Sutton, and Marco Tamayo, and Michelle Tan, and Ibrahim Tastekin, and Christoph Treiber, and David Vacek, and Georg Vogler, and Scott Waddell, and Wanpeng Wang, and Rachel I. Wilson, and Mariana F. Wolfner, and Yiu-Cheung E. Wong, and Anthony Xie, and Jun Xu, and Shinya Yamamoto, and Jia Yan, and Zepeng Yao, and Kazuki Yoda, and Ruijun Zhu, and Robert P. Zinzen

*Science*, 375 (6584), eabk2432. • DOI: 10.1126/science.abk2432

**Cell type diversity in a whole fly**

The fruit fly *Drosophila melanogaster* has served as a premier model organism for discovering fundamental and evolutionarily conserved biological mechanisms. Combining recent advances in single-cell sequencing with powerful fly genetic tools holds great promise for making further discoveries. Li *et al*. present a single-cell atlas of the entire adult fly that includes 580,000 cells and more than 250 annotated cell types. Cells from the head and body recapitulated cell types from 15 dissected tissues. In-depth analyses revealed rare cell types, cell-type-specific gene signatures, and sexual dimorphism. This atlas provides a resource for the *Drosophila* community to study genetic perturbations and diseases at single-cell resolution. —BAP