Aging Fly Cell Atlas Identifies Exhaustive Aging Features at Cellular Resolution

Authors:
Tzu-Chiao Lu[1,2#], Maria Brbić[3,13#], Ye-Jin Park[1,2,4,14], Tyler Jackson[1,2,5], Jiaye Chen[1,2,6], Sai Saroja Kolluru[7,8,9], Yanyan Qi[1,2], Nadja Sandra Katheder[10], Xiaoyu Tracy Cai[10], Seungjae Lee[11], Yen-Chung Chen[12], Niccole Auld[1,2,5], Doug Welsch[1,2], Samuel D'Souza[9], Angela Oliveira Pisco[9], Robert C. Jones[7], Jure Leskovec[13], Eric C. Lai[11], Hugo J. Bellen[2,4,14], Liqun Luo[15], Heinrich Jasper[10*], Stephen R. Quake[7,8,9*], Hongjie Li[1,2*]

Affiliations:
1 Huffington Center on Aging, Baylor College of Medicine, Houston, TX 77030, USA.
2 Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA.
3 School of Computer and Communication Sciences, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland
4 Program in Development, Disease Models & Therapeutics, Baylor College of Medicine, Houston, TX 77030, USA.
5 Program in Cancer Cell Biology, Baylor College of Medicine, Houston, TX 77030, USA.
6 Program in Quantitative & Computational Biosciences, Baylor College of Medicine, Houston, TX 77030, USA.
7 Department of Bioengineering, Stanford University, Stanford, CA, USA.
8 Department of Applied Physics, Stanford University, Stanford, CA, USA.
9 Chan Zuckerberg Biohub, San Francisco CA, USA.
10 Regenerative Medicine, Genentech, Inc., South San Francisco, CA 94080, USA.
11 Developmental Biology Program, Sloan Kettering Institute, 1275 York Ave, New York, NY 10065, USA.
12 Department of Biology, New York University, New York, NY 10013, USA.
13 Department of Computer Science, Stanford University, Stanford, CA 94305, USA.
14 Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, TX 77030, USA.
15 Howard Hughes Medical Institute, Department of Biology, Stanford University, Stanford, CA 94305, USA.

# equal contribution
* Corresponding author: jasperh@gene.com (H.J.); steve@quake-lab.org (S.R.Q.); hongjie.li@bcm.edu (H.L.)

Abstract
Aging is characterized by a decline in tissue function, but the underlying changes at cellular resolution across the organism remain unclear. Here, we present the Aging Fly Cell Atlas (AFCA), a single-nucleus transcriptomic map of the whole aging *Drosophila*. We characterize 162 distinct cell types and perform an in-depth analysis of changes in tissue cell composition, gene expression, and cell identities. We further develop aging clock models to predict the fly age and show that ribosomal gene expression is a conserved predictive factor for age. Combining all aging features, we find unique cell type-specific aging patterns. This atlas provides a valuable resource for studying fundamental principles of aging in complex organisms.

Short title: Aging Fly Cell Atlas

Key words: aging, *Drosophila*, single-nucleus RNAseq, cell atlas

Introduction

Aging is characterized by the progressive decline of tissues across the entire body. It is a major risk factor for a wide range of diseases, including cardiovascular diseases, cancers, and neurodegenerative diseases (López-Otín et al., 2013; Niccoli and Partridge, 2012). Aging phenotypes have been observed and described for centuries and a number of different aging hypotheses have been proposed (e.g., Harman, 2003). However, critical questions remain largely unaddressed in complex organisms: How does aging impact cell composition and the maintenance of specific cell types? Do different cell types age at the same rate? Can we use one cell type's transcriptome to predict age? What genes and signaling pathways drive aging in different cell types?

The fruit fly *Drosophila melanogaster* has been at the basis of many key discoveries in genetics, neurobiology, development, and aging. About 75% of human disease-associated genes have functional homologs in the fly (Baldridge et al., 2021; Bellen et al., 2010). Many of the age-related functional changes in humans are also observed in flies, including a decline in motor activity, learning and memory, cardiac function, and fertility (Grotewiel et al., 2005). Hence, a proper description of the molecular and genetic basis of the age-related decline in flies should provide an important resource for aging studies not only in flies but also in other organisms.

The recent development of single-cell RNA sequencing (scRNA-seq) technologies and the establishment of the Fly Cell Atlas (FCA) (Li et al., 2022), a single-cell transcriptomic atlas of *Drosophila* (5-day-old; 5d hereafter), have made it possible to investigate aging phenotypes across the whole organism at single-cell resolution. Here, we present the Aging Fly Cell Atlas (AFCA), a single-cell transcriptomic map describing changes that occur in most tissues over the lifetime of the fly, including male and female transcriptomic data to identify sex differences. We performed an in-depth analysis of age-related gene expression and cell composition changes across the entire fly, as well as cell type-specific and common pathways that correlate with aging. Interestingly, we observed a significant increase of fat body nuclei with age and our *in vivo* data validated the observation. We also observed an increase of apoptotic markers in old indirect flight muscles, which may underlie the age-related decrease of muscle nuclei. Furthermore, we developed aging clock models that predict the animal's age from the single-cell transcriptomic data. In addition, we found aging variances in expressed gene number, as well as cell-type identity. Our analysis revealed that different cell types are uniquely impacted by different aging features. This aging atlas provides a valuable resource for the *Drosophila* and aging community as a reference to study aging and age-related diseases. To make the data easy to access and use, we developed a website portal for data visualization and custom analyses, and made data available at the CELLxGENE portal (**fig. S1 and S2**). All resources can be accessed at www.flyagingatlas.org.

Results

Single-Nucleus Transcriptomes of the Entire Fly from Different Ages

To generate the Aging Fly Cell Atlas (AFCA), we applied the same single-nucleus RNA sequencing (snRNA-seq) pipeline that was used for the Fly Cell Atlas (FCA, 5d adults) (Li et al., 2022) and profiled the whole head and body from three additional ages (30d, 50d and 70d). These

time points were chosen to cover the trajectory of a fly's life based on the lifespan curve generated from the same cohort of flies (**Fig. 1A**), including 70d old flies, the estimated equivalent of 80-90 year old humans. Male and female flies were sequenced separately, allowing the investigation of sexual dimorphism during aging (**Fig. 1B**; see below). To achieve the most reliable analyses of aging features, we performed preprocessing of the new aging snRNA-seq data similar to the young FCA data, including index-hopping removal, normalization, doublet removal, and batch-effect correction (**fig. S1A**). Consistent with the previous scRNA-seq study of the aging fly brain (Davie et al., 2018), we found that young and old cells have similar distribution, suggesting the whole organisms largely maintains their cell-type identities during aging (**Fig. 1C**). Overall, by combining the FCA with the new aging data, we obtained more than 868,000 cells covering all 17 broad cell classes (**Fig. 1, D and E**). The detected numbers of expressed genes and UMIs were largely consistent across different ages (**fig. S3**). The most abundant cell classes were neurons, epithelial cells, muscle, fat cells, and cells from the female reproductive system. To explore comprehensive aging features across the whole fly, we next annotated those broad cell classes into detailed cell types.

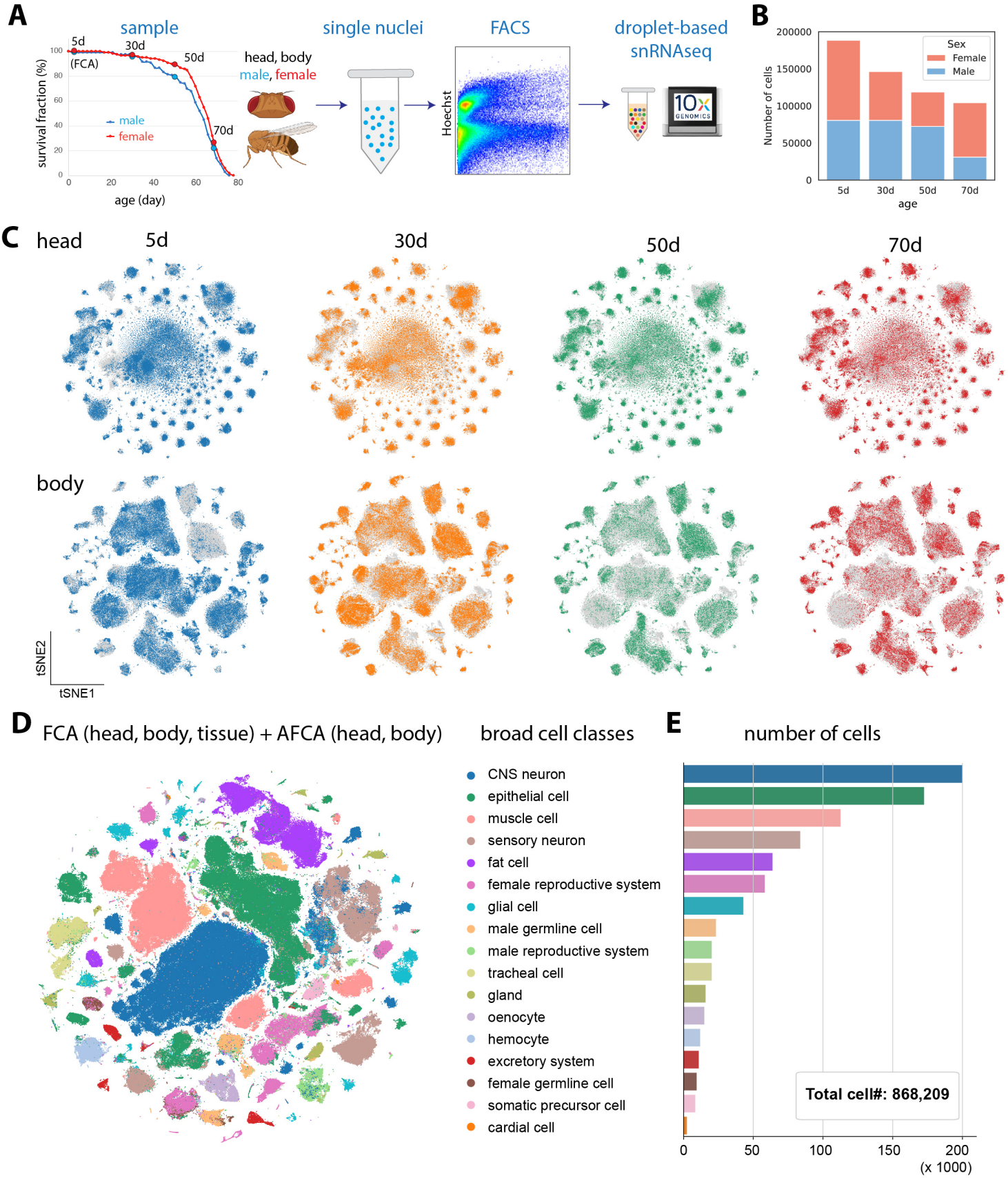
A
sample
single nuclei
FACS
droplet-based snRNAseq
5d
(FCA)
30d
50d
70d
survival fraction (%)
age (day)
male
female
head, body
male, female
Hoechst
10x GENOMICS
B
Number of cells
age
Sex
Female
Male
C
head
body
5d
30d
50d
70d
tSNE2
tSNE1
D
FCA (head, body, tissue) + AFCA (head, body)
broad cell classes
CNS neuron
epithelial cell
muscle cell
sensory neuron
fat cell
female reproductive system
glial cell
male germline cell
male reproductive system
tracheal cell
gland
oenocyte
hemocyte
excretory system
female germline cell
somatic precursor cell
cardial cell
E
number of cells
Total cell#: 868,209
(x 1000)

Figure 1. Overview of aging fly cell atlas (AFCA).
A) Flowchart of snRNA-seq experiment. Wild-type male and female flies are collected at 30, 50, and 70 days after eclosion, and head and body are separated. 5d samples are from the previous FCA data. Fly nuclei are extracted and collected using FACS for 10x Genomics based snRNA-seq.
B) Number of nuclei collected from different ages. Bars for the number of male and female cells are shown in different colors.
C) tSNE visualizations of the head and body samples from different ages. Datasets from different ages are shown in different colors.
D) tSNE visualizations showing broad cell classes of datasets integrated across different time points. Different colors denote broad cell classes.
E) Cell number of each broad cell class shown in Fig 1D. Different colors denote broad cell classes.

AFCA Cell Type Annotation and Resource for Studying Cell Type-Specific Aging

Since cell type-specific aging analysis largely depends on the accurate annotation of cell types, we took several complementary approaches to ensure that our new AFCA data are annotated with high confidence. In particular, we first co-clustered our aging data with the annotated FCA data, from either the head/body or individual tissues. Then we transferred AFCA annotations using two independent methods: a cluster-centered method and a supervised machine-learning based method (**fig. S4 and S5**, Methods). Overall, we found that these two approaches agree well with an 80% overlap of annotations between these two independent methods across the organism (**Fig. 2A**). The discrepancies of the non-overlapping annotations were mostly due to uncharacterized cell types in the FCA data or cell types with aging differences. Next, we manually validated each cell type annotation using cell type-specific markers. Marker validation confirmed the accuracy of our automatic annotation procedure, with a few exceptions such as indirect flight muscles due to age-related loss of specific markers (**fig. S6A-B**) and gut cell types due to the high similarity between intestinal stem cells and renal stem cells (**fig. S6C**). These cell types were then manually added and corrected (**fig. S5 and S7**). Overall, we characterized 162 distinct cell types, including 90 cell types from the aging head and 72 cell types from the aging body (**Fig. 2B, fig. S8 and S9**). Note that we could not reveal all 251 cell types annotated in the previous FCA study, because many FCA annotations are from dissected tissues with deep sequencing (e.g., 37 cell types from the testis with more than 43,000 nuclei sequenced and 18 types of olfactory receptor neurons from the antenna with more than 37,000 nuclei sequenced).

For complex tissues, like the brain, more cell types are resolved when significantly more cells are sequenced (Janssens et al., 2022; La Manno et al., 2021; Zeng, 2022). Indeed, 17 new neuronal cell types emerged after combining young and old head data with distinct cell markers (**Fig. 2C)**. Among them, 4 are GABAergic neurons (*Gad1*+), 2 are Glutamatergic neurons (*VGlut*+), and the remaining 11 types are cholinergic neurons (*VAChT*+) (**fig. S10**).

Next, we assessed if the AFCA provides a reliable resource to investigate age-related changes in certain tissues or cell types. We focused on the fly gut as a case study, where the somatic stem cell lineage and its aging have been well characterized in the past decade (Li and Jasper, 2016). In a healthy young fly gut, intestinal stem cells (ISCs) maintain gut homeostasis through proper proliferation and differentiation. In old guts, ISCs exhibit a high proliferation rate, and their daughter cells, enteroblasts (EBs), do not properly differentiate into mature enterocytes, leading to a dysplasia phenotype (Jasper, 2020; Micchelli and Perrimon, 2006; Ohlstein and Spradling, 2006). We first extracted six major gut cell types from the body data – ISCs, EBs, differentiating enterocytes, anterior enterocytes, posterior enterocytes, and enteroendocrine cells, and performed pseudotime trajectory analysis (Jacomy et al., 2014; Wolf et al., 2019). Interestingly, there is a significant increase of ISCs and EBs along with a decrease of fully differentiated enterocytes, consistent with previous *in vivo* studies (**Fig. 2D, fig. S11**) (Rodriguez-Fernandez et al., 2020). This trajectory analysis also allowed us to identify genes that showed different dynamic patterns along the ISC lineage between young and old flies (**fig. S11E**). In summary, the detailed annotations in our AFCA data offer a valuable resource to explore cell type- and tissue-specific aging signatures.

All the AFCA annotations and data can be accessed through our web portal www.flyagingatlas.org. We also provide a customized platform which allows users to flexibly survey different features, including cell-type, age, and sex variations (see **fig. S2**).

Cell Composition Changes During Aging

In complex organisms, aging can affect cellular composition in different ways, such as changing stem cell proliferation or differentiation processes, altering cell identity, or inducing cell death. We assessed whether and how aging impacts cellular composition across the whole fly. Note that our data are profiles of single nuclei, so our measurements are based on nuclei composition. Since the nuclei are extracted from the whole head and body with minimal sampling bias, the ratio for each cell type in our sequencing data should largely reflect their composition *in vivo*.

To perform reliable analyses, we focused on cell types that have more than 500 nuclei in total (112 cell types after filtering). We then calculated the composition changes by comparing three older ages (30d, 50d, 70d) to young flies (5d). The top 5 increased cell types are cardia cells (proventriculus from the gut), fat body cells, male accessory gland main cells, EBs, and adult salivary gland cells (**Fig. 2E**). Comparing two consecutive ages showed similar results (**fig. S12**). Age-related increases of proventriculus cells, male accessory gland main cells, and EBs from the gut have been reported previously (Biteau et al., 2008; Regan et al., 2016; Rezaei et al., 2015), confirming the quality of our data and analysis. Fat body cells are one of the most abundant cell types in *Drosophila*. They are polyploid, filled with lipid droplets, and tightly attached to the abdominal cuticle. These features make it difficult to isolate them or to compare their compositions using traditional imaging or cell-isolation and counting methods. To our knowledge, an increase of fat body nuclei in old flies has not been reported. We were able to validate this observation by dissecting young and old fly abdomen filet and manually counting fat body nuclei (**Fig. 2F**).

Fat body cells are postmitotic cells, and no adult stem cells or progenitor cells have been reported to regenerate adult fat body cells (Arrese and Soulages, 2010). To examine why fat body nuclei were increased in old flies, we first checked the number of nuclei within single cells using a fat body-specific GAL4 driving cell membrane GFP (*cg-GAL4* > *UAS-CD8GFP*). Interestingly, many aged fat body cells exhibited an increase in cell size and contained multiple nuclei per cell (**Fig. 2G**). The multinucleated phenotype can potentially be caused by cell membrane fusion as reported in other cell types (Losick et al., 2013), but cell fusion cannot explain the increase in the number of nuclei observed by snRNA-seq and in labeled tissue (**Fig. 2, E and F**). It has been reported that polyploid enterocytes from the fly gut can undergo nuclear cleavage without mitosis after starvation-induced stem cell loss (a process termed amitosis) (Lucchetta and Ohlstein, 2017). To test this possibility, we performed immunohistochemistry to detect the nuclear lamina protein, Lamin C (LamC), and a mitosis marker, Phospho-Histone H3 (pH3), in fat body cells. We did not detect any mitotic events from more than 60 flies across different ages, but we did observe many cases where two nuclei were localized very close to each other and they were negative for the mitotic marker (**Fig. 2H**). 3D reconstruction of confocal images confirmed that these nuclei were present in the same cell without a separating cell membrane (Movies S1-S3). Such events were captured across different ages (**fig. S13**). Together, these data suggest that fat body cells undergo nuclear division without cytokinesis across different ages, leading to multinucleated cells and an increase in the number of nuclei in older flies. The formation of multinucleated cells has been documented in human hepatocytes (Chen et al., 2012; Kudryavtsev et al., 1993) and it will be of interest to further investigate molecular mechanisms and physiological functions (see Discussion).

Among the top 5 decreased cell types are three types of muscles – indirect flight muscle, visceral muscle, and other muscle cells (mostly skeletal muscle) (**Fig. 2E**). Loss of muscle mass and strength or sarcopenia is a conserved aging phenotype across different mammals, including humans (Doherty, 2003). Due to the unique fiber structure and multinucleation of muscle cells, it is difficult to count total muscle nuclei from the whole animal. Our *in vivo* staining data confirmed the age-related degeneration of indirect flight muscles as well as a significant increase of an apoptosis marker Caspase3 in old flies (**Fig. 2I and fig. S14**), consistent with a previous study (Jeon et al., 2015). Germ line cells also showed a significant decrease (**Fig. 2E, fig. S12**), presumably contributing to the decline of fecundity in old flies. The majority cell types from the head, mostly neurons, showed no or minimal cellular composition changes (**Fig. 2E; fig. S12**).

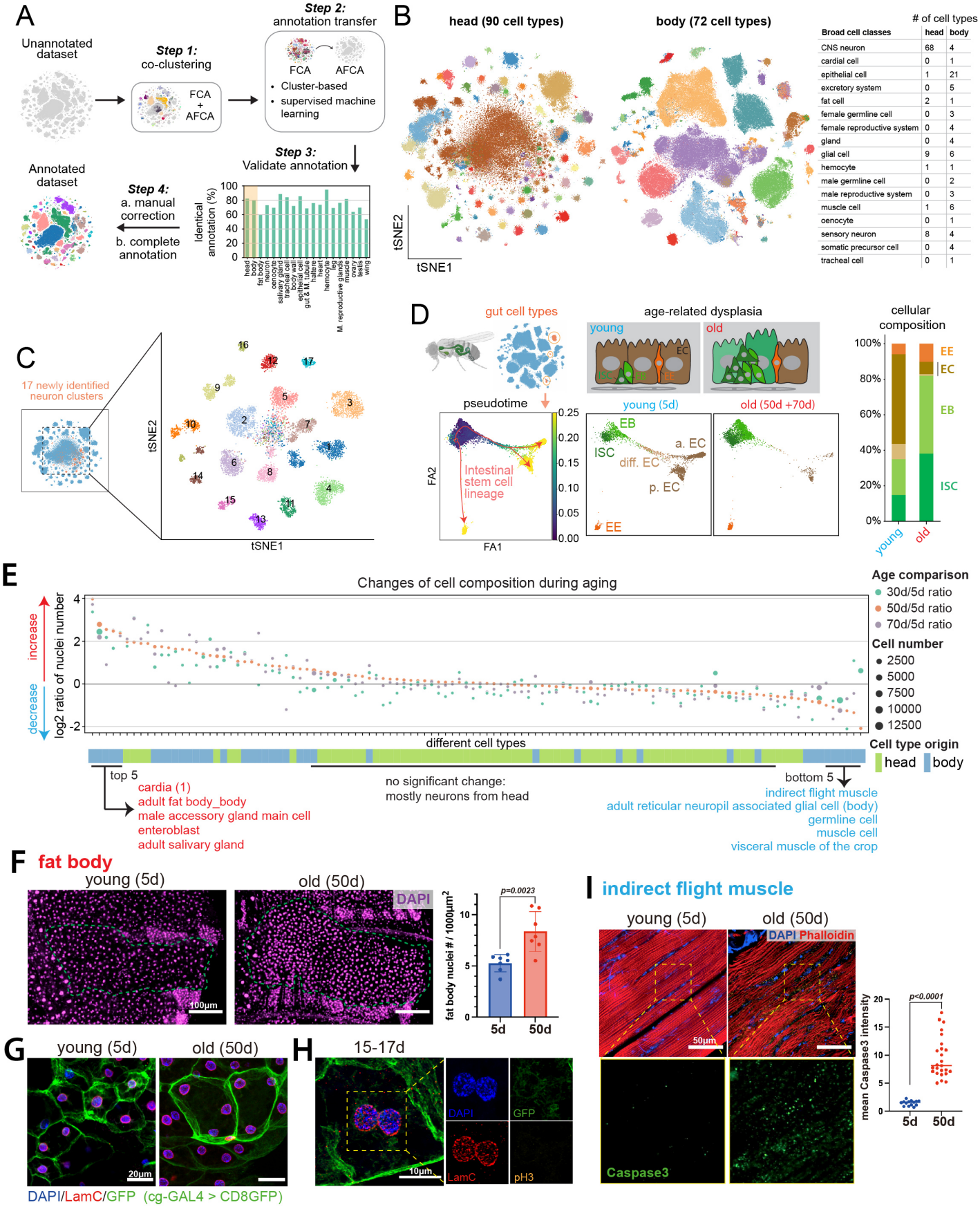
A
Unannotated dataset
Step 1: co-clustering
FCA + AFCA
Step 2: annotation transfer
FCA
AFCA
• Cluster-based
• supervised machine learning
Step 3: Validate annotation
Identical annotation (%)
Annotated dataset
Step 4: a. manual correction
b. complete annotation
B
head (90 cell types)
body (72 cell types)
tSNE2
tSNE1
# of cell types
Broad cell classes | head | body
CNS neuron | 68 | 4
cardial cell | 0 | 1
epithelial cell | 1 | 21
excretory system | 0 | 5
fat cell | 2 | 1
female germline cell | 0 | 3
female reproductive system | 0 | 4
gland | 0 | 4
glial cell | 9 | 6
hemocyte | 1 | 1
male germline cell | 0 | 2
male reproductive system | 0 | 3
muscle cell | 1 | 6
oenocyte | 0 | 1
sensory neuron | 8 | 4
somatic precursor cell | 0 | 4
tracheal cell | 0 | 1
C
17 newly identified neuron clusters
tSNE2
tSNE1
D
gut cell types
age-related dysplasia
young
old
EC
ISC
EB
EE
cellular composition
pseudotime
Intestinal stem cell lineage
FA2
FA1
young (5d)
old (50d +70d)
a. EC
diff. EC
p. EC
E
Changes of cell composition during aging
increase
decrease
log2 ratio of nuclei number
different cell types
Age comparison
30d/5d ratio
50d/5d ratio
70d/5d ratio
Cell number
2500
5000
7500
10000
12500
Cell type origin
head
body
top 5
cardia (1)
adult fat body_body
male accessory gland main cell
enteroblast
adult salivary gland
no significant change:
mostly neurons from head
bottom 5
indirect flight muscle
adult reticular neuropil associated glial cell (body)
germline cell
muscle cell
visceral muscle of the crop
F fat body
young (5d)
old (50d)
DAPI
100μm
fat body nuclei # / 1000μm²
p=0.0023
5d
50d
G
young (5d)
old (50d)
20μm
DAPI/LamC/GFP (cg-GAL4 > CD8GFP)
H
15-17d
DAPI
GFP
LamC
pH3
10μm
I indirect flight muscle
young (5d)
old (50d)
DAPI Phalloidin
50μm
Caspase3
mean Caspase3 intensity
p<0.0001
5d
50d

Figure 2. AFCA resource and changes of cell composition during aging.
A) Flowchart of transferring annotations from FCA to AFCA. See fig. S4-S7 and Methods for details.
B) Cell types annotated in the AFCA head and body shown by tSNE. Number of annotated cell types corresponding to the broad cell classes are shown in the table. See fig. S8 and S9 for detailed annotations.
C) Identification of 17 new neuronal clusters after combining AFCA and FCA head data. See fig. S10 for top marker genes.
D) Pseudotime and cellular composition of ISC and ISC-differentiated cell types. The ratio of ISCs and EBs increases with age, while the ratio of ECs decreases. ISC, intestinal stem cell; EB, enteroblast; EC, enterocyte; EE, enteroendocrine cell; a. EC, anterior EC; p. EC, posterior EC; diff. EC, differentiating EC.
E) Changes of cellular composition during aging. Each dot represents one cell type. Each color means the comparison between one aged sample and the young sample (5d). Dot sizes reflect the cell numbers of the corresponding cell type from the aged population. Tissue origins are indicated.
F) Comparison of the number of nuclei of fat body from young and old flies. Nuclei are stained by DAPI and counted in each fly. The nucleus numbers are significantly increased in the 50d population (t-test, 50d vs. 5d, P value=0.0023). Error bar, standard deviation (SD).
G) Representative confocal images showing nuclei in young and old fat body cells. The fat body cell membrane is labeled by *cg-GAL4 > UAS-CD8GFP*. Nuclei are stained by DAPI and the LamC antibody.
H) Fat body cell with segregating nuclei stained by pH3, DAPI, LamC, and GFP. pH3 signal is not detected in the segregating nuclei.
I) Indirect flight muscle stained with cleaved-Caspase3 antibody, DAPI, and Phalloidin. Cleaved-Caspase3 signals are significantly increased in the aged population (t-test, 50d vs. 5d, P value<0.0001). Median numbers indicated.

Differentially Expressed Genes (DEGs)
Altered gene expression is another consequence of aging. To assess such changes, we performed the differentially expressed gene (DEG) analysis between young and old flies (Methods) and ranked cell types by the number of DEGs between young and old flies (**Fig. 3A; fig. S15A**). Again, we focused on 112 cell types with more than 500 cells for a reliable analysis. In the body, the cell type with the highest number of DEGs was the fat body, while in the head, the most affected cell type was the outer photoreceptor, including R1 to R6 (**Fig. 3A**).

To explore the dynamics of cell type-specific changes, we further examined the time window during which cell types change the most by computing DEG numbers between two neighboring ages (5d/30d, 30d/50d, and 50d/70d; **fig. S15B**) and normalizing their ratios (**Fig 3B; fig S16**). This analysis revealed several interesting insights (**Fig 3, B and C**). Specifically, about 80% of the cell types showed major changes (> 50% of DEGs) during the first time window, suggesting that 30d old flies have captured a large portion of age-related gene changes for these cell types,

as for example observed in visceral muscle of the crop. Some cell types, like the male accessory gland, showed minimal changes in the last time window, indicating these cell types reach their maximum transcriptomic changes around 50d. However, 5.3% of types showed dramatic changes (> 50% of DEGs) in the last time window, such as intestinal stem cells and cardia cells, suggesting they age at a slower rate during the first 50d. Other cell types showed a steady rate of change in DEG, such as outer photoreceptors and fat body cells. Hence, this analysis indicates that different cell types age in different ways.

By integrating cellular composition changes and DEG analysis, we further determined which cell types were affected by those two parameters (**Fig 3D**). Significantly affected cell types ('outliers') fell into four categories: 1) cells showing changes for both, such as fat body cells from the body, pericerebral adult fat mass (fat cells from the head), and male accessory gland main cells; 2) cells showing a high number of DEGs but minimal composition changes, such as eye outer photoreceptors and cone cells, consistent with a previous study reporting that the age-related fly visual decline is not due to the loss of photoreceptors (Hall et al., 2017); 3) cells showing decreased nucleus number and a moderated number of DEGs, such as indirect flight muscles; 4) cells showing increased nucleus number but minimal DEGs, such as cardia.

Since we sequenced male and female flies separately, the dataset allowed us to explore how aging affects the sexual dimorphism across cell types. We first observed that female marker yolk protein genes (*Yp1*, *Yp2*, and *Yp3*) showed significant decrease during aging in most female cell types, while the male markers (*roX1* and *roX2*) maintained high expression levels in most male cells (**fig. S17**), suggesting aging indeed impacts male and female cells differently. It has been shown that different cell types exhibit a varied number of DEGs between males and females at a young age (Li et al., 2022). How do these numbers change during aging? To address this question, we calculated the number of DEGs between males and females and examined their dynamics along aging. Typically, if one cell type showed a high (e.g., fat body cells) or low (e.g., Johnston organ neurons) number of DEGs at 5d, it maintained the high or low number during aging (**fig. S18A**). However, some cell types showed age-specific sex differences. For example, three cell types from the head, pericerebral adult fat mass, skeletal muscle, and hemocyte, all showed low DEGs at the young age, but high DEGs in old flies (**fig. S18**). In contrast, three glial populations showed a significant decrease of DEGs with age (**fig. S18C**).

Analysis of Gene Pathways

Next, we asked which genes and pathways are enriched in those DEGs. Gene ontology (GO) analysis was performed for both up- and down-regulated genes (**fig. S19**). Most GO terms were cell type-specific (**fig. S19A**): only about 20% GO terms were shared by >5 cell types out of 112 cell types for down-regulated genes and 40% for up-regulated genes. Interestingly, we found that one GO term from down-regulated genes, called 'cytoplasmic translation', was shared by almost all cell types. Cytoplasmic translation refers to the ribosome-mediated process for protein synthesis. Indeed many transcripts of genes encoding ribosomal proteins were decreased across cell types, consistent with previous studies showing age-related reduction of protein synthesis in different tissues (see below) (Gonskikh and Polacek, 2017). There were no globally shared GOs among the up-regulated genes. Instead, they were restricted to specific groups of cells, like signal

transduction seen in neuronal cell types and protein phosphorylation enriched in different non-neuronal cells (**fig. S19**).

In addition to the shared GOs, we focused on GOs enriched within a few cell types to understand the cell type-specific regulations (**Fig. 3E and 3F**). For example, fat body cells from body and head (pericerebral adult fat mass) shared metabolic-related GO terms from up-regulated genes, such as lipid homeostasis and triglyceride homeostasis (**Fig. 3E**), suggesting a coordinated increase of these two tissues for metabolic regulation. For indirect flight muscles, a reduction of locomotor behavior was observed and is likely to be caused by the dysregulation of muscle structure and the reduction of their nucleus number (**Fig. 3E**; also see Fig. 2E). Also, sex reproduction-related genes were strongly decreased in male accessory gland main cells, suggesting a decline in reproductive ability in males (Ruhmann et al., 2018). In summary, we observed that genes involved in “cytoplasmic translation” are commonly decreased during aging in many cell types and most other GO terms showed cell type-specific patterns corresponding to the functions of these cell types.

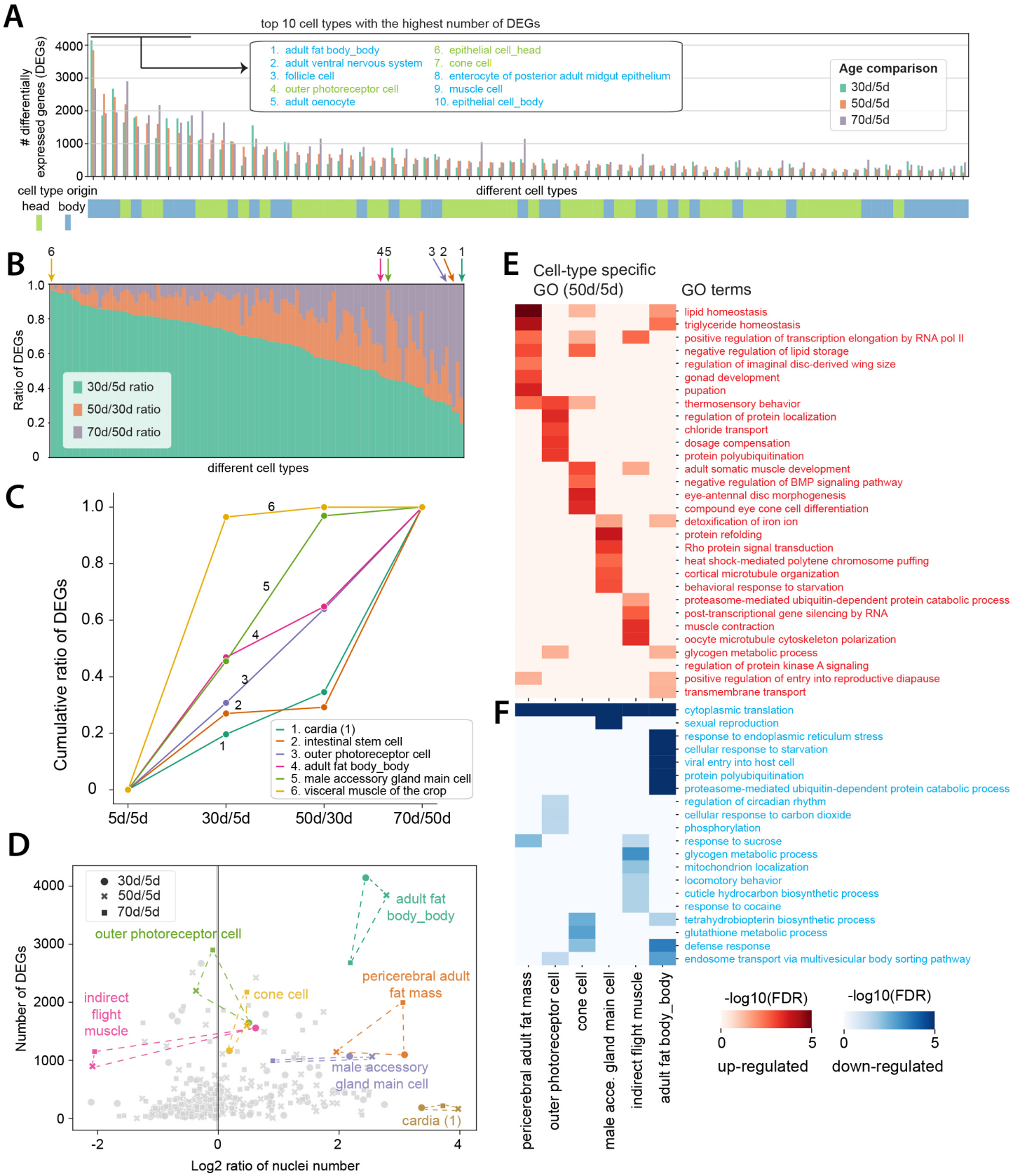
A
top 10 cell types with the highest number of DEGs
differentially expressed genes (DEGs)
1. adult fat body_body
2. adult ventral nervous system
3. follicle cell
4. outer photoreceptor cell
5. adult oenocyte
6. epithelial cell_head
7. cone cell
8. enterocyte of posterior adult midgut epithelium
9. muscle cell
10. epithelial cell_body
Age comparison
30d/5d
50d/5d
70d/5d
cell type origin
head
body
different cell types
B
Ratio of DEGs
30d/5d ratio
50d/30d ratio
70d/50d ratio
different cell types
C
Cumulative ratio of DEGs
1. cardia (1)
2. intestinal stem cell
3. outer photoreceptor cell
4. adult fat body_body
5. male accessory gland main cell
6. visceral muscle of the crop
5d/5d
30d/5d
50d/30d
70d/50d
D
30d/5d
50d/5d
70d/5d
Number of DEGs
Log2 ratio of nuclei number
adult fat body_body
outer photoreceptor cell
pericerebral adult fat mass
cone cell
indirect flight muscle
male accessory gland main cell
cardia (1)
E
Cell-type specific GO (50d/5d)
GO terms
lipid homeostasis
triglyceride homeostasis
positive regulation of transcription elongation by RNA pol II
negative regulation of lipid storage
regulation of imaginal disc-derived wing size
gonad development
pupation
thermosensory behavior
regulation of protein localization
chloride transport
dosage compensation
protein polyubiquitination
adult somatic muscle development
negative regulation of BMP signaling pathway
eye-antennal disc morphogenesis
compound eye cone cell differentiation
detoxification of iron ion
protein refolding
Rho protein signal transduction
heat shock-mediated polytene chromosome puffing
cortical microtubule organization
behavioral response to starvation
proteasome-mediated ubiquitin-dependent protein catabolic process
post-transcriptional gene silencing by RNA
muscle contraction
oocyte microtubule cytoskeleton polarization
glycogen metabolic process
regulation of protein kinase A signaling
positive regulation of entry into reproductive diapause
transmembrane transport
F
cytoplasmic translation
sexual reproduction
response to endoplasmic reticulum stress
cellular response to starvation
viral entry into host cell
protein polyubiquitination
proteasome-mediated ubiquitin-dependent protein catabolic process
regulation of circadian rhythm
cellular response to carbon dioxide
phosphorylation
response to sucrose
glycogen metabolic process
mitochondrion localization
locomotory behavior
cuticle hydrocarbon biosynthetic process
response to cocaine
tetrahydrobiopterin biosynthetic process
glutathione metabolic process
defense response
endosome transport via multivesicular body sorting pathway
pericerebral adult fat mass
outer photoreceptor cell
cone cell
male acce. gland main cell
indirect flight muscle
adult fat body_body
-log10(FDR)
up-regulated
-log10(FDR)
down-regulated

Figure 3. Differentially expressed genes (DEGs)
A) Number of DEGs from different cell types. Each aged group is compared with the young population (5d). Each line shows the number of DEGs from the indicated age comparison. Cell types are ranked by DEG numbers from high to low (50d vs 5d). Top 10 cell types are indicated.
B) Ratio of DEGs from each age comparison. Dynamic age comparisons, including 30d vs. 5d, 50d vs. 30d, and 70d vs. 50d, identify different numbers of DEGs from each age period (Methods). Arrows point out representative cell types that would be further compared in Fig. 3C.
C) Cumulative ratios of DEGs from cell types indicated in Fig. 3B.
D) Combination of DEG number and change of nucleus number illustrate different aging patterns.
E-F) Top 5 cell type-specific GOs from selected cell types. **E)** GOs enriched in the selected cell types based on up-regulated DEGs. **F)** GOs enriched in the selected cell types based on down-regulated DEGs.

Aging Clock to Predict the Biological Age
To predict the biological age of an animal or human, a number of different aging clock models have been recently developed using epigenetic markers and transcriptomic data (Horvath and Raj, 2018; Buckley et al., 2022; Davie et al., 2018). Here, we investigated whether our snRNA-seq data can be used to develop an aging clock, a cell type-specific age predictor. To perform a more accurate prediction, we focused on cell types that have more than 200 cells at each age point, with a total 64 cell types after filtering. For each cell type, we trained an Elastic Net regression model (Zou and Hastie 2005) to predict age by using its transcriptome as features (Methods). We measured predictive performance using the coefficient of determination or R^2 which represents the proportion of the variation in the age variable that was predictable from the transcriptome (**Fig 4A**). The average performance across all cell types was high (average R^2=0.79 for body and 0.84 for head). Next, we asked what types of cells and what genes can be best used to predict the fly's age.

Outer photoreceptor cells and oenocytes showed the highest predictive scores in head and body, respectively (**Fig. 4. B, C, fig. S20**). *Drosophila* oenocytes are large secretory cells present in the abdomen which perform liver-like functions including lipid storage and metabolomics functions, similar to fat body cells (Ghosh et al., 2020). We confirmed that high scores were not caused by higher cell numbers (**fig. S21A**). We further asked whether there is a difference in predictive performance between different time points. To explore this, we trained a logistic regression classifier to distinguish between transcriptomes of consecutive time points. We found that the classifier achieved near-perfect performance in distinguishing between 5d and 30d old flies in both head and body, indicating that the largest transcriptomic differences are present between the first two ages (**Fig. 4D**).

Next, we focused on identifying aging clock genes that are used to predict the animal's age (Methods). First, we found that most aging clock genes are used in a cell type-specific manner – for example, among the top 480 genes, only 8 genes are shared for more than 20 cell types (**Fig. 4E**). Interestingly, 70 out of 94 fly ribosomal protein genes were identified as aging clocking genes,

showing age-related reduction in different cell types (**Fig. 4, E to G**). This data is consistent with our previous GO analysis where 'cytoplasmic translation' is reduced in almost all cell types. Decreased protein turnover, which can be caused by the reduction of ribosomal proteins, is a prevalent feature of aging (Basisty et al., 2018). Many studies have demonstrated that the overall level of protein synthesis is reduced with age across many species, including flies, mice, rats, and humans (Anisimova et al., 2018). Our data suggest that reduction of ribosomal expression contributes to the age-related decrease of protein synthesis and turnover. To further investigate the relationship of aging clock genes across species, we identified aging clock genes for Mouse Aging Cell Atlas (Tabula Muris Senis) (Tabula Muris Consortium, 2020). Interestingly, we observed that a large portion of overlapping aging clock genes from the fly and mouse encode ribosomal proteins (**Fig. 4H; fig S22, S23**).

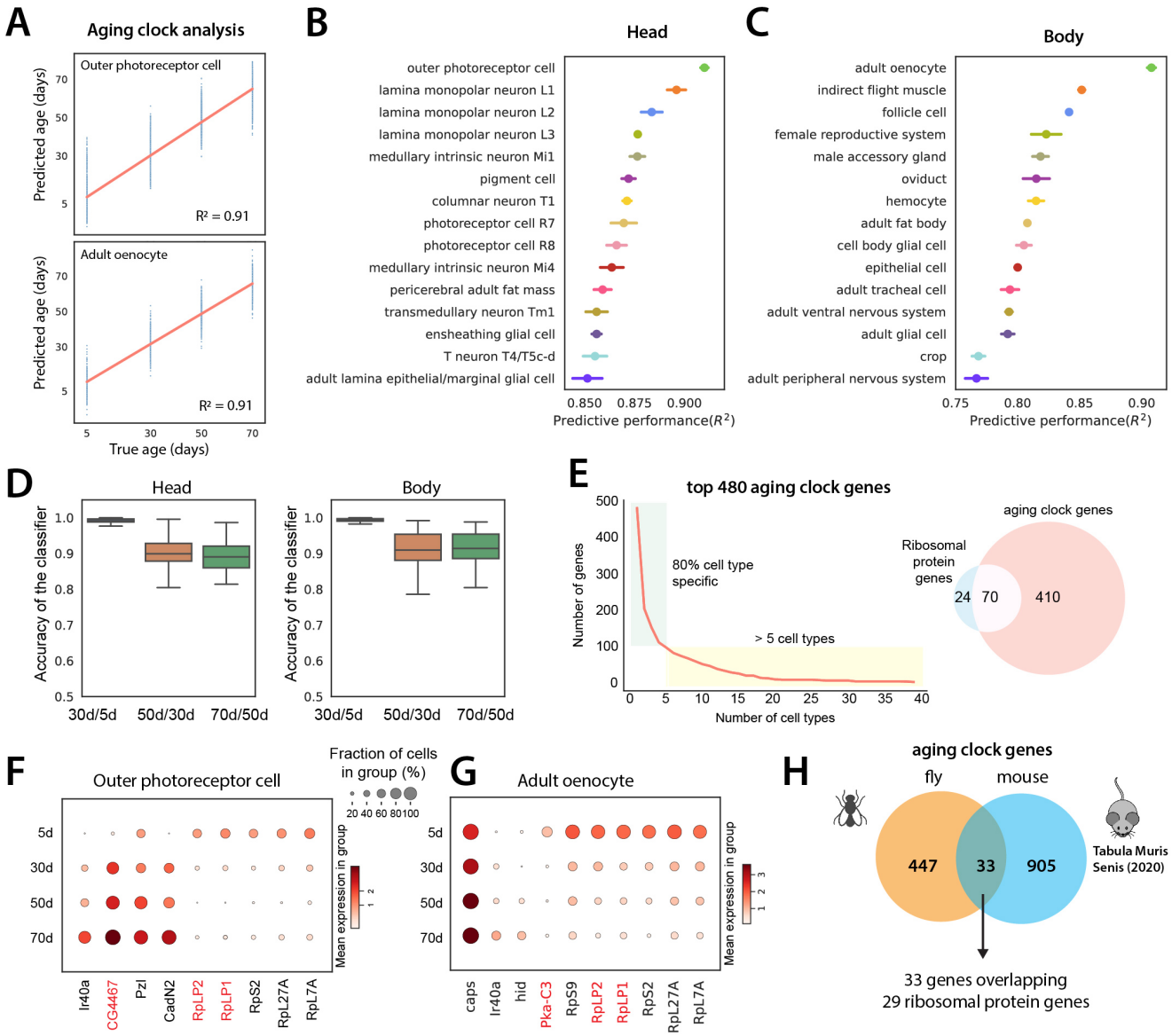
A
Aging clock analysis
Outer photoreceptor cell
Adult oenocyte
Predicted age (days)
R² = 0.91
True age (days)
B
Head
outer photoreceptor cell
lamina monopolar neuron L1
lamina monopolar neuron L2
lamina monopolar neuron L3
medullary intrinsic neuron Mi1
pigment cell
columnar neuron T1
photoreceptor cell R7
photoreceptor cell R8
medullary intrinsic neuron Mi4
pericerebral adult fat mass
transmedullary neuron Tm1
ensheathing glial cell
T neuron T4/T5c-d
adult lamina epithelial/marginal glial cell
Predictive performance(R^2)
C
Body
adult oenocyte
indirect flight muscle
follicle cell
female reproductive system
male accessory gland
oviduct
hemocyte
adult fat body
cell body glial cell
epithelial cell
adult tracheal cell
adult ventral nervous system
adult glial cell
crop
adult peripheral nervous system
Predictive performance(R^2)
D
Head
Body
Accuracy of the classifier
30d/5d
50d/30d
70d/50d
E
top 480 aging clock genes
Number of genes
80% cell type specific
> 5 cell types
Number of cell types
aging clock genes
Ribosomal protein genes
24
70
410
F
Outer photoreceptor cell
Fraction of cells in group (%)
Mean expression in group
5d
30d
50d
70d
Ir40a
CG4467
Pzl
CadN2
RpLP2
RpLP1
RpS2
RpL27A
RpL7A
G
Adult oenocyte
caps
Ir40a
hid
Pka-C3
RpS9
RpLP2
RpLP1
RpS2
RpL27A
RpL7A
H
aging clock genes
fly
mouse
447
33
905
Tabula Muris Senis (2020)
33 genes overlapping
29 ribosomal protein genes

Figure 4. Aging clock analysis.

A) Example of aging clocks for outer photoreceptor cells (top) and adult oenocytes (bottom). Aging clocks are cell type-specific regression models that predict age from transcriptome data. Red line is the fitted regression line. Blue dots represent individual predictions where each dot corresponds to one cell. We measure performance as the proportion of the variance for an age variable that's explained by transcriptome (R^2).

B, C) Predictive performance of cell type-specific aging clocks for B) head cell types and C) body cell types. Performance is measured as the mean R^2. 15 cell types with the highest scores are shown. Error bars are estimated as a standard deviation over 5 runs.

D) Accuracy of logistic regression models trained to distinguish transcriptome between two consecutive time points. Boxplots show distribution across head cell types (left) and body cell types (right).

E) Number of aging clock genes as a function of the number of cell types (left). 80% of genes appear in less than 5 cell types. Out of 480 genes identified as aging clock genes, 70 encode ribosomal proteins (right).

F-G) Examples of aging clock genes for outer photoreceptor cell F) and adult oenocyte G). Cell type-specific genes that appear in less than 5 cell types are shown in black color, while genes that appear in at least 5 cell types are marked in red color.

H) Venn diagram comparing aging clock genes in fly and mouse. Aging clock genes in mice are derived from Tabula Muris Senis dataset. 33 genes are 1-1 orthologs between the two species, and 29 of these genes encode ribosomal proteins. See fig. S22 for gene names.

Comprehensive Aging Features

Through the above analyses, we noticed that different cell types are sensitive to different aging measurements. For example, the outer photoreceptors showed a high number of DEGs, but a minimal change of cell number; germline cells showed decreased cell numbers, but a very low number of DEGs. To gain a better understanding of cell type-specific aging, we investigated more aging features, including expressed gene number or transcript number (measured by UMI) changes, and decline of cell identity (**Fig. 5A**).

Similar to the aging clock analysis, we only focused on cell types with at least 200 cells from each age for a reliable analysis. Gene and UMI numbers were previously found to decrease in the fly aging brain (Davie et al., 2018). Indeed, we observed that the CNS neurons decrease the gene or UMI numbers during aging (**fig. S24A**). To understand whether such a reduction is a general aging feature across different cell types, gene and UMI numbers from all cell types were compared between different ages (**Fig. 5B, fig. S24B**). The overall trend of gene and UMI numbers was largely consistent. Focusing on expressed gene numbers, we found that around 80% of cell types decreased expressed gene numbers per cell, but 20% increased the expressed gene number per cell, suggesting that aging affects expressed gene numbers in a cell type-specific manner. Head hemocytes and pericerebral adult fat mass increased most in expressed gene numbers, while oenocytes, ventral nervous system, and fat body decreased most (**Fig. 5B, fig S24C**). The variation of expressed gene number per cell was less likely caused by the differences in

sequencing depth since the sequencing saturation is over 90% for most of our samples (**fig. S24D**). Even though head and body fat body cells both increased the nucleus number (**Fig. 3D**), their expressed gene numbers were changed in the opposite trends during aging, supporting the idea that different aging features are differentially regulated.

Loss of cell-type identity has been shown to occur during aging for certain cell types (Izgi et al., 2022; Li et al., 2016). However, how cell identity is changed across the entire organism during aging remains uncharacterized. To characterize the cell identity changes, we developed a measurement to combine two ratios – loss of original markers and gain of new markers – by comparing old to young populations (**Fig. 5C**, left panel) (Methods). We measured these two factors and ranked cell types by their cell identity decline score (**Fig. 5C**, right panel). The gene, *Nlg1*, which was used as a marker gene to annotate the indirect flight muscle, showed a dramatic decrease with age, similar to a number of other young marker genes (**Fig. 5C, fig. S6A**). Meanwhile, many other genes began to appear in this cell type with aging (**fig. S6A**). Other than the indirect flight muscle, pericerebral adult fat mass, epithelial cell from the head, male accessory gland main cell, and outer photoreceptor cell were also found to largely decrease cell-type identity (**fig. S25-S26**).

Correlation and Ranks of Aging Features

Together with the previously analyzed three aging features – cell composition (**Fig. 2E**), DEG analysis (**Fig. 3**), and aging clock scores (**Fig. 4**) – we examined five different aging features in total. To understand the overall correlation of different aging features, we ranked each feature from the least age-related change to the most (Methods). After integrating five aging features, their correlations were compared using Spearman's correlation and clustered using the correlation scores (**Fig. 5F**). Among the five features, DEG number and cell identity were highly correlated with each other, suggesting cell types with high number of DEGs would usually fluctuate in the expression of marker genes. On the other hand, change of cell number, and expressed gene number were more correlated with each other, while the aging clock score seemed to be more independent of other features.

To categorize which cell types have more aging features, we summed different feature ranks and sorted cell types by the total rank sums, higher rank meaning more "aged" (**Fig. 5G, fig. S27**). Interestingly, the top four cell types include three adipose cell types, oenocytes, fat body cells, and pericerebral adult fat mass (fat cells from the head), as well as indirect flight muscle cells, suggesting that those cells age faster than other cell types. Following them are two reproduction-related cell types, male accessory gland main cells and follicle cells from the ovary, indicating that reproductive systems also undergo fast aging. Generally, neurons and glia from the nervous system age slower compared to other cell types (**Fig. 5G**). In summary, our analysis provided the first exhaustive analysis of different aging features and revealed the aging rates of different cell types across the entire fly.

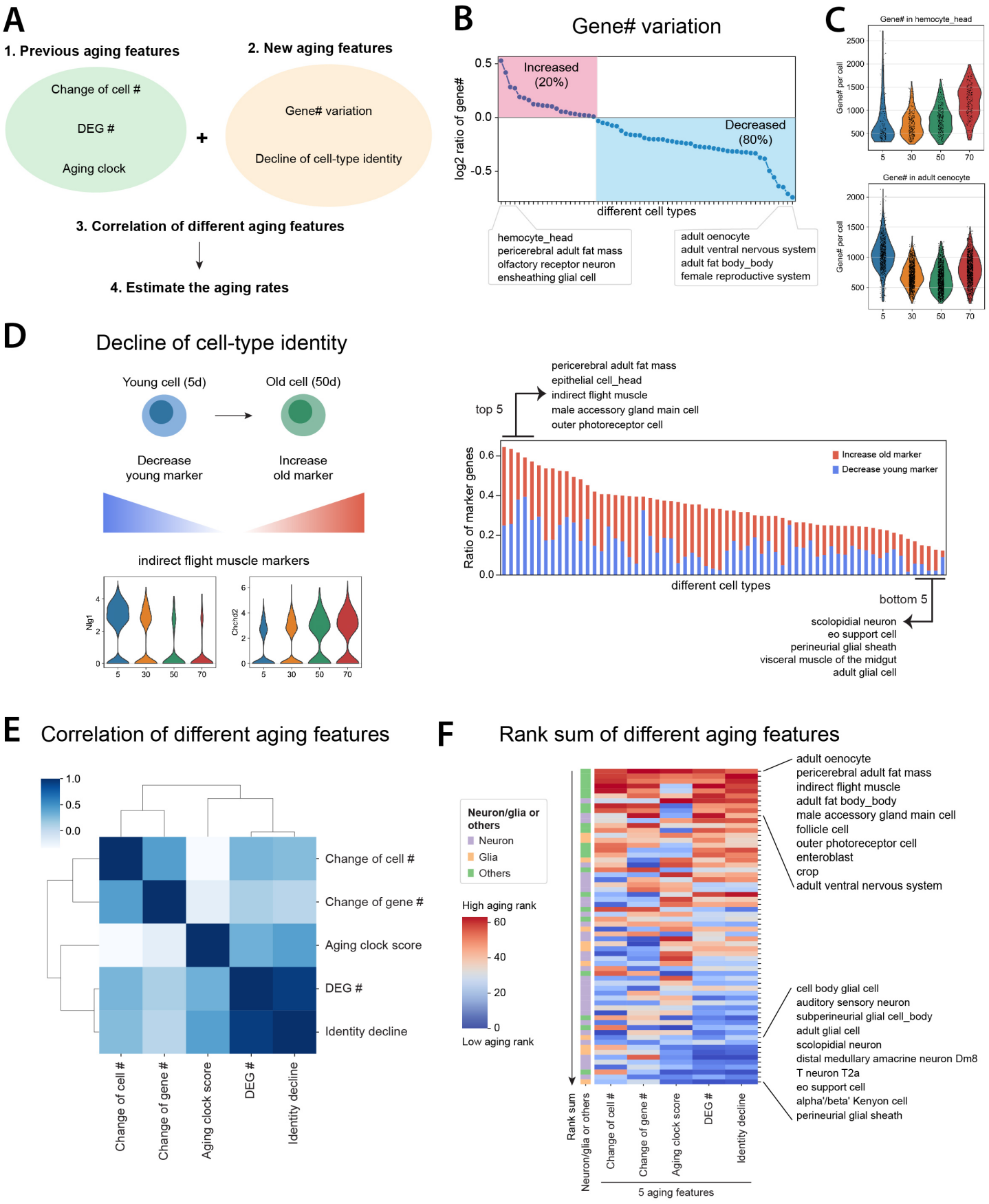
A
1. Previous aging features
Change of cell #
DEG #
Aging clock
+
2. New aging features
Gene# variation
Decline of cell-type identity
3. Correlation of different aging features
4. Estimate the aging rates
B
Gene# variation
Increased (20%)
Decreased (80%)
log2 ratio of gene#
different cell types
hemocyte_head
pericerebral adult fat mass
olfactory receptor neuron
ensheathing glial cell
adult oenocyte
adult ventral nervous system
adult fat body_body
female reproductive system
C
Gene# in hemocyte_head
Gene# in adult oenocyte
Gene# per cell
D
Decline of cell-type identity
Young cell (5d)
Old cell (50d)
Decrease young marker
Increase old marker
indirect flight muscle markers
Nlg1
Chchd2
top 5
pericerebral adult fat mass
epithelial cell_head
indirect flight muscle
male accessory gland main cell
outer photoreceptor cell
Ratio of marker genes
Increase old marker
Decrease young marker
different cell types
bottom 5
scolopidial neuron
eo support cell
perineurial glial sheath
visceral muscle of the midgut
adult glial cell
E
Correlation of different aging features
Change of cell #
Change of gene #
Aging clock score
DEG #
Identity decline
F
Rank sum of different aging features
Neuron/glia or others
Neuron
Glia
Others
High aging rank
Low aging rank
adult oenocyte
pericerebral adult fat mass
indirect flight muscle
adult fat body_body
male accessory gland main cell
follicle cell
outer photoreceptor cell
enteroblast
crop
adult ventral nervous system
cell body glial cell
auditory sensory neuron
subperineurial glial cell_body
adult glial cell
scolopidial neuron
distal medullary amacrine neuron Dm8
T neuron T2a
eo support cell
alpha'/beta' Kenyon cell
perineurial glial sheath
Rank sum
Neuron/glia or others
5 aging features

Figure 5. Systematic comparisons of different aging features.
A) Flowchart of comparing different aging features.
B) Age variation of detected gene numbers per cell. Expressed gene numbers per cell from each cell type are compared between old (50d) and young (5d) flies. The red block shows cell types with increased gene number during aging, while the blue block includes cell types with decreased one.
C) Two cell types, hemocyte from the head and oenocyte, have the highest increase and decrease of expressed gene numbers per cell.
D) Decline of cell identity during aging. The left panel illustrates two different mechanisms of decreasing cell identity. The right panel shows the ratio of marker genes decreasing cell identity. Each line in the right panel represents one cell type.
E) Spearman's correlation of different aging features. The correlations of five different aging features are compared in 64 cell types.
F) Rank sum of different aging features. The heatmap shows the overall rank sum scores from different cell types. High aging ranks are shown in red, while low aging ranks are shown in blue. Neuron- or glia-related classification are shown beside the heatmap.

Discussion

In complex organisms, different cell types possess distinct characteristics of their corresponding functions. For example, somatic stem cells need to proliferate and differentiate to replenish the tissue, while postmitotic neurons remain alive and functional during the lifespan of the organism. Hence, aging causes a physical and functional decline of all cell types in very different manners. Understanding how different cell types are impacted by aging is critical for designing interventions that delay aging and improve health. Molecular genetic approaches in *Drosophila* have helped elucidate fundamental mechanisms and principles in genetics, development, and neurobiology. The AFCA provides a resource for the aging community to investigate aging and age-related diseases at a single-cell resolution.

The cell atlas approach is emerging as a powerful tool to systematically study aging in different organisms, including *C. elegans* (Roux et al., 2022), mouse (Tabula Muris Consortium, 2020), and humans (Sziraki et al., 2022). A recent study performed cross-species analysis with single-cell RNAseq data from three different species, including *Drosophila* from 11d to 41d (Wang et al., 2022). However, this study uses a single-cell RNAseq protocol, and thus, many cell types which cannot be efficiently isolated as intact cells, are not revealed (33 annotated cell types versus 162 in our study) (**fig. S28**). For example, they annotated 1 glial type (15 types in AFCA), 1 type of sensory neuron (12 types in AFCA), and 5 types of CNS neurons (72 types in AFCA) (**fig. S28A**). As a consequence, easily-isolated cell types are overrepresented (e.g., 38.9% cells from the dataset are follicle cells from the ovary) (**fig. S28B**). Also, male and female flies are not separated in this study. Our study employed an unbiased single-nucleus RNAseq approach, allowing us to survey cell composition changes more faithfully. We sequenced male and female flies separately, allowing to study sexual dimorphism along aging (**fig. S17 and S18**). The AFCA atlas contains

samples from older populations (50d and 70d) covering the whole lifespan. Overall, our AFCA provides a complementary but more thorough and complete dataset for studying aging features across the whole organism.

One interesting observation is the significant increase of fat body nuclei during aging. *Drosophila* fat body is a liver-like tissue that stores fat and serves as a detoxifying and immune responsive organ. Adult fat body cells are postmitotic polyploid cells without a stem cell or progenitor population (Arrese and Soulages, 2010). How do they increase their nucleus number? Our observations suggest that these polyploid cells increase their number of nuclei by nuclear cleavage without cytokinesis, forming multinucleated cells. To complete karyokinesis, the nuclear envelope needs to be reassembled or reorganized. Interestingly, fat body cells have been shown to decrease their nuclear envelope integrity due to the loss of Lamin-B during aging (Chen et al., 2014). It will be of interest to study whether the loss of nuclear envelope integrity is associated with the multinucleated phenomenon. On the other hand, multinucleated cells have also been observed in the fly male accessory gland (Bertram et al., 1992, Box et al., 2022) and subperineurial glia cells (Unhavaithaya and Orr-Weaver, 2012), suggesting multinucleated cells play specific roles in *Drosophila*. Other than *Drosophila*, multinucleated cells have also been observed in other species, including mushrooms (Gehrmann et al., 2018), plants (von Aderkas et al., 2005), and human and mouse liver cells (Chen et al., 2012). Understanding the formation and regulation of multinucleated cells in the aging organism may provide new insights into an evolutionarily conserved phenomenon, as well as into the potential roles of multinucleated cells in age-related diseases.

We focused on five different aging features in this study. Although these features cover several key aspects of age-related changes, the picture is not complete. Additional aging measurements may reveal more specific aging patterns. For example, we investigated the change of alternative polyadenylation (APA) patterns, which can reflect short or long 3' UTR usage for different isoforms (Lee et al., 2022), and found that neuronal extended 3' isoforms were progressively depleted during aging. This phenotype is more obvious at 70d and more pronounced in females than in males (**fig. S29**). These data imply a global change in post-transcriptional regulation in aging neurons. Thus, future specific analysis may help elucidate additional aging patterns.

One major goal of this study is to characterize how different cell types age across the organism. Our analysis using five different aging features provides several interesting insights. First, different cell types have very distinct aging patterns. For example, the fat body cells showed high ranks for 4 aging features but low for aging clock score, while scolopidia neurons showed low ranks for 4 aging features but high changes in expressed gene number (**Fig. 5F**). This observation is not unexpected considering that each cell type carries a unique function. Secondly, we observed a divergence in the contribution of individual cell types to a tissue's aging phenotype. For example, in the female reproductive system, follicle cells were ranked very high (6th of 64 cell types), but germline cells were ranked near the bottom (53th of 64 cell types) (**fig. S27A**). This indicates that age-related decline of female fertility may be due to aging of follicle cells. Thirdly, the top ranked cell types include all adipose cells. This is surprising and we do not fully understand the underlying mechanisms. It may be linked to the fact that these cell types play multiple critical roles in different

physiological conditions, such as lipid storage and metabolism, immune responses, and interorgan communication with muscles and gut (Arrese and Soulages, 2010; Chatterjee and Perrimon, 2021; Zhao and Karpac, 2021).

Acknowledgments:
We thank Stein Aerts and Jasper Janssens for helping preprocess the data and their constructive advice on aging data annotation, Norbert Perrimon for his comments and suggestions during the initiation of this project, Hugo Bellen lab members for feedback and suggestions, Laura Buttitta for comments and discussion on polyploid phenotypes, Andrey Parkhitko for discussion on metabolic pathways and aging. Y.C.C. is grateful to Claude Desplan for support. We thank Chung-Yi Liang, Chuangye Qi, Madeline Burns, and all other lab members in Hongjie Li lab for comments and feedback along the project. We also thank FCA Consortium and the fly community for their enthusiastic support.

Funding:
M.B. acknowledges the EPFL support. T.J. was supported by the Baylor College of Medicine Cancer and Cell Biology program grant T32 (GM136560) during the preparation of the manuscript. H.J.B. is supported by NIH/NIA, R01 AG07326, the Huffington foundation, and the endowment of the Chair of the Neurological Research Institute. L.L. is an investigator of the Howard Hughes Medical Institute and supported by NIH (R01-DC005982). S.L. was supported by a training award from the NYSTEM contract #C32559GG and the Center for Stem Cell Biology at MSKCC. E.C.L. was supported by NIH grants R01-GM083300 and R01-HL135564, and by the MSK Core Grant P30-CA008748. Y.C.C. was supported by the MacCracken Program at New York University, by a NYSTEM institutional training grant (Contract #C322560GG), and by Scholarship to Study Abroad from the Ministry of Education, Taiwan. H.L. is a CPRIT Scholar in Cancer Research (RR200063) and supported by National Institutes of Health (R00AG062746) and BCM Junior Faculty Award. This work is supported by Longevity Impetus Grant (H.L.).

Author contributions:
Conceptualization: T.C.L., L.L., H.J., S.R.Q., H.L.
Sample preparation: N.K., X.T.C., H.L.
Sequencing: S.S.K., Y.Q., R.C.J., H.L.
Computational analysis: T.C.L., M.B., J.L., T.J., D.W.
Data portal and AFCA website: J.C., T.C.L., S.D, A.P.
APA analysis: S.L., Y.C.C., E.C.L.
Data validation: Y.J.P., H.J.B., T.J., Y.Q., N.A.
Writing: T.C.L., M.B., H.L.
Review and Editing: All authors
Supervision: H.J., S.R.Q., H.L.
Funding Acquisition: H.J., S.R.Q., H.L.

Competing interests: H.J., N.K., and X.T.C. are employees of Genentech, Inc. The other authors declare no competing interests.

Data and materials availability:

All data are available for querying at https://flyagingatlas.org. Raw FASTQ files and processed h5ad files, including expression matrices and cell-type annotations, can be downloaded from NCBI GEO: GSE218661. Analysis codes for generating the figures can be downloaded from https://github.com/hongjie-lab/AFCA. FCA data were downloaded from https://www.flycellatlas.org/.

Materials and Methods

Fly sample and lifespan

Fly head and body samples were from F1 of two wild type fly stains: female w^{1118} and male Oregon R (*OreR*). Flies were maintained on standard cornmeal medium at 25℃ with 12-hour light–dark cycle. For the lifespan experiment, progenies were collected 3 days after the first fly hatched. Flies were allowed to mate for 3 days. Then male and female flies were separated into different cages (120 flies per cage) for lifespan or sample collection at 25°C. Fly food was changed every 2 days. Dead flies were counted every 2 days. Lifespan curves were generated in Excel.

Single-nucleus RNA sequencing

Fly heads and bodies were dissected at different ages, 30d, 50d and 70d, put into 1.5ml RNAase free Eppendorf tube, flash-frozen using liquid nitrogen, and then stored at –80℃. We put 100 heads or 50 bodies in each tube. Single-nucleus suspensions were prepared following the protocol described in Fly Cell Atlas study (Li et al., 2022). Next, we collect nuclei by FACS. We used the SONY SH800 FACS sorter for collecting nuclei. Nuclei were stained by Hoechst-33342 (1:1000; >5min). Since polyploidy is common for many fly tissues, we observed different populations of nuclei according to DNA content (Hoechst signal). In order to include all cell populations with different nuclear sizes, we have included all nuclear populations from the FACS. Nuclei were collected into a 1.5ml tube with 200ul 1x PBS with 0.5% BSA as the receiving buffer (RNase inhibitor added). For each 10x Genomics run, 100k–400k nuclei were collected. Nuclei were spun down for 10 min at 1000g at 4℃, and then resuspended using 40ul or desired amount of 1x PBS with 0.5% BSA (RNase inhibitor added). 2ul nucleus suspension was used for counting the nuclei with hemocytometers to calculate the concentration. When loading to the 10x controller, we always target 10k nuclei for each channel. We observed that loading 1.5 folds more nuclei as recommended by the protocol allowed us to recover about 10k cells after sequencing. For example, if the concentration is 1500 nuclei per ul for one sample, we treat it as 1000 nuclei per ul when loading to the 10x controller.

Next, we performed snRNA-seq using the 10x Genomics system with 3' v3.1 kits with following settings. All PCR reactions were performed using the Biorad C1000 Touch Thermal cycler with 96-Deep Well Reaction Module. 13 cycles were used for cDNA amplification and 16 cycles were used for sample index PCR. As per 10x protocol, 1:10 dilutions of amplified cDNA and final libraries were evaluated on a bioanalyzer. Each library was diluted to 4 nM, and equal volumes of 18 libraries were pooled for each NovaSeq S4 sequencing run. Pools were sequenced using 100 cycle run kits and the Single Index configuration. Read 1, Index 1 (i7), and Read 2 are 28 bp, 8 bp and 91 bp respectively. A PhiX control library was spiked in at 0.2 to 1% concentration. Libraries were sequenced on the NovaSeq 6000 Sequencing System (Illumina).

For each age, we sequenced the head for 12 runs (6 for male and 6 for female) and body for 12 runs (6 for male and 6 for female). To collect enough nuclei for 6 runs, we used 100 heads and 50 bodies. For three ages, we sequenced 24*3 = 72 runs. We have recently published a protocol describing the single-nucleus RNAseq in *Drosophila* and all above steps have been detailed (McLaughlin et al., 2022).

Sequencing read alignment

The raw FASTQ files were first filtered for the index-hopping reads by using the index-hopping-filter developed by the 10x Genomics (version 1.0.1). Before mapping reads, the *Drosophila melanogaster* genome (FlyBase r6.31) was indexed with a pre-mRNA GTF, which was established by FCA (Li et al., 2022), using the Cell Ranger (version 4.0.0). FASTQ files removing the index-hopping reads were mapped to the *Drosophila melanogaster* genome (r6.31) using Cell Ranger Count (version 4.0.0) to generate the count matrix.

AFCA annotation

In the FCA dataset, there are 251 cell types annotated, many of which were characterized from dissected tissues. For cell types from the head, besides the whole head sequencing, there are only two dissected tissues: antennae and proboscis. So, most cell types from the aging head can be annotated based on the FCA whole head data. For cell types from the body, most cell type annotations are from tissues (only 33 annotated cell types from the FCA whole body data). To annotate aging body data, we need to transfer annotations from the body and individual tissues (e.g., gut, wing, leg, testis, ovary, fat body etc). Thus, the aging head and body data were annotated using slightly different strategies and we characterized them separately. For both, we first integrated FCA and AFCA datasets and corrected age and sex differences using Harmony (Korsunsky et al., 2019). Next, we characterized cell types using two different approaches: co-clustering and supervised machine learning (ML)-based classification. Then, we validated those two approaches using cell type-specific markers and performed manual correction when necessary.

For the head data, FCA and AFCA head cells were co-clustered and 150 principal components (PCs) were selected for clustering. We followed the co-cluster-based approaches utilized in FCA brain-head integration (Li et al., 2022) and annotations from the FCA cells were transferred to the clusters that contained at least 15% of cells with the same annotation. Alternatively, if more than 15% of the annotation-specific cells were enriched in the clusters, those clusters were assigned to the corresponding annotations. Different Leiden resolutions were applied for the transfer and the one with the maximum number of transferred annotations was chosen. If annotations from two approaches were not consistent, cluster-dominant annotations would be assigned to the clusters. Besides cluster-based transfer, we also included a supervised ML method. We utilized the Logistic Regression classifier to transfer the FCA annotations to AFCA cells. FCA cells were used as the training dataset, while AFCA samples were considered as test data. By comparing to the annotations from these two approaches, annotations were manually confirmed by examining the expression of cell type-specific genes to prevent the loss of transfer due to aging. Also, annotations were unified and denoised in the cluster level. In total, 90 cell types were annotated, including 17 newly identified neuron types.

To annotate AFCA body cells, FCA and AFCA body cells were first co-clustered using 50 PCs and 33 annotations from FCA body were transferred to the AFCA body cells using approaches

similar to the head transfer. To include annotations from different body tissues, we co-clustered cells from FCA body, AFCA body, and all FCA tissues to identify clusters with tissue-specific annotations (**fig. S4**). The cluster-based method enabled us to quickly identify tissue-specific annotations from integrated body and tissue cells. Clusters with tissue-specific annotations were chosen for further sub-clustering and tissue annotations were transferred similarly using cluster- and supervised ML-based approaches. Due to the similarity between some gut and malpighian tubule cell types, especially for the intestinal stem cell and renal stem cell (**fig. S6C and S7**), we combined annotations from these two tissues together for the transfer. Some cell types, like muscle cells, were present in multiple tissues and we considered subsetting those cells based on the expression of cell-type markers, instead of selecting the annotation-enriched clusters (**fig. S5**). After subsetting marker-positive cells, FCA annotations were similarly transferred. After transferring all tissue annotations, 72 cell types were annotated in the AFCA body.

Gut cell type case study

The analysis of the gut cells in the AFCA data was performed using SCANPY and scFates python packages (Faure et al., 2022; Wolf et al., 2018). The intestinal stem cell lineage was isolated from the data and includes intestinal stem cells, enteroblasts, adult differentiating enterocytes, enteroendocrine cells, anterior enterocytes, and posterior enterocytes. The pseudotime was generated by combining the young (FCA age 5 sample) and old (AFCA age 50 and 70 samples). To generate this plot, the partition-based graph abstraction (PAGA) function was applied to infer the connectivity of clusters then the ForceAtlas2 algorithm was implemented to spatially overlay the cells onto the connectivity plot generated by PAGA (Jacomy et al., 2014; Wolf et al., 2019). To compare the cell compositions of each age, the cells of different ages were plotted separately next to the pseudotime. The percentages of cells comprising the intestinal stem cell lineage were extracted from the data and then plotted on Fig 2D.

The changes in gene expression were generated using the scFates python package (Faure et al., 2022). Cells were isolated from each age and then run separately in the scFates pipeline. Briefly, scFates uses the ElPiGraph algorithm to learn the topography of the data. Then, the pseudotime was constructed from the graph inferred by the ElPiGraph algorithm. The genes differentially expressed along the trajectory were determined using a cubic spline regression model and then these results were compared to an unrestrained model. The Benjamini-Hochberg correction was then applied to adjust for multiple comparisons. A significantly altered gene along the trajectory between each age, Rbfox1, is shown in Fig S11.

Fat body and indirect flight muscle staining

To label the fat body tissue with GFP, cg-GAL4 (Bloomington Drosophila Stock Center (BDSC), #7011) virgins were crossed with UAS-unc84GFP (from Dr. Gilbert L. Henry) or UAS-CD8GFP (from Dr. Liqun Luo) males to obtain cg-GAL4>unc84GFP or cg-GAL4>CD8GFP flies. For indirect flight muscle labeling, Act88F-GAL4 (BDSC #38461) virgins were crossed with UAS-unc84GFP to obtain Act88F-GAL4>unc84GFP flies. Flies were cultured at 25°C to desired ages.

For fat body dissection, fly abdomen filet were prepared in 1X PBS using Vannas Spring Scissors (FST, 3mm Cutting Edge, 15000-00) and Dumont #55 forceps (FST, 11255-20). For indirect flight muscle dissection, flies in 1.5ml Eppendorf tube were snap-freezed in liquid nitrogen, placed on the ice rack. After the head, leg and wing were removed, the abdomen was clamped with forceps and thorax was cut in the sagittal section by a razor. Tissues were fixed with 4% paraformaldehyde for 20 min at room temperature (RT), washed with 1X PBS, blocked in 5% NGS blocking buffer (1XPBS, 0.3% Triton X-100, 5% NGS) for 2 h at RT, incubated with primary antibodies (in 5% NGS blocking buffer) overnight at 4°C, and then washed 5 times with 0.3% PBST (1XPBS, 0.3% Triton X-100) before incubation with secondary antibodies for 2 h at RT. Tissues were thoroughly rinsed in PBST, stained with DAPI(1:1000) for 15 min, washed with 1XPBS and mounted with SlowFade™ Gold Antifade Mountant (Thermo Fisher, S36936).

Mouse anti-LamC (1:100, Developmental Studies Hybridoma Bank (DSHB), LC28.26), chicken anti-GFP (1:1000, Aves Labs), rabbit anti-pH3 (1:1000, Cell Signaling Technology, #9701), rabbit anti-cleaved-Caspase3 (1:100, Cell Signaling Technology, #9661) were used as primary antibodies. For secondary antibodies, we used Cy™3 AffiniPure Donkey Anti-Mouse IgG (H+L) (1:250, Jackson ImmunoResearch, AB_2340813), Alexa Fluor® 488 AffiniPure Donkey Anti-Chicken IgY (IgG) (H+L) (1:250, Jackson ImmunoResearch, AB_2340375), Alexa Fluor™ 647 Donkey anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody (1:250, Invitrogen, A31573). Alexa Fluor 647 Phalloidin (1:250, Invitrogen, A22287) was used to stain F-actin in muscle.

Images were obtained with Leica STELLARIS 5 confocal microscope. Images were obtained as Z series with the same interval. Z series images were merged by ImageJ (Image-Stacks-Z projection-Max Intensity), and then the mean fluorescence intensity was measured (Analyze-Measure-Mean gray value) or nucleus number was counted by Point Tool. Quantification graphs are generated by GraphPad Prism. P-values from unpaired t-test. Error bar, standard deviation (SD).

DEG analysis

To identify genes that are differentially expressed in the aged population, we applied the Wilcoxon Rank Sum test (rank_genes_groups() in SCANPY) to compare the gene expression between the aged (30d, 50d, and 70d) and the young population (5d). Genes with FDR lower than 0.05 would be considered to be differentially expressed.

Change of cell composition

Cell numbers of each cell type were counted and then separated by age. To get the ratio of one cell type from one age, the cell numbers of a specific cell type and age were further divided by the total cell number of the corresponding age. The cell-type ratios from the aged population (30d, 50d, 70d) were further divided by ratios from the young population (5d) to get the relative ratios in log2 scale.

GO analysis

Age-specific DEGs, including up- or down-regulated genes, from each cell type were applied for the Gene Ontology (GO) analysis using GOATOOLS (version 1.2.3) (Klopfenstein et al., 2018). The gene association file (version FB2022_04) was downloaded from the FlyBase. Redundant GO terms were removed by REVIGO (Supek et al., 2011). GO terms from Biological Process (BP) were mainly used for our analysis (**Fig. 3E, 3F, and fig. S16**). GOs restricted within 5 cell types were considered as cell-type specific GOs (**Fig. 3E, 3F, and fig. S19A**).

Aging clock analysis

For aging clock analysis, we focused on 64 head and body cell types that have at least 200 cells at each time point. For each cell type satisfying the condition, we then built aging clock models – regression models that predict age from the transcriptome. In particular, for each cell type we trained the elastic net model (Zou & Hastie 2005) by using cells' transcriptome as explanatory (independent) variables and age as the response (dependent) variable. We randomly split the dataset into a train and test dataset where we used 70% of cells to train the model, and 30% of cells to validate the model. We compared the predictions to the true age and measured performance of the model as R^2, *i.e.*, proportion of the variance for a dependent variable that's explained by an independent variable or variables. We repeated the analysis over five random seeds, resulting in different train and test data splits. To find genes that correlate with age, we first filtered all genes with zero coefficient in the regression model. For the remaining genes, we calculated Pearson correlation between its value and age and retained those genes for which the correlation coefficient was higher than 0.3. This analysis resulted in 480 genes. We repeated the same analysis for the Mouse Aging Cell Atlas (Tabula Muris Senis) (Tabula Muris Consortium, 2020)**.**

To compare models between two consecutive stages, we trained three logistic regression models for each cell type to distinguish between (i) 5 days and 30 days old fly, (2) 30 days and 50 days old fly, and (iii) 50 days and 70 days old fly. We split the data into train and test dataset by using 70% data for training the model and 30% data for evaluation by preserving the percentage of samples for each class (stratified sampling). We evaluated classification performance using accuracy.

Cell-identity analysis

Change of cell-identity was based on the expression changes of marker genes between different ages. Cell-type markers from young (5d) and old animals (50d) were identified separately using the Wilcoxon Rank Sum test (rank_genes_groups() in SCANPY). Top 200 genes that are differentially expressed between cell types were defined as marker genes for each cell type. Expression of marker genes were further compared between young and old population using Wilcoxon Rank Sum test to identify differentially expressed markers (FDR < 0.05). Young marker genes with significant reduction of expression would be considered to decrease the cell-type identity. On the other hand, old marker genes with significant increase of expression were also considered to decrease the identity. Together, we used the average of these two ratios to indicate the change of cell identity.

AFCA website

The AFCA Website was developed with the Shiny package (v1.6.0) in R language (v4.0.5). We included three major datasets: Head dataset, Body dataset, and Head and Body dataset, where they contain the snRNA-Seq analytical results from head tissues, body tissues, and combined respectively. For the Cell type and Gene Expression tabs, the data processing and visualization were performed with the Tidyverse (v1.3.0) and ggplot2 (v3.3.3) packages in R. For the customized analysis tab, its features and functions were powered by the ShinyCell (v2.1.0) package.

Rank sum of aging features

We measured the absolute differences of each aging feature between the young (5d) and old (50d) population and ranked the cell-type differences. High ranks represented high aging differences, while low ranks indicated low differences between young and old population. To have a better estimation, we focused on cell types with more than 200 nuclei from each age. Correlations of different aging features were compared using the Spearman's correlation. To know which cell types showing high aging differences supported by different features, we summed the ranks of five different aging features and ranked the sums. High rank sums of different aging features represented high aging differences, and vice versa.

Analysis of alternative polyadenylation (APA) analysis

Our overall analysis strategy was based on our recent scAPA study of the FCA dataset (Lee et al., 2022). We used LABRATsc (https://github.com/TaliaferroLab/LABRAT) to quantify APA from the scRNA-seq data with current FlyBase *Drosophila melanogaster* gene annotation (version r6.45). Two modes were used for either single-cell level (cellbycell) or cell-type specific (subsampleClusters) quantification using '--mode' parameter. We used read coverage thresholds of at least 3 or 100 counts per gene for single-cell level or cell-type specific quantification respectively to minimize noise and false positives. LABRAT calculates ψ values reflecting 3' isoform usage, with "0" indicating exclusive usage of the most proximal pA site, while "1" indicates exclusive usage of the most distal pA site. Ψ values for each annotated cell-type, and for individual cells, were used for downstream analysis.

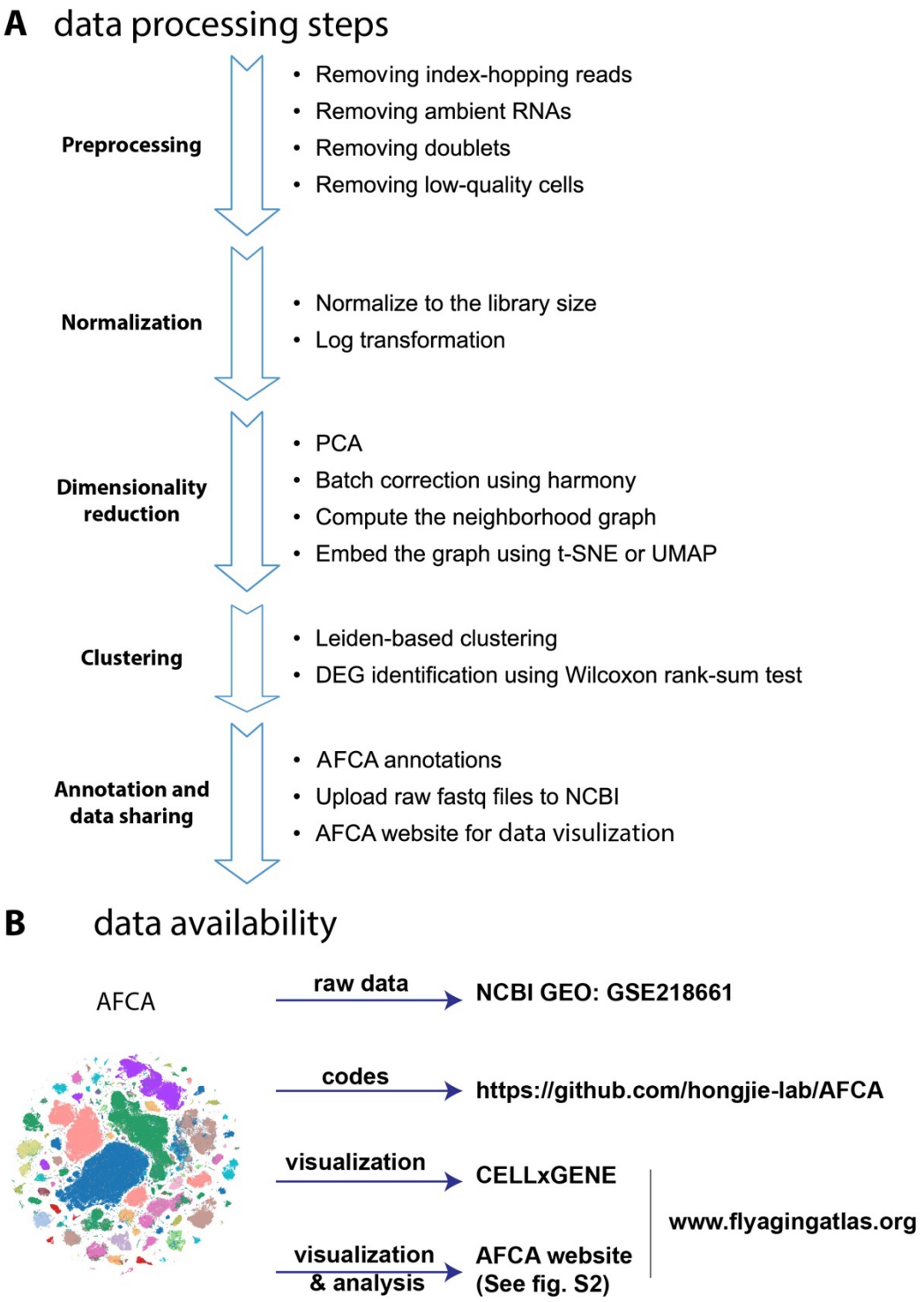

Fig. S1. General pipeline of snRNA-seq analysis and data availability.
(A) Data processing steps, including preprocessing, normalization, dimensionality reduction, clustering, and annotation and data sharing. (B) Data available from different sources, including NCBI, GitHub, CELLxGENE, and AFCA website.

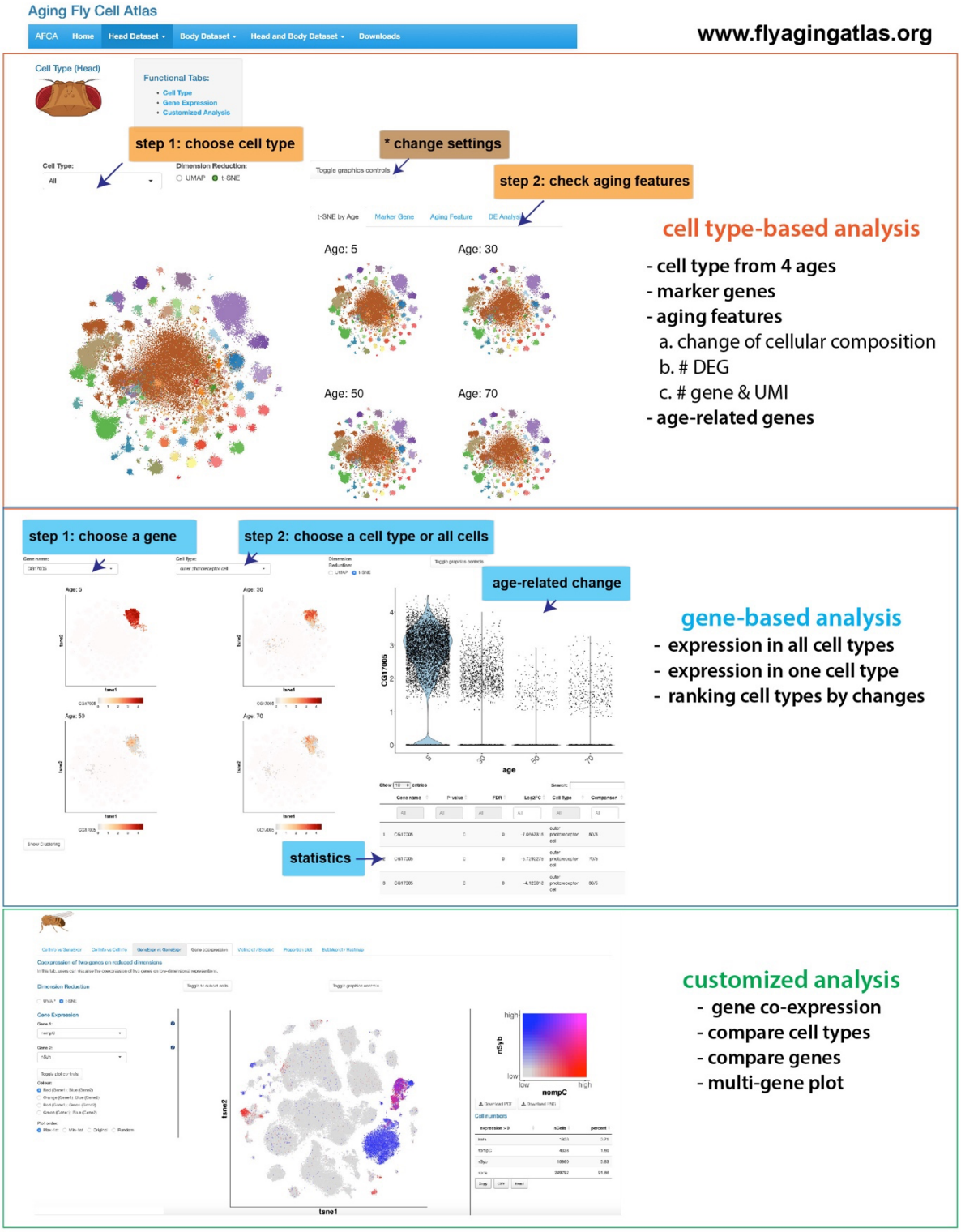

Fig. S2. AFCA web portal for surveying the age-related changes.
AFCA web portal provides cell type-based, gene-based, and customized analysis.

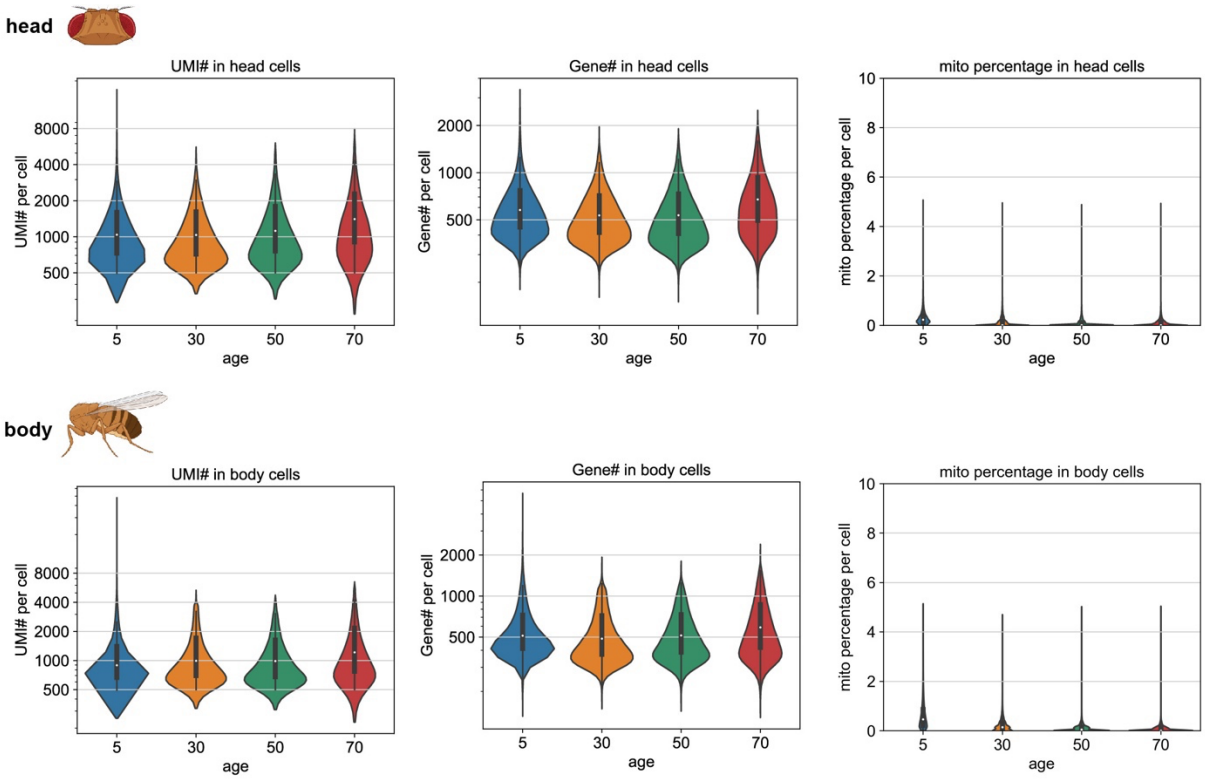

Fig. S3. General quality of snRNA-seq results.
UMI number, expressed gene number, and mitochondria transcript ratio from the AFCA data.

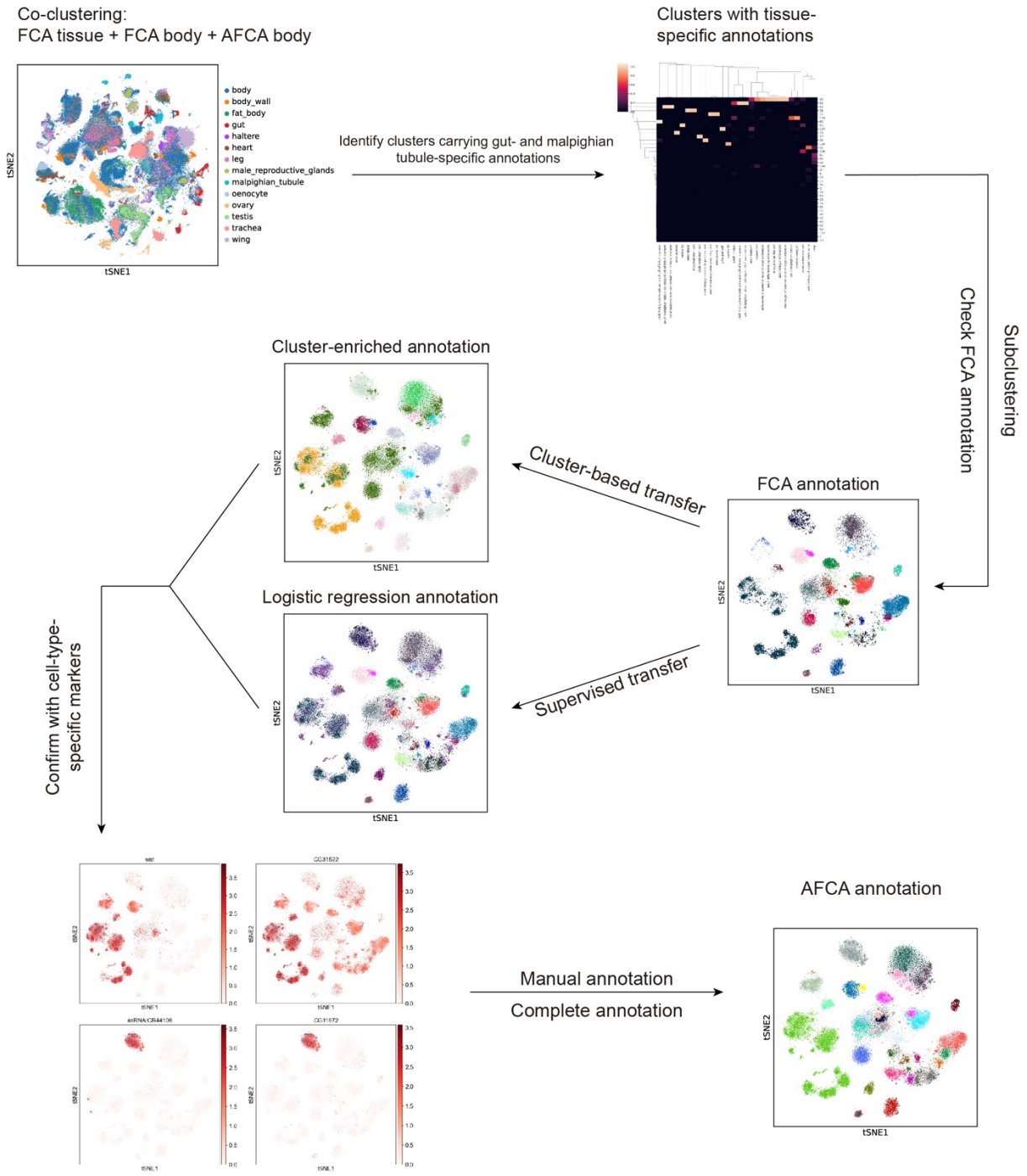

Fig. S4. Illustration of transferring the FCA annotation to the AFCA cells.
Flowchart of transferring the annotations from FCA to AFCA. FCA and AFCA data are first co-coustered. Co-clustered data can be further extracted for tissue-specific cell types, e.g. gut- and malpighian tubule-specific cell types. After subclustering targeted cells, FCA annotations are transferred to AFCA cells using cluster-based or supervised machine-learning approaches. Results from two transfer methods are compared and ambiguous transfers are further checked

for the expression of the corresponding marker genes. After confirming the expression of the marker genes, annotations are corrected and finalized.

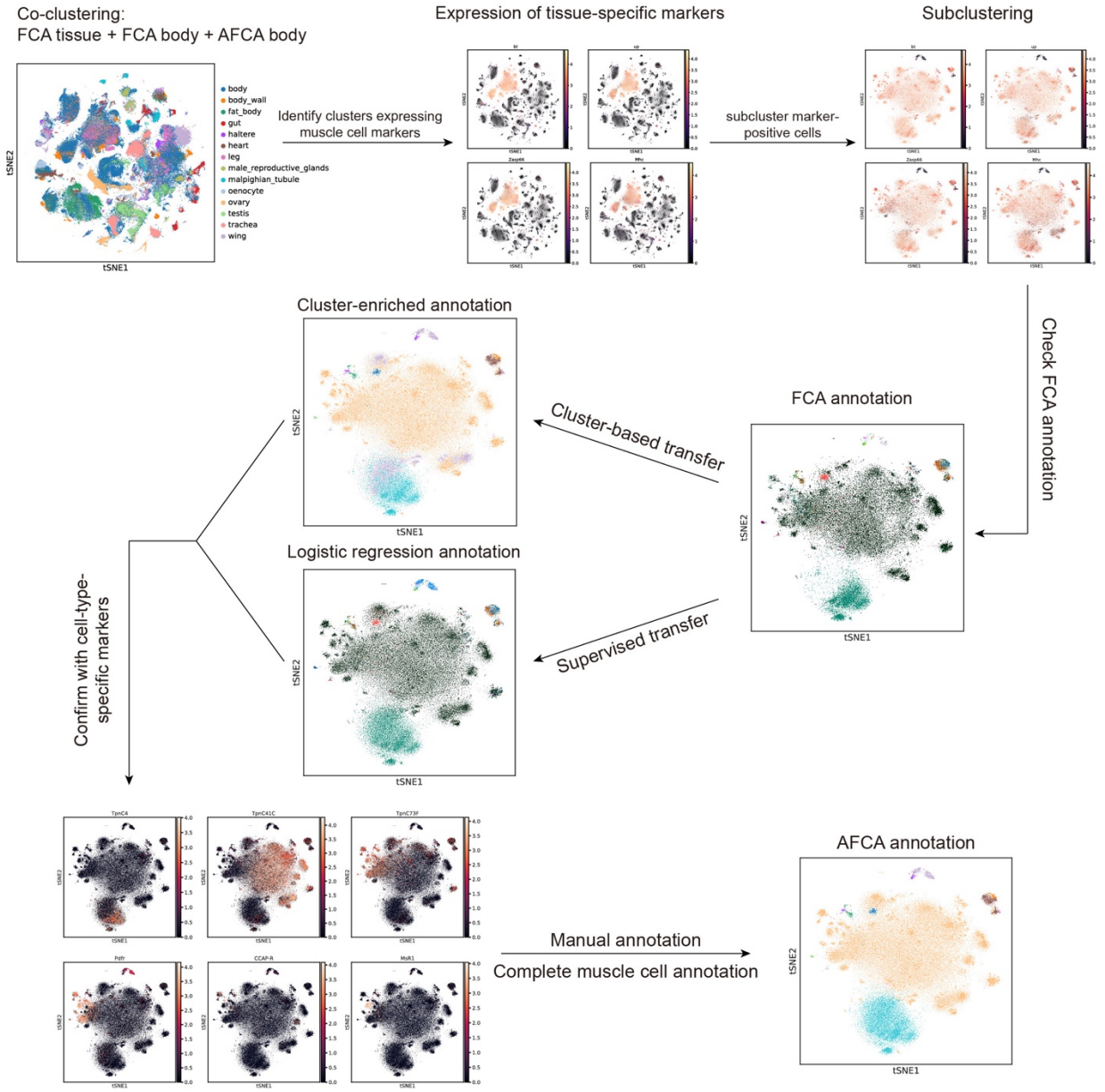

Fig. S5. Illustration of transferring the FCA annotation to the AFCA muscle cells. Flowchart of transferring the muscle-type annotations from FCA to AFCA. Some cell types, like muscle cells, can be found in multiple tissues and we consider first extracting those cells based on the expression of general markers of muscle cells. After subclustering the corresponding cells, all remaining analyses are similar to approaches used in **Fig. S4**.

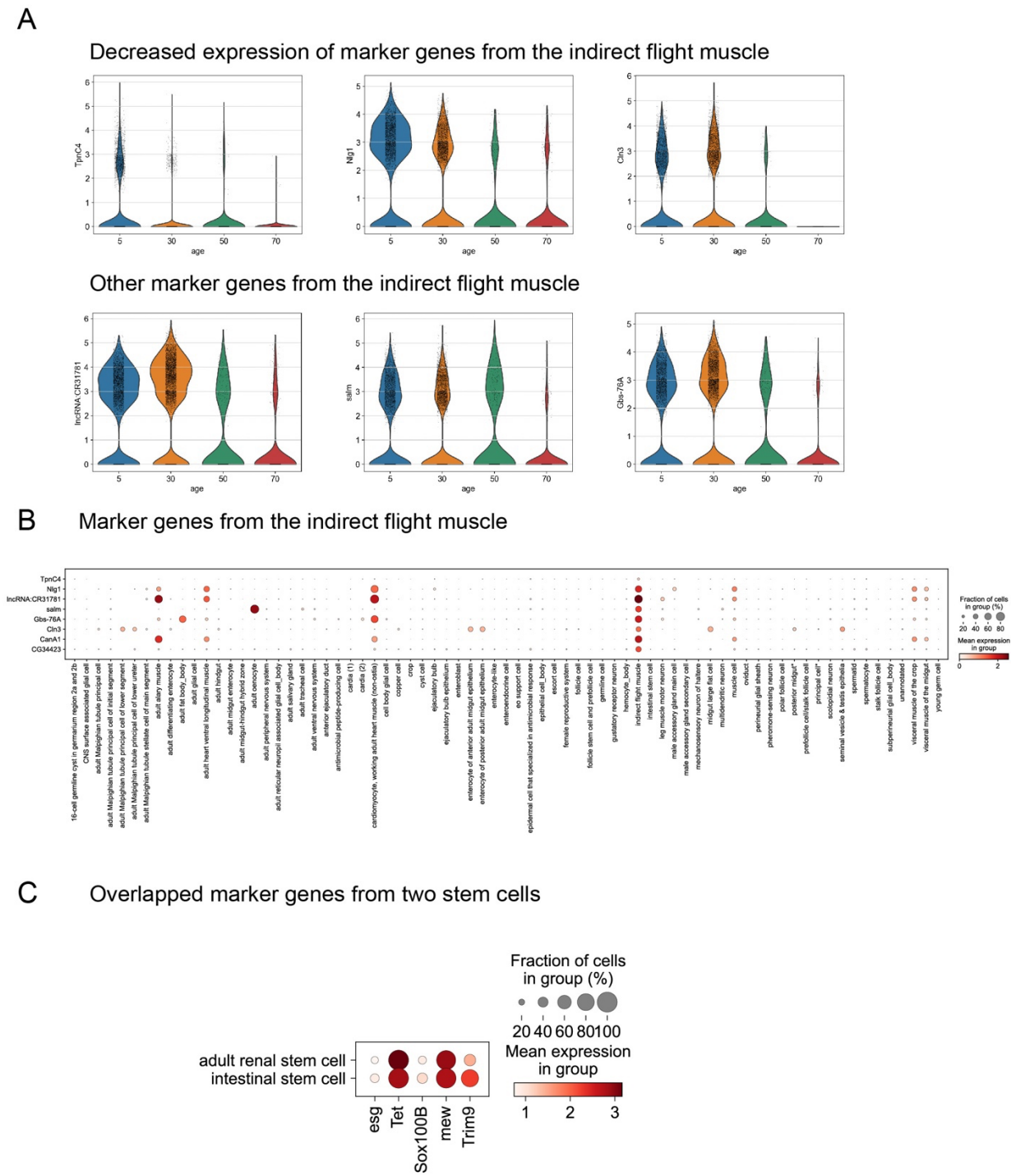

Fig. S6. Ambiguous annotations from the aging samples.
(A, B) Marker genes of the indirect flight muscle. (A) Loss of marker genes in the indirect flight muscle. (B) Markers specific for the indirect flight muscle. (C) Overlapped marker genes from renal stem cells and intestinal stem cells.

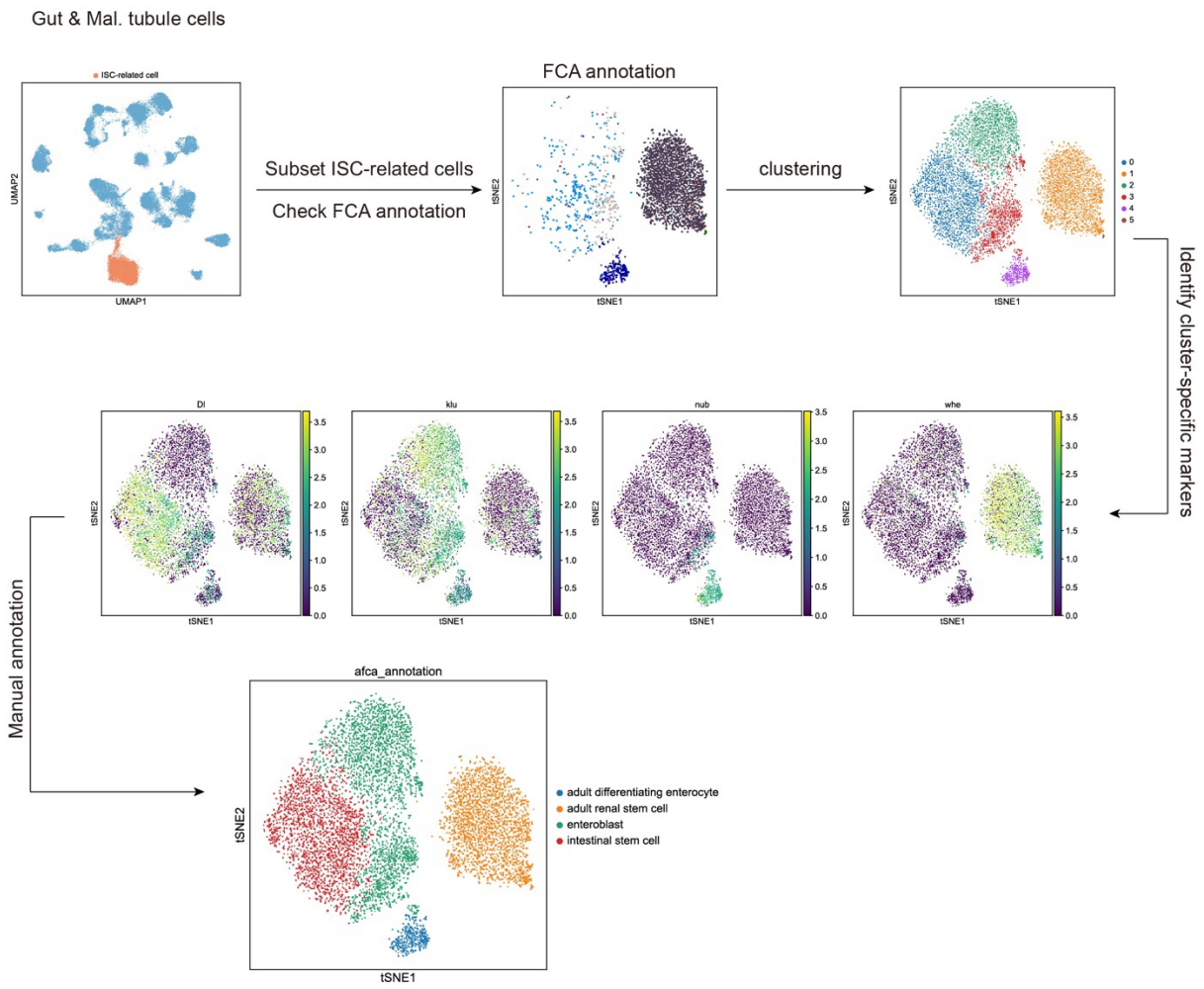

Fig. S7. Illustration of manually annotating ISCs, EBs, and renal stem cells.
Flowchart of manual annotation of stem cell population. ISCs, EBs, and differentiating enterocyte cells are selected from the gut and malpighian tubule dataset shown in Fig. S3. Marker genes enriched in each cluster are identified and confirmed.

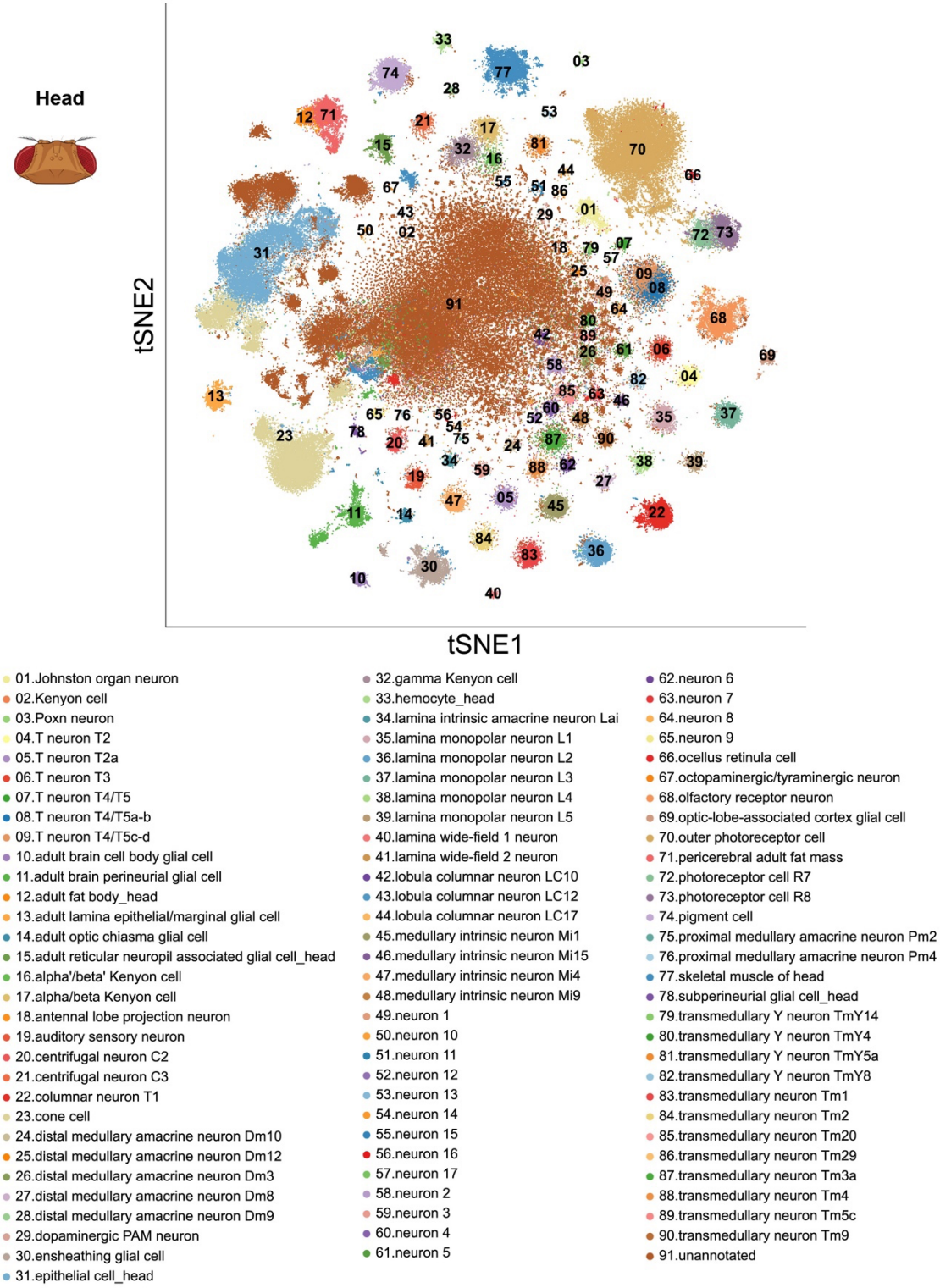

Fig. S8. Detailed cell types annotated in the head sample shown by tSNE.
90 head cell types are shown in the tSNE plot.

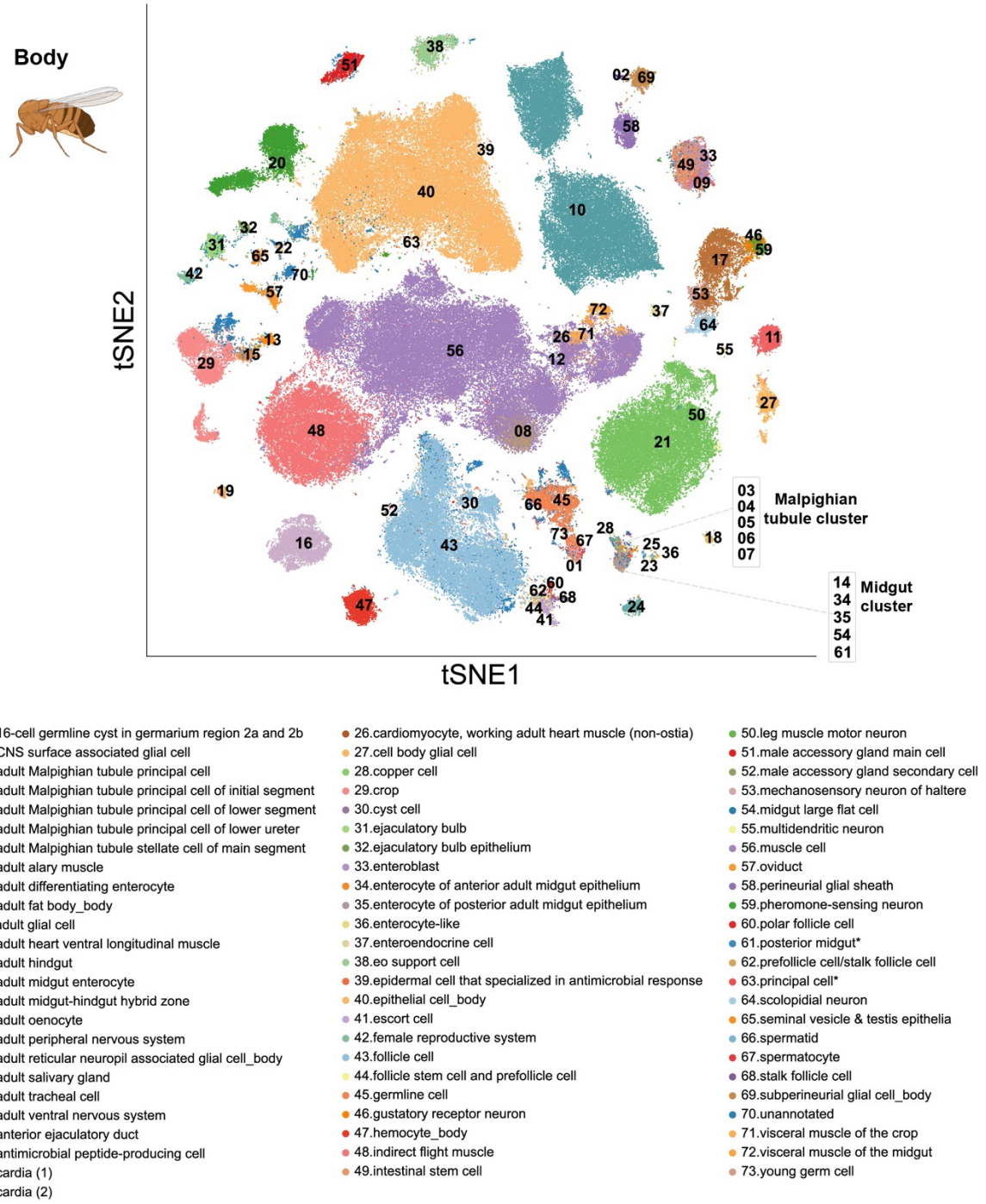

Fig. S9. Detailed cell types annotated in the body sample shown by tSNE.
72 body cell types are shown in the tSNE plot.

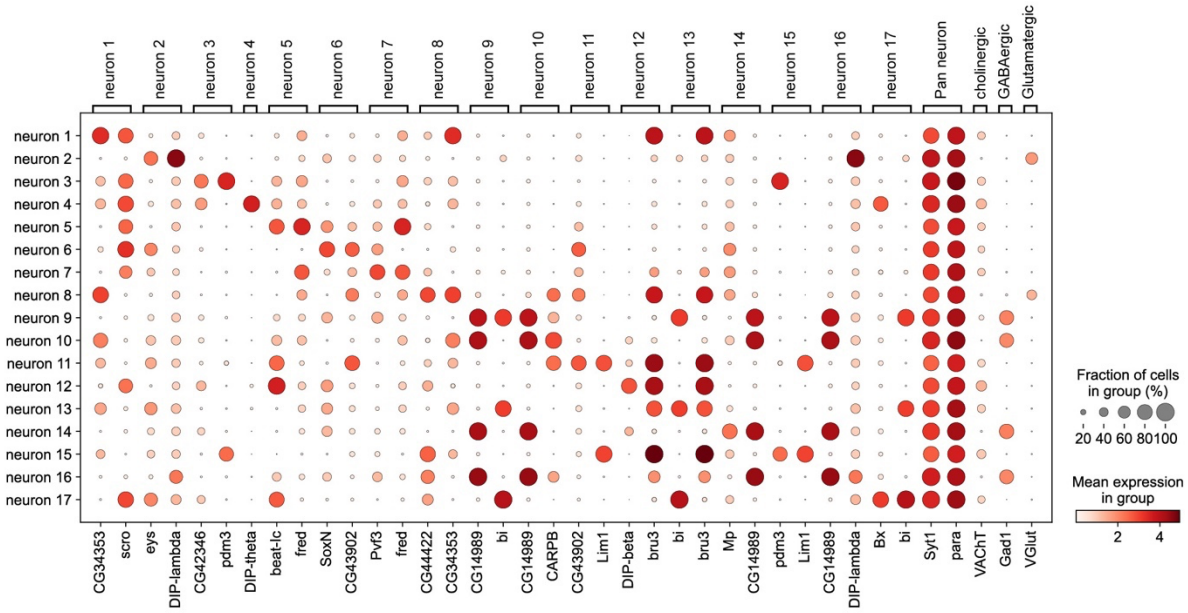

Fig. S10. Marker genes identified from the newly identified neuron clusters.
Markers of 17 newly identified neuron types are shown. Expression of general neuronal markers, Syt1, and para, indicates the neuron specificity. Expression of VAChT, Gad1 and VGlut indicates they are cholinergic, GABAergic, or Glutamatergic neurons.

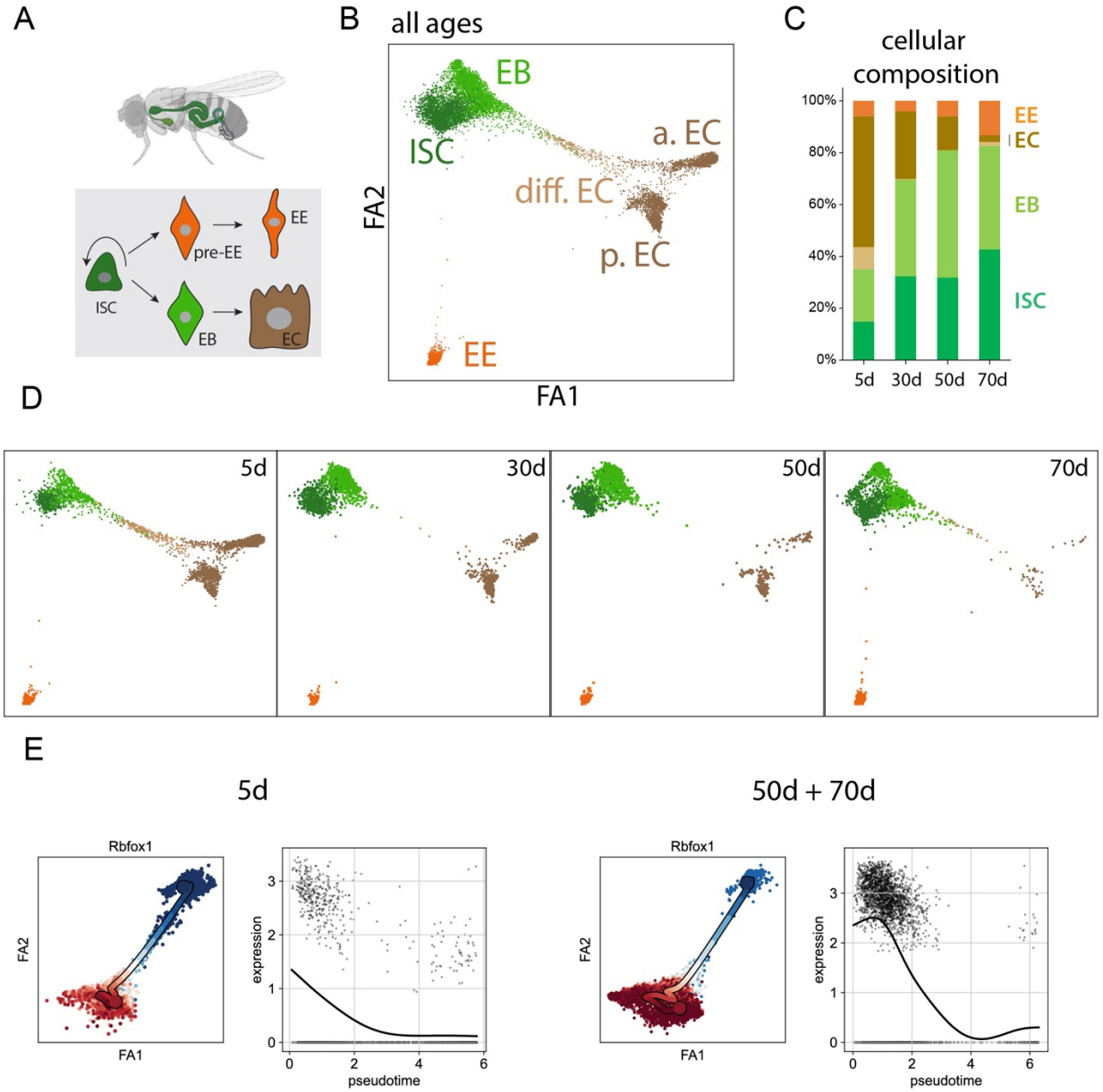

Fig. S11 Pseudotime inference and cellular composition of ISC and ISC-differentiated cell types.

(A) Illustration of ISC differentiation process. (B) Cell types shown in different colors. (C) Changes of cellular composition in the ISC lineage. (D) Age-specific profile of different cell types. (E) Rbfox1 gene shows different patterns along the young and old ISC lineage. ISC, intestinal stem cell; EB, enteroblast; EC, enterocyte; EE, enteroendocrine cell; a. EC, anterior EC; p. EC, posterior EC; diff. EC, differentiating EC.

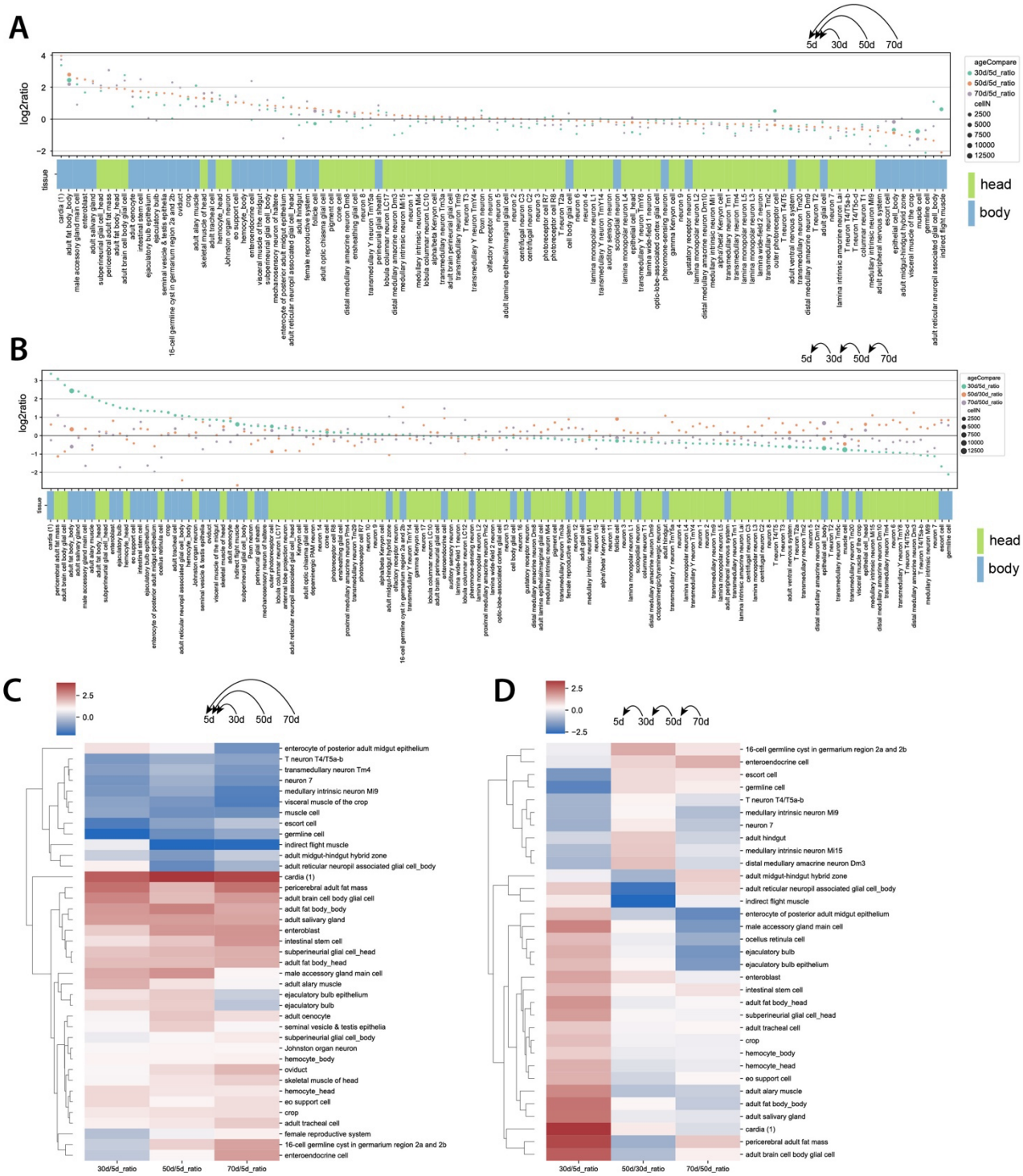

Fig. S12 Changes of cell-type composition during aging.

(A) Changes of cellular composition by comparing the nucleus ratio of the old sample (30d, 50d, and 70d) to the ratio of the young one (5d). (B) Changes of cellular composition by comparing the nucleus ratio between two consecutive ages. (C) Cell types with more than 2-fold changes of

cellular composition by comparing the old samples to the young ones. (D) Cell types with more than 2-fold changes of cellular composition by comparing the sample to the previous age.

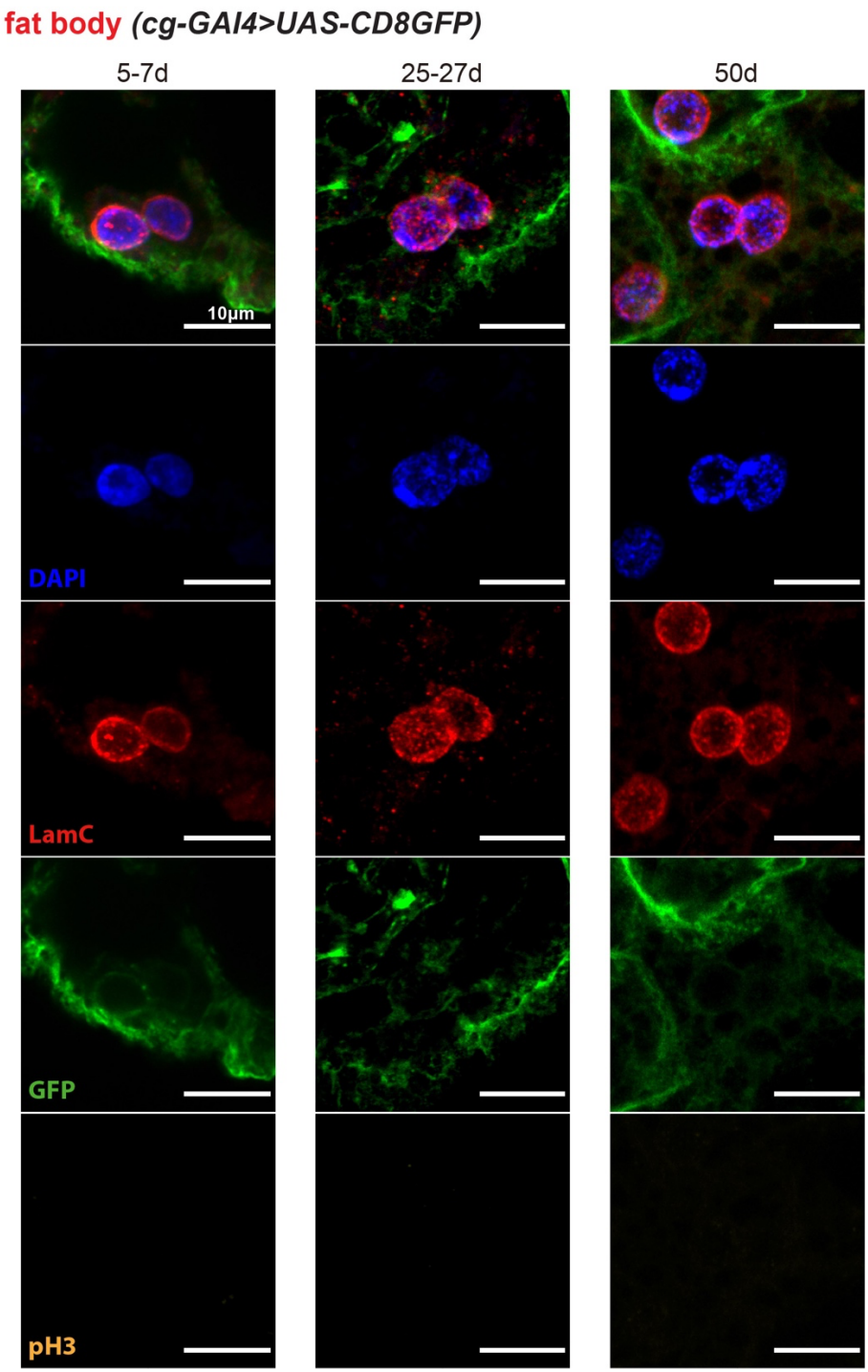

Fig. S13. Examples of dividing nuclei without the mitosis marker.
Dividing nuclei from flies with different ages showing the absence of the mitosis marker (pH3). Fat body nuclei are collected from different ages, including 5-7d, 25-27d, and 50d. Dividing nuclei are stained by DAPI and LamC. The fat body membranes are labeled by *cg-GAL4 > UAS-CD8GFP*.

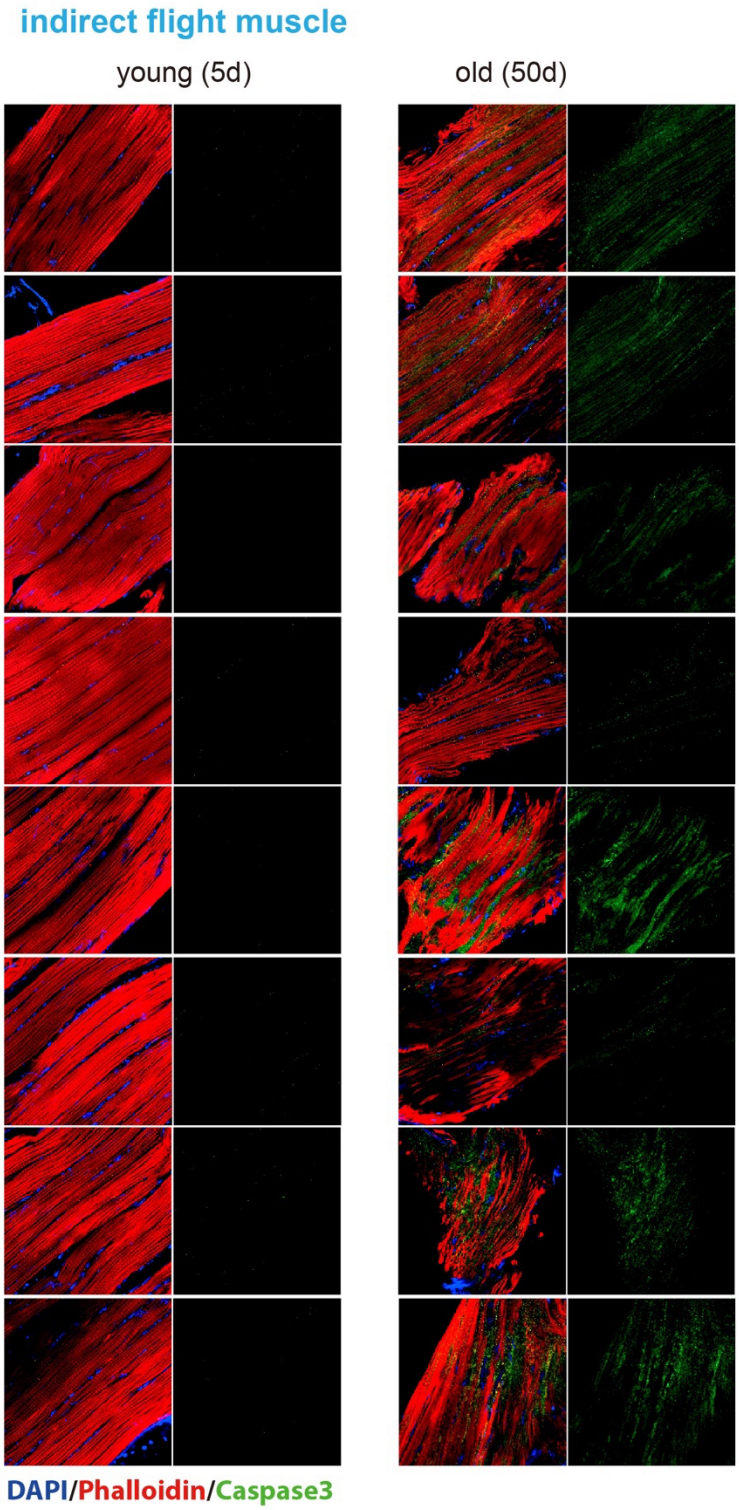

Fig. S14 Apoptosis signals form the aged indirect flight muscle.
Examples of indirect flight muscle stained by the apoptosis marker in the young and old flies. The indirect flight muscles are stained for the apoptosis events using cleaved-Caspase3 antibody. Nuclei are stained by DAPI. Actin filaments in the muscle cells are labeled by Phalloidin.

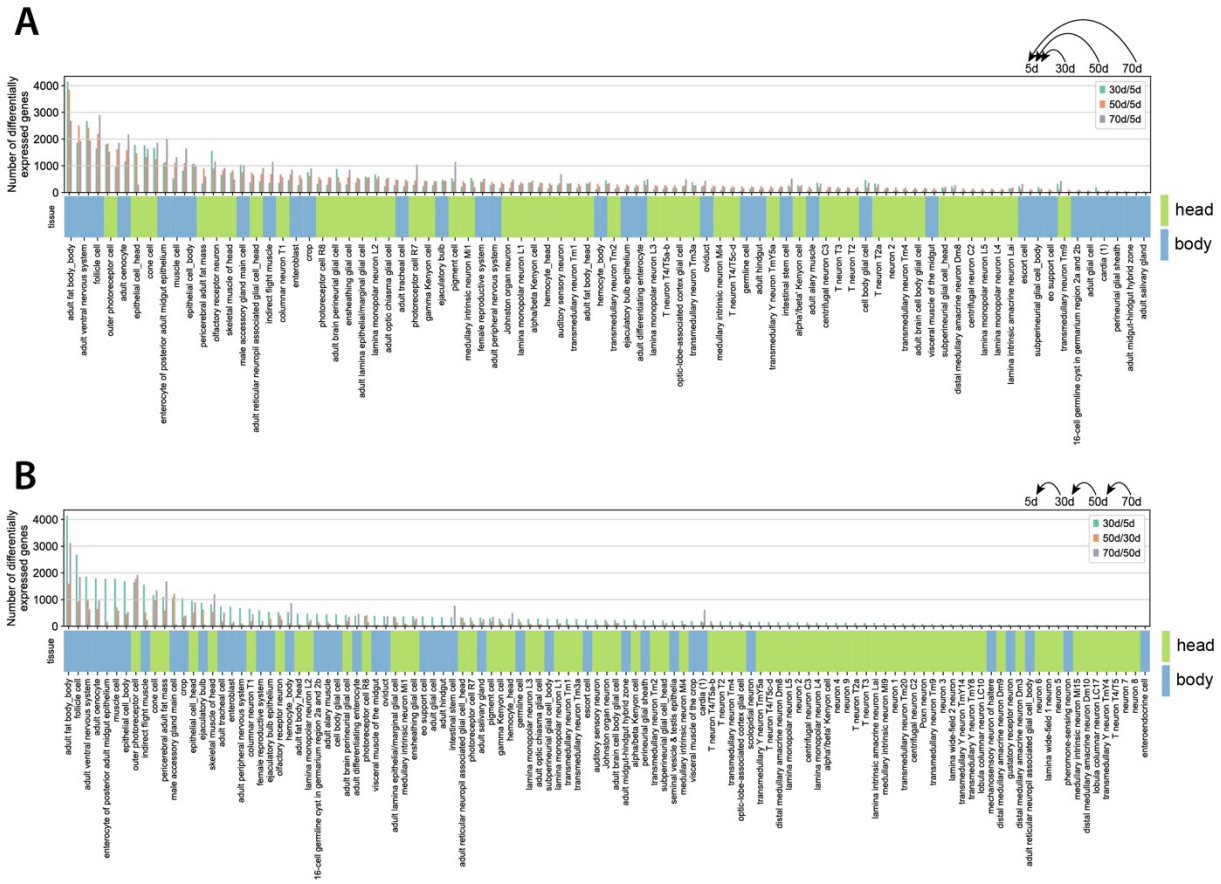

Fig. S15 DEG numbers detected in each cell type.
(A) DEG numbers from different cell types. The aged groups (30d, 50d, and 70d) are compared with the young population (5d). (B) Two consecutive ages are compared to identify DEGs in each cell type.

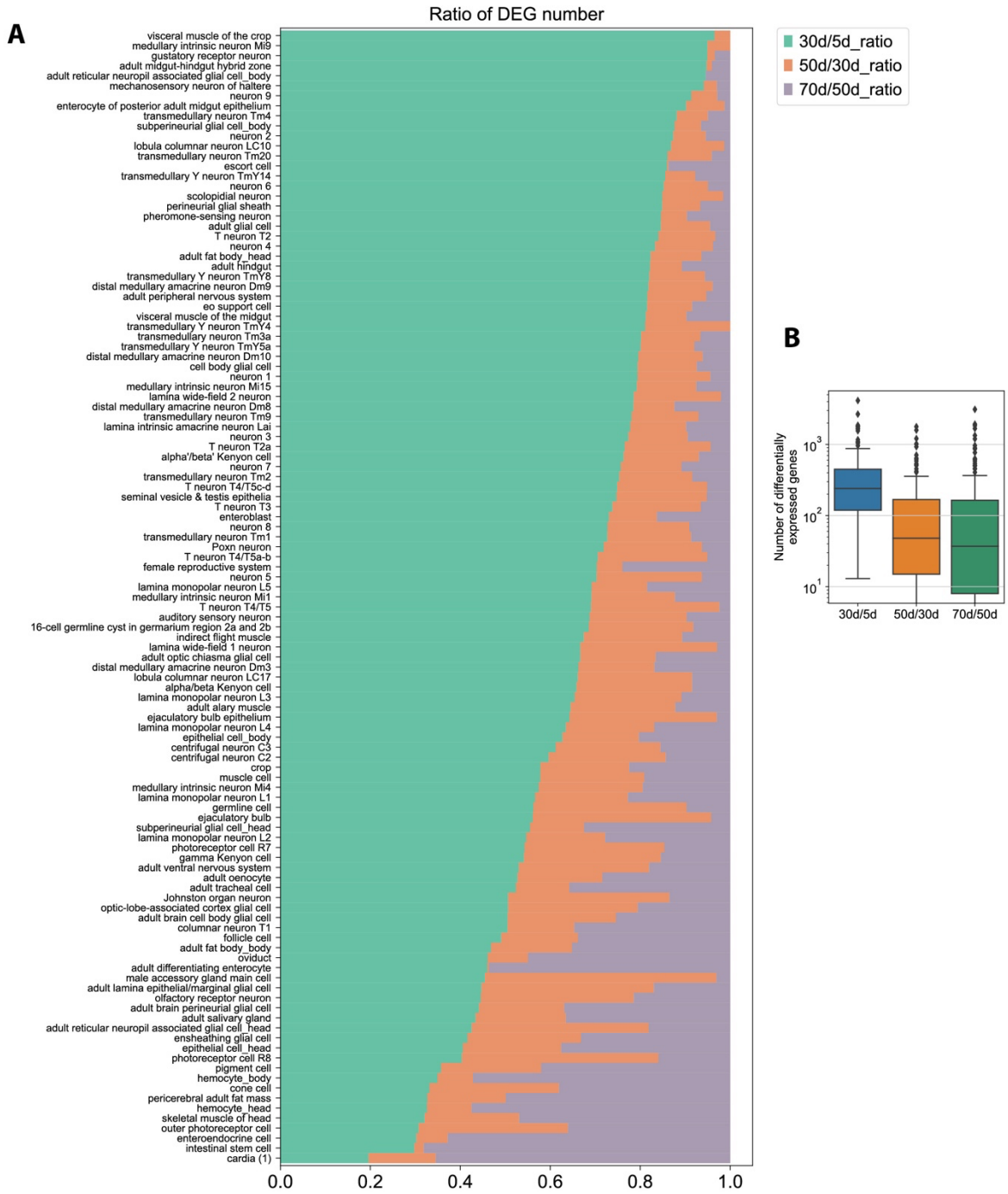

Fig. S16 Detailed ratio of DEG number across different age periods.
(A) The ratios of DEG numbers in different cell types. Each age period is shown in a different color. (B) Number of differentially expressed genes detected in the different age period.

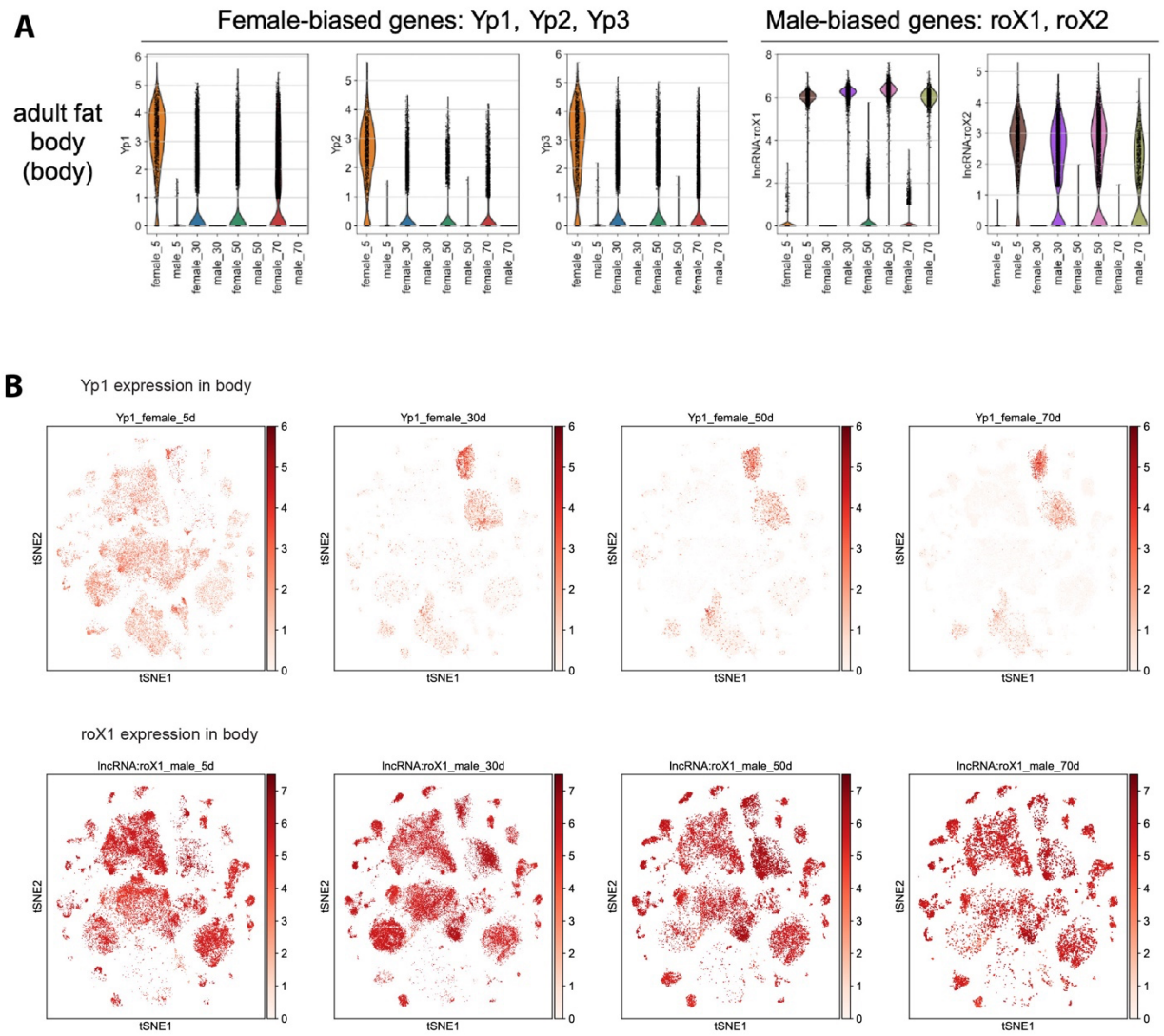

Fig. S17 Sex-specific expression of genes.

(A) Expression of female- and male-biased genes from the fat body cells during aging. (B) Expression of Yp1 and roX1 in body cells shown by tSNE.

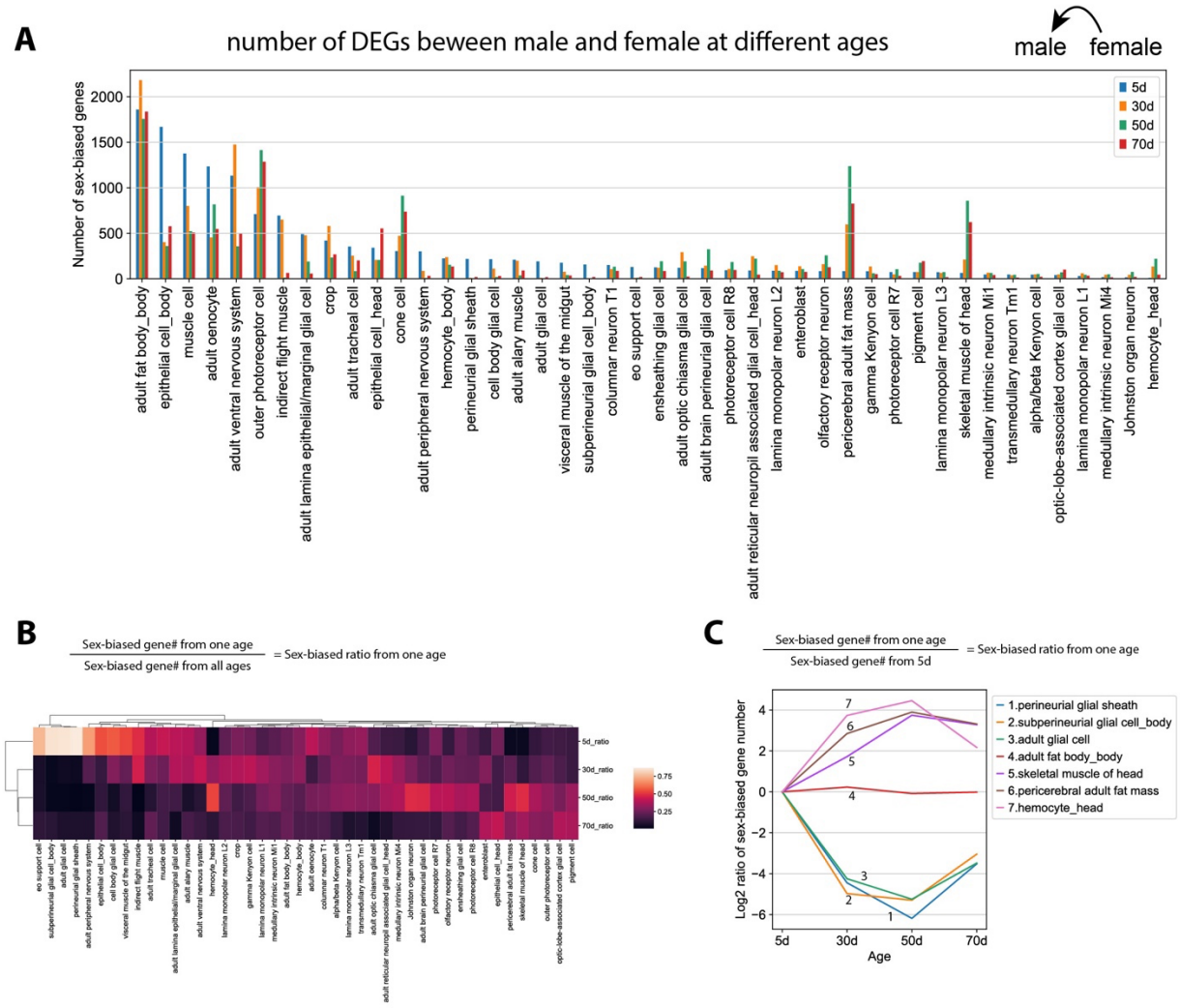

Fig. S18 Sex-biased genes from each age.

(A) The number of sex-specific genes from each age. (B) Sex-biased ratios from different ages. (C) Illustration of sex-biased ratios from several cell types.

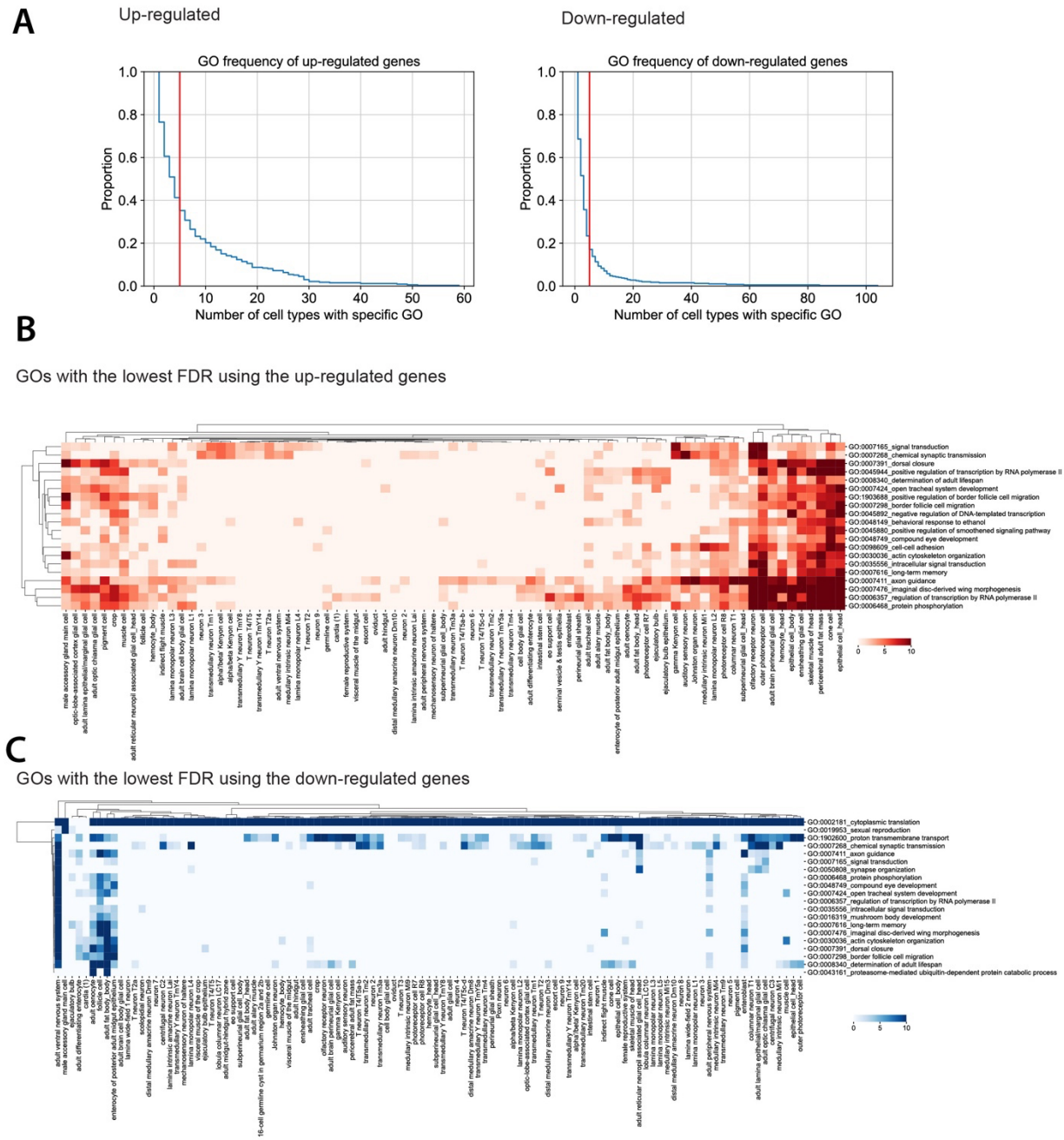

Fig. S19 GO frequency and the most significant GOs from the up- and down-regulated genes.

(A-B) Up- and down-regulated DEGs are identified by comparing 50d flies to 5d flies. (A) GO frequencies from up- and down-regulated genes. GO frequency indicates how many cell types have the corresponding GO. The red line indicates five cell types with the corresponding GO. ((B) GOs with the lowest FDR from the up-regulated genes. (C) GOs with the lowest FDR from the down-regulated genes.

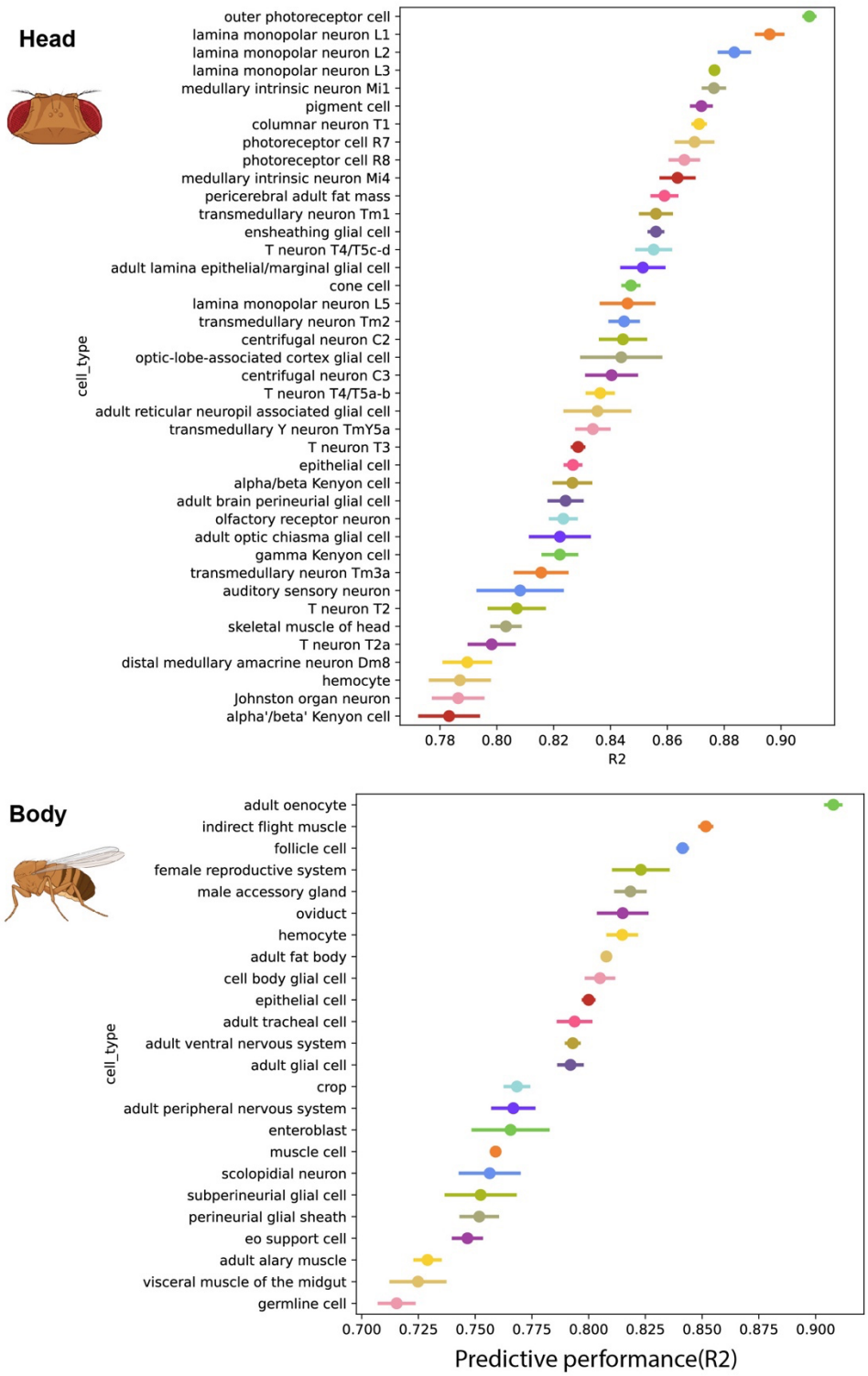

Fig. S20 Predictive performance of aging clocks.
Predictive performances of aging clocks from each cell type. Performances of head and body cell types are shown separately.

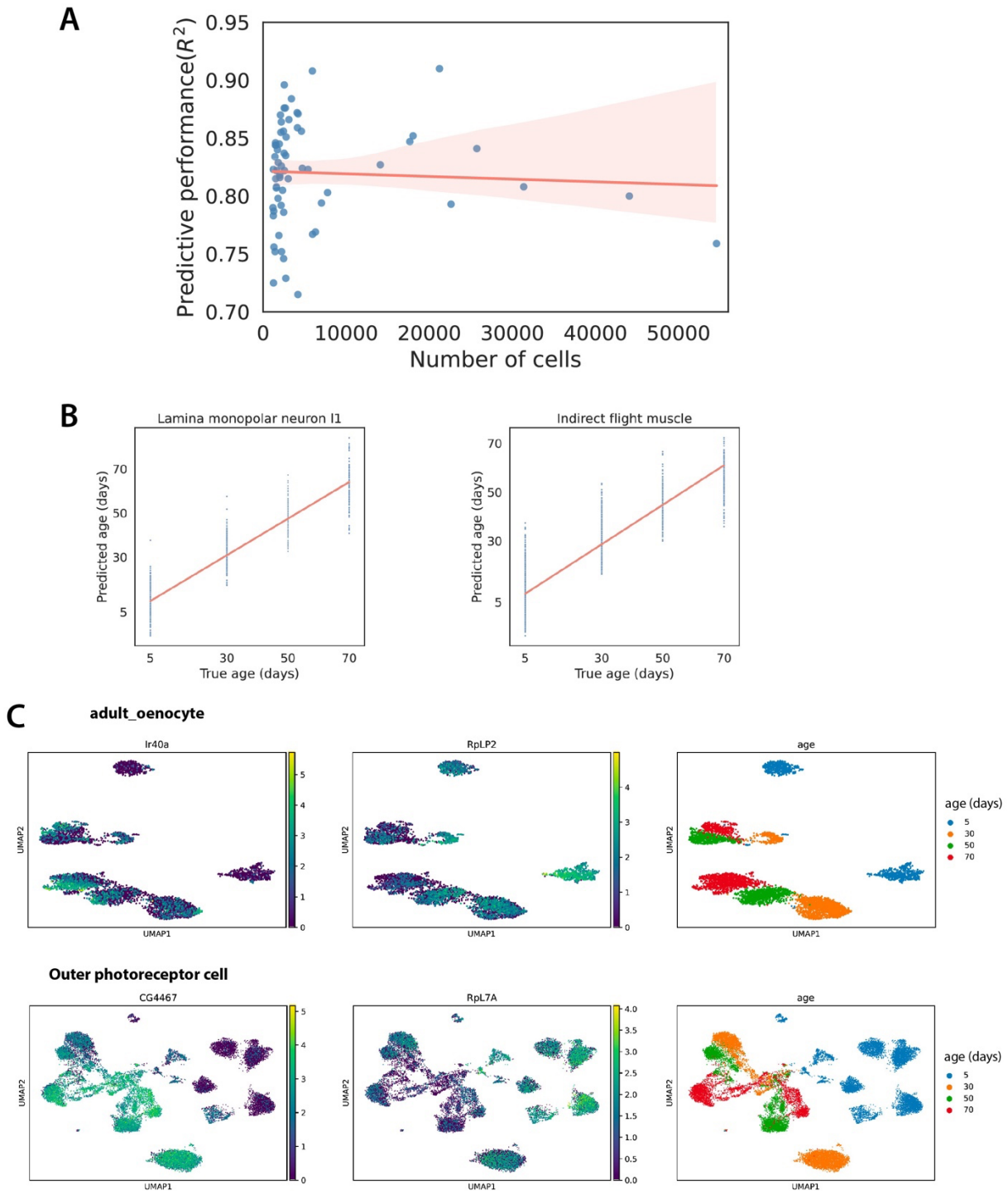

Fig. S21 Aging clock profiles.

(A) The cell numbers from each cell type are not correlated to the predictive performance. (B) Examples of two cell types with good prediction of true age. (C) Expression of aging clock genes in two different cell types.

fly_gene	mouse_ortholog	# cell types in fly	# cell types in mouse
rpl13a	rpl13a	8	24
rps25	rps25	13	18
rps29	rps29	5	24
rps26	rps26	10	10
rps21	rps21	3	15
rps15	rps15	15	2
rps23	rps23	6	11
rpl27a	rpl27a	15	1
rps7	rps7	13	2
rpl21	rpl21	11	2
rpl35a	rpl35a	3	10
rpl14	rpl14	11	1
rps12	rps12	7	5
rps8	rps8	11	1
rpl23	rpl23	8	3
rpl38	rpl38	1	10
rpl10ab	rpl10a	9	1
rps16	rps16	6	3
rps3	rps3	3	6
sta	rpsa	7	1
rpl13	rpl13	3	4
rpl31	rpl31	4	3
rps17	rps17	2	5
rplp0	rplp0	4	2
rps20	rps20	5	1
rps24	rps24	4	1
rps30	fau	3	2
atpsynbeta	atp5b	2	2
rpl11	rpl11	3	1
rps11	rps11	3	1
rps6	rps6	3	1
hsc70-3	hspa5	1	2
rpl24	rpl24	1	1

Fig. S22 Aging clock genes present in both fly and mouse.
1-1 orthologs of 33 aging clock genes conserved between fly and mouse. Number of cell types shown in each species indicate that the aging clock genes are detected as aging clock genes in the corresponding number of cell types.

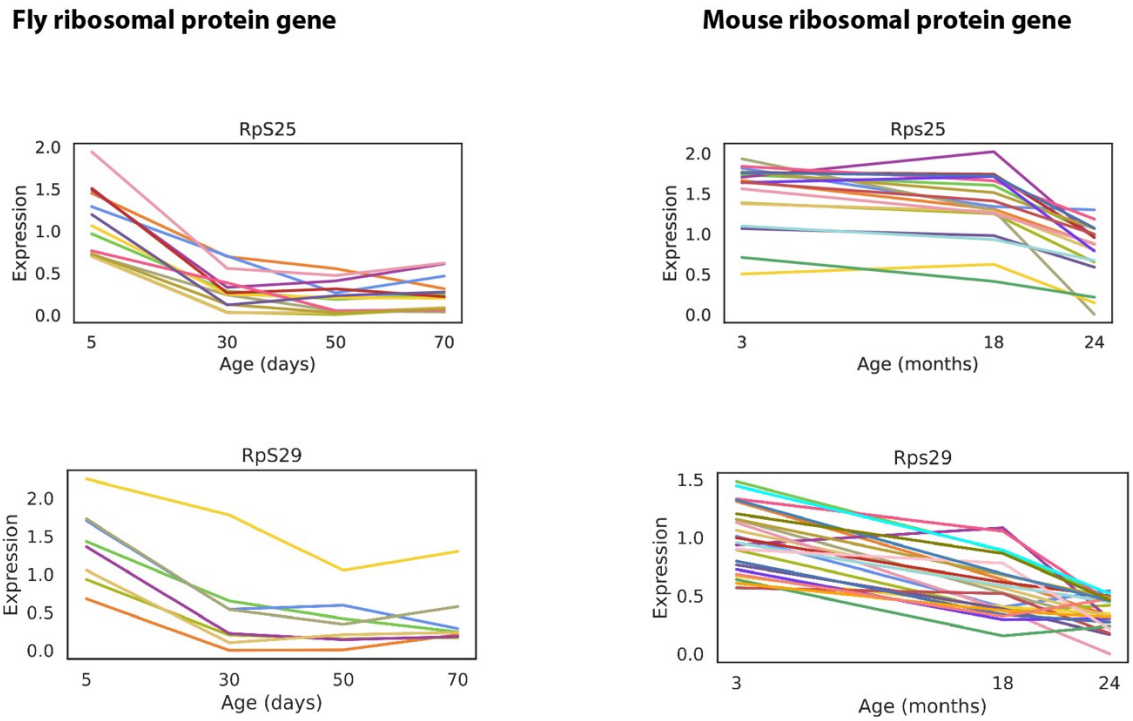

Fig. S23 Examples of cross-species comparison of aging clock genes.
Expression of two orthologous aging clock genes.

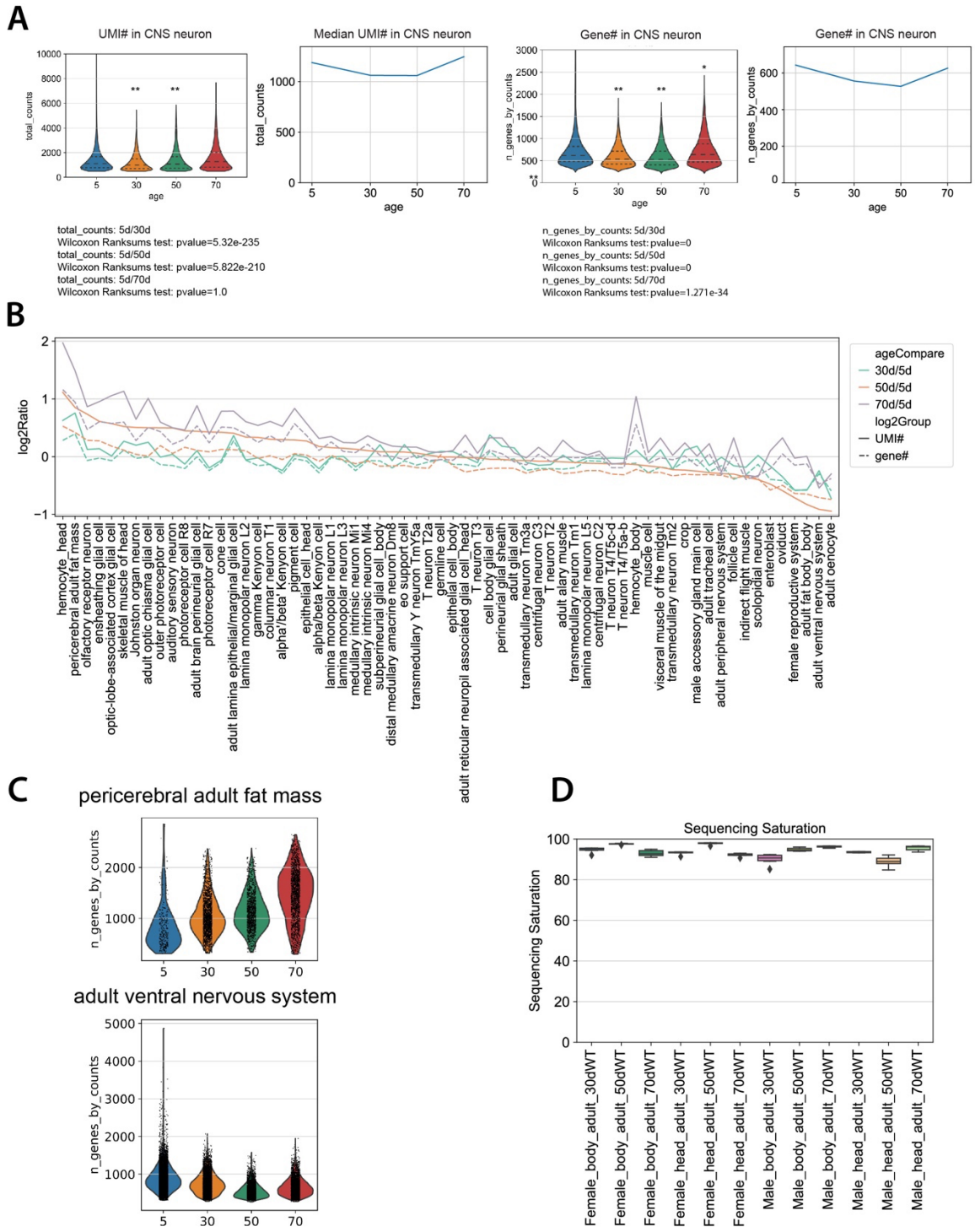

Fig. S24 Variance of UMI and expressed gene numbers during aging.
(A) Numbers of expressed genes and UMIs per cell are decreased in the aged CNS neurons (30d and 50d). (B) Log2 ratio of expressed gene or UMI number per cell during aging. (C) Changes of expressed gene number in pericerebral adult fat mass and adult fat body from the body. (D) Sequencing saturation of different batches of samples.

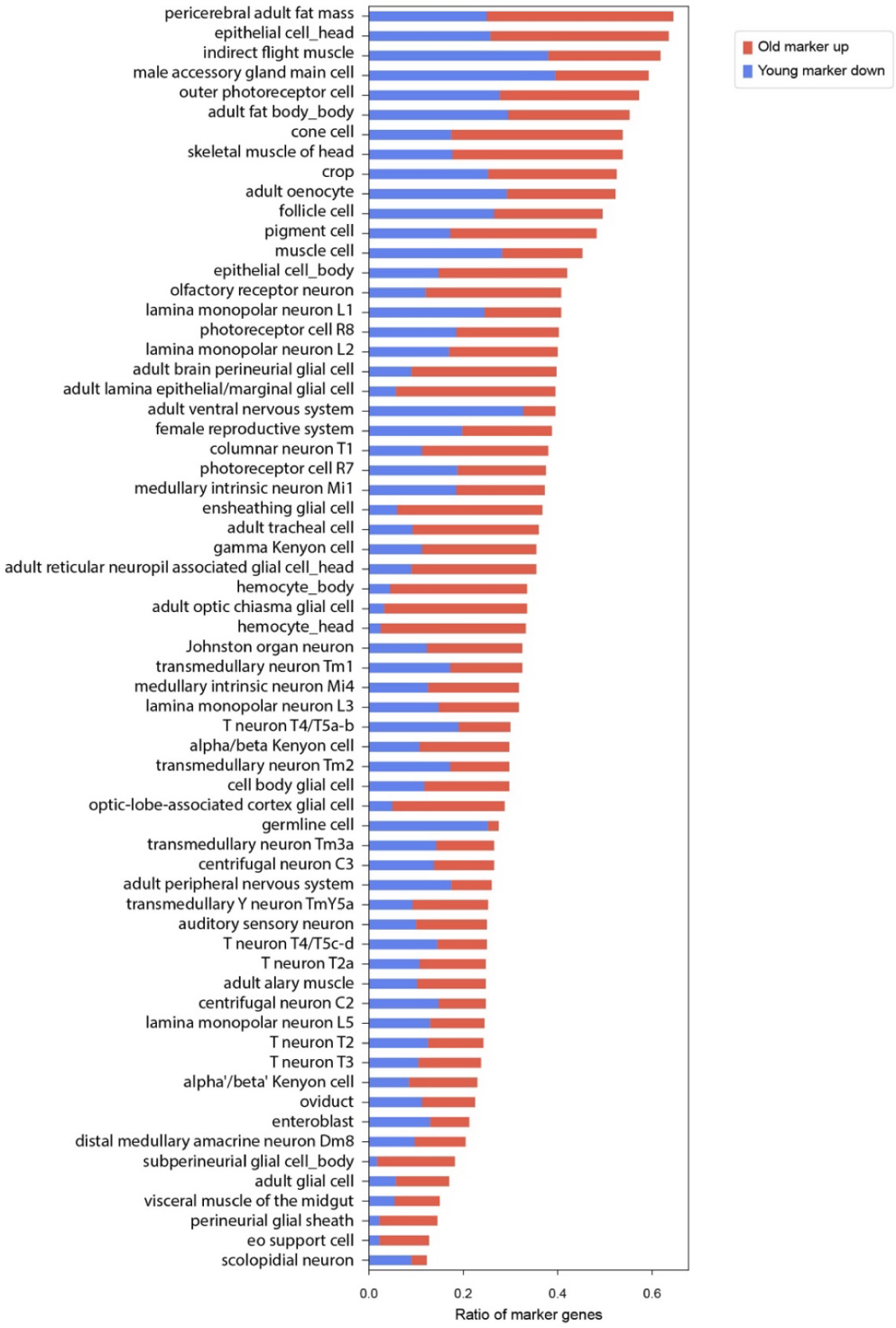

Fig. S25 Changes of cell-type identity during aging.
Cell types ranked by the decline of cell-type identity during aging. Decrease of young markers and increase of old markers are shown separately in each cell type.

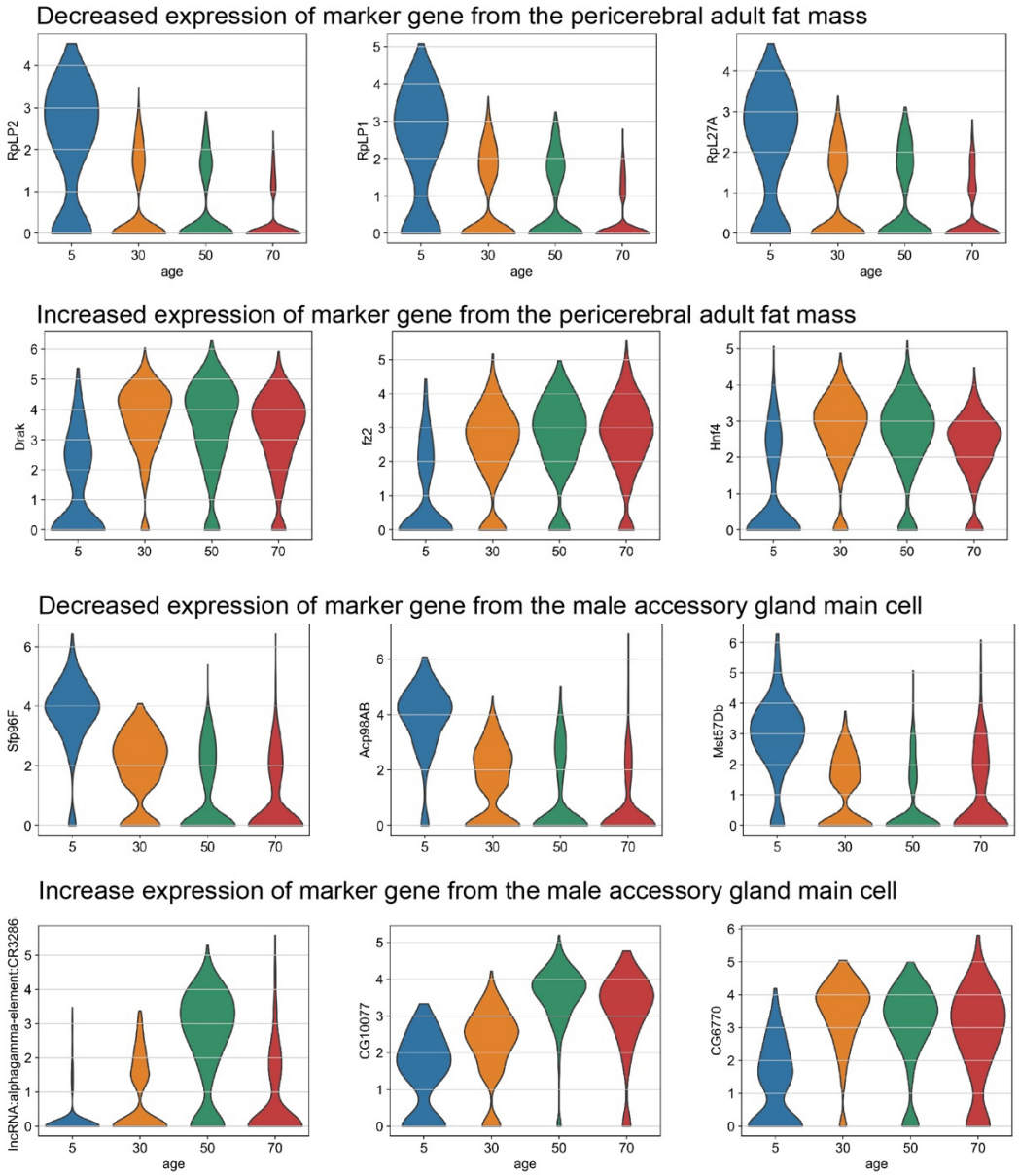

Fig. S26 Marker genes that affect cell-type identities during aging.
Examples of marker genes that caused the decline of cell-type identities.

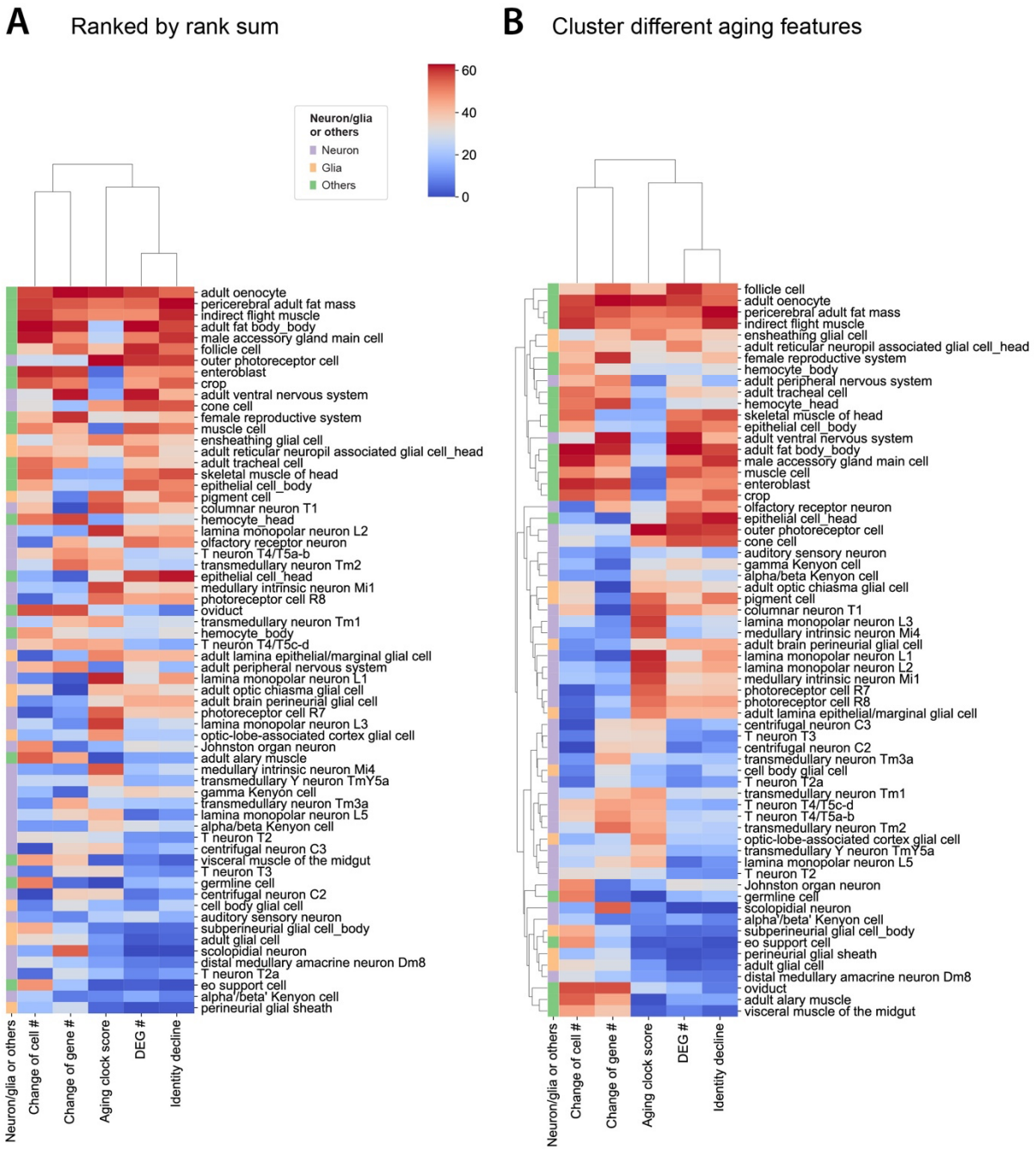

Fig. S27 Rank sums of different aging features.

(A) Cell types ranked by the sum of different aging feature scores. (B) Clustering of rank sum scores across different cell types.

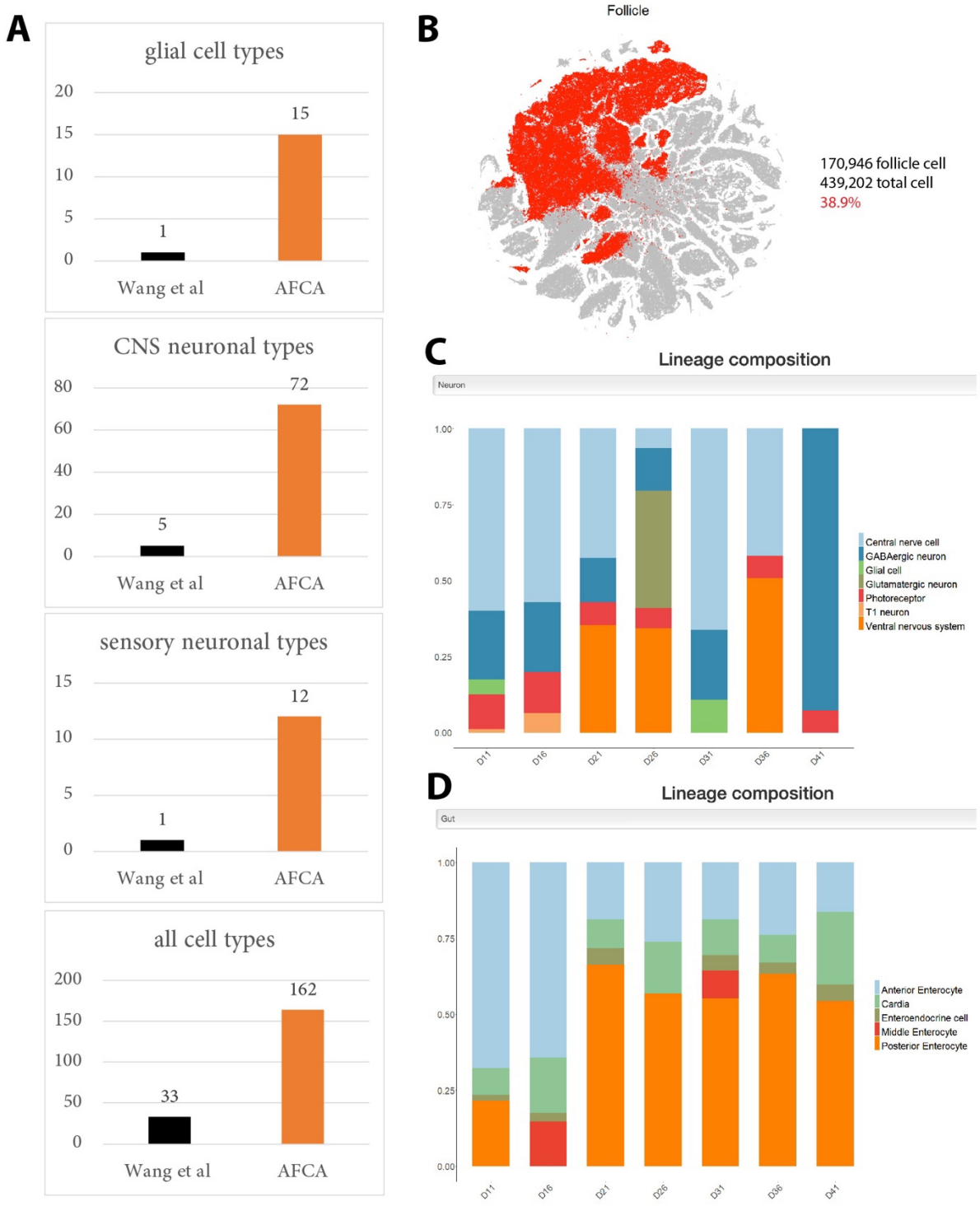

Fig. S28 Comparisons between the AFCA and the *Drosophila* cell landscape (DCL) datasets.

(A) A comparison of the number of annotated cell types between our AFCA and DCL (Wang et al., 2022). The AFCA dataset annotates significantly more cell types than DCL. For example, the AFCA annotates a higher number of cell types in glial, CNS neuronal, and sensory neuronal cells

compared to the DCL. (B) 38.9% of the DCL cells belong to the follicle cells of the ovary, indicating a sampling bias of this cell type. (C) The neuronal lineage composition of the DCL dataset. Note that glial cells are only detected in 2 out of 7 ages, and GABAergic neurons are not detected in 36d but account for 90% at 41d. (D) The gut lineage composition of the DCL dataset. Note that no intestinal stem cells are annotated. B-D plots are from the DCL website portal, http://bis.zju.edu.cn/DCL/.

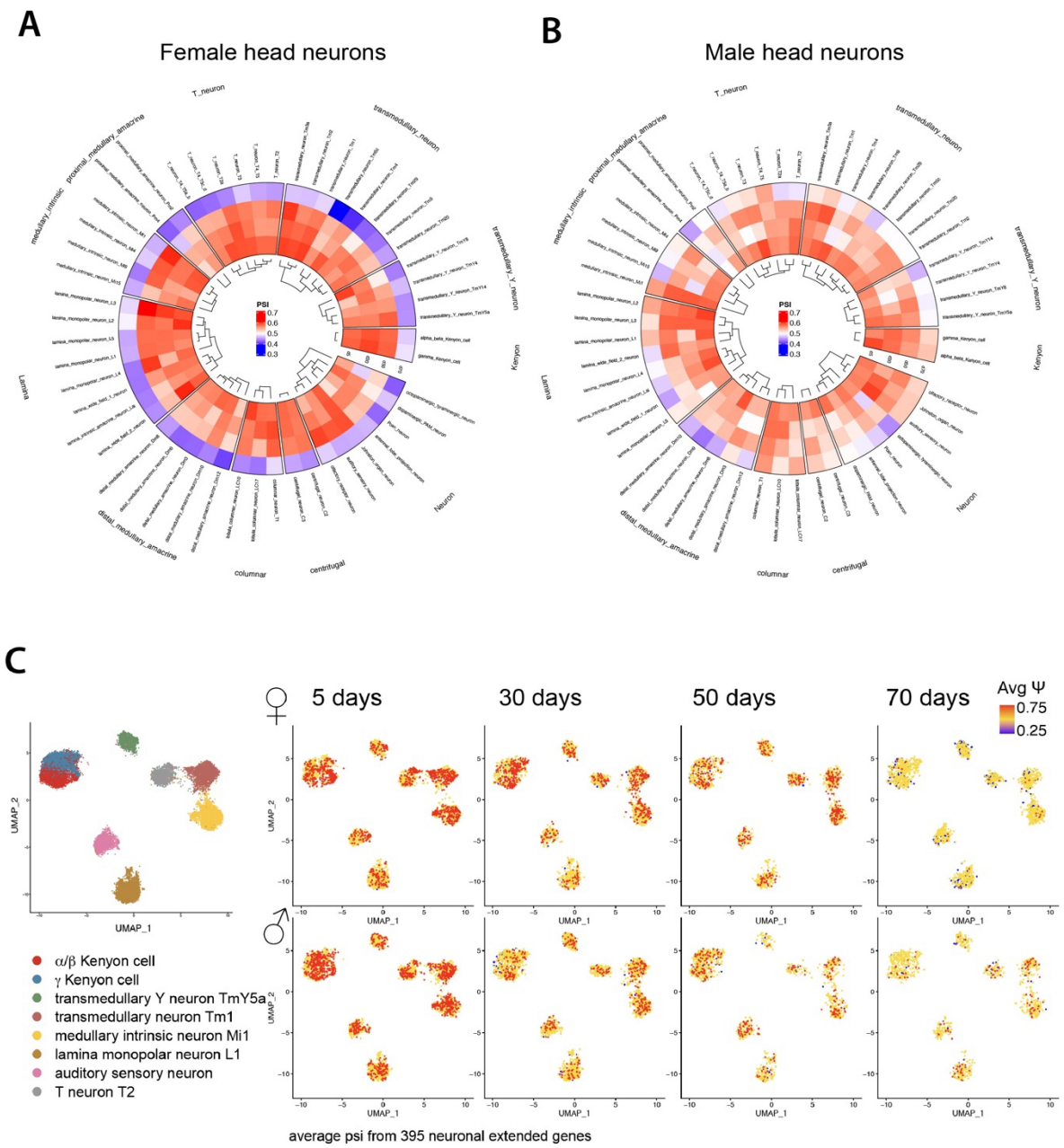

Fig. S29 Comparison of alternative polyadenylation (APA) scores across different ages. (A-B) The circular heatmaps summarize APA trends from the annotated neuronal extended genes across different time points in females (A) and males (B). We use the LABRATsc package to calculate ψ and average value for 391 neuronal extended genes, which are plotted in each of the neuronal types from the AFCA head data. (C) UMAP of different neuronal types designated by the AFCA. Average 3' isoform usage of 391 neural extended genes is plotted in individual cells by different aging time-points. Both analyses show that neuronal extended 3' isoforms are progressively depleted during aging. This phenotype is stronger in females at 70d.

Movies S1-S3. 3D reconstructions of two attached nuclei in the same fat body cell. Attached nuclei are stained by DAPI and LamC. The membranes of fat body cells are labeled by *cg-GAL4 > UAS-CD8GFP*.

Reference:

1. Anisimova AS, Alexandrov AI, Makarova NE, Gladyshev VN, Dmitriev SE. 2018. Protein synthesis and quality control in aging. Aging (Albany NY) 10:4269–4288. doi:10.18632/aging.101721
2. Arrese EL, Soulages JL. 2010. Insect fat body: energy, metabolism, and regulation. Annu Rev Entomol 55:207–225. doi:10.1146/annurev-ento-112408-085356
3. Baldridge D, Wangler MF, Bowman AN, Yamamoto S, Undiagnosed Diseases Network, Schedl T, Pak SC, Postlethwait JH, Shin J, Solnica-Krezel L, Bellen HJ, Westerfield M. 2021. Model organisms contribute to diagnosis and discovery in the undiagnosed diseases network: current state and a future vision. Orphanet J Rare Dis 16:206. doi:10.1186/s13023-021-01839-9
4. Basisty N, Meyer JG, Schilling B. 2018. Protein turnover in aging and longevity. Proteomics 18:e1700108. doi:10.1002/pmic.201700108
5. Bellen HJ, Tong C, Tsuda H. 2010. 100 years of Drosophila research and its impact on vertebrate neuroscience: a history lesson for the future. Nat Rev Neurosci 11:514–522. doi:10.1038/nrn2839
6. Bertram MJ, Akerkar GA, Ard RL, Gonzalez C, Wolfner MF. 1992. Cell type-specific gene expression in the Drosophila melanogaster male accessory gland. Mech Dev 38:33–40. doi:10.1016/0925-4773(92)90036-j
7. Biteau B, Hochmuth CE, Jasper H. 2008. JNK activity in somatic stem cells causes loss of tissue homeostasis in the aging Drosophila gut. Cell Stem Cell 3:442–455. doi:10.1016/j.stem.2008.07.024
8. Buckley MT, Sun E, George BM, Liu L, Schaum N, Xu L, Reyes JM, Goodell MA, Weissman IL, Wyss-Coray T, Rando TA, Brunet A. 2022. Cell type-specific aging clocks to quantify aging and rejuvenation in regenerative regions of the brain. doi:10.1101/2022.01.10.475747
9. Chatterjee N, Perrimon N. 2021. What fuels the fly: Energy metabolism in Drosophila and its application to the study of obesity and diabetes. Sci Adv 7. doi:10.1126/sciadv.abg4336
10. Chen H-Z, Ouseph MM, Li J, Pécot T, Chokshi V, Kent L, Bae S, Byrne M, Duran C, Comstock G, Trikha P, Mair M, Senapati S, Martin CK, Gandhi S, Wilson N, Liu B, Huang Y-W, Thompson JC, Raman S, Leone G. 2012. Canonical and atypical E2Fs regulate the mammalian endocycle. Nat Cell Biol 14:1192–1202. doi:10.1038/ncb2595
11. Chen H, Zheng X, Zheng Y. 2014. Age-associated loss of lamin-B leads to systemic inflammation and gut hyperplasia. Cell 159:829–843. doi:10.1016/j.cell.2014.10.028
12. Davie K, Janssens J, Koldere D, De Waegeneer M, Pech U, Kreft Ł, Aibar S, Makhzami S, Christiaens V, Bravo González-Blas C, Poovathingal S, Hulselmans G, Spanier KI, Moerman T, Vanspauwen B, Geurs S, Voet T, Lammertyn J, Thienpont B, Liu S, Aerts S. 2018. A Single-Cell Transcriptome Atlas of the Aging Drosophila Brain. Cell 174:982-998.e20. doi:10.1016/j.cell.2018.05.057
13. Doherty TJ. 2003. Invited review: Aging and sarcopenia. J Appl Physiol 95:1717–1727. doi:10.1152/japplphysiol.00347.2003
14. Faure L, Soldatov R, Kharchenko PV, Adameyko I. 2022. scFates: a scalable python package for advanced pseudotime and bifurcation analysis from single cell data. Bioinformatics. doi:10.1093/bioinformatics/btac746
15. Gehrmann T, Pelkmans JF, Ohm RA, Vos AM, Sonnenberg ASM, Baars JJP, Wösten HAB, Reinders MJT, Abeel T. 2018. Nucleus-specific expression in the multinuclear mushroom-forming fungus Agaricus bisporus reveals different nuclear regulatory programs. Proc Natl Acad Sci USA 115:4429–4434. doi:10.1073/pnas.1721381115

16. Ghosh AC, Tattikota SG, Liu Y, Comjean A, Hu Y, Barrera V, Ho Sui SJ, Perrimon N. 2020. Drosophila PDGF/VEGF signaling from muscles to hepatocyte-like cells protects against obesity. eLife 9. doi:10.7554/eLife.56969
17. Gonskikh Y, Polacek N. 2017. Alterations of the translation apparatus during aging and stress response. Mech Ageing Dev 168:30–36. doi:10.1016/j.mad.2017.04.003
18. Grotewiel MS, Martin I, Bhandari P, Cook-Wiens E. 2005. Functional senescence in Drosophila melanogaster. Ageing Res Rev 4:372–397. doi:10.1016/j.arr.2005.04.001
19. Hall H, Medina P, Cooper DA, Escobedo SE, Rounds J, Brennan KJ, Vincent C, Miura P, Doerge R, Weake VM. 2017. Transcriptome profiling of aging Drosophila photoreceptors reveals gene expression trends that correlate with visual senescence. BMC Genomics 18:894. doi:10.1186/s12864-017-4304-3
20. Harman D. 2003. The free radical theory of aging. Antioxid Redox Signal 5:557–561. doi:10.1089/152308603770310202
21. Horvath S, Raj K. 2018. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. Nat Rev Genet 19:371–384. doi:10.1038/s41576-018-0004-3
22. Izgi H, Han D, Isildak U, Huang S, Kocabiyik E, Khaitovich P, Somel M, Dönertaş HM. 2022. Inter-tissue convergence of gene expression during ageing suggests age-related loss of tissue and cellular identity. eLife 11. doi:10.7554/eLife.68048
23. Jacomy M, Venturini T, Heymann S, Bastian M. 2014. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. PLoS ONE 9:e98679. doi:10.1371/journal.pone.0098679
24. Janssens J, Aibar S, Taskiran II, Ismail JN, Gomez AE, Aughey G, Spanier KI, De Rop FV, González-Blas CB, Dionne M, Grimes K, Quan XJ, Papasokrati D, Hulselmans G, Makhzami S, De Waegeneer M, Christiaens V, Southall T, Aerts S. 2022. Decoding gene regulation in the fly brain. Nature 601:630–636. doi:10.1038/s41586-021-04262-z
25. Jasper H. 2020. Intestinal stem cell aging: origins and interventions. Annu Rev Physiol 82:203–226. doi:10.1146/annurev-physiol-021119-034359
26. Jeon H-J, Kim Y-S, Park J-S, Pyo J-H, Na H-J, Kim I-J, Kim C-M, Chung HY, Kim ND, Arking R, Yoo M-A. 2015. Age-related change in γH2AX of Drosophila muscle: its significance as a marker for muscle damage and longevity. Biogerontology 16:503–516. doi:10.1007/s10522-015-9573-0
27. Kalmar T, Lim C, Hayward P, Muñoz-Descalzo S, Nichols J, Garcia-Ojalvo J, Martinez Arias A. 2009. Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells. PLoS Biol 7:e1000149. doi:10.1371/journal.pbio.1000149
28. Klopfenstein DV, Zhang L, Pedersen BS, Ramírez F, Warwick Vesztrocy A, Naldi A, Mungall CJ, Yunes JM, Botvinnik O, Weigel M, Dampier W, Dessimoz C, Flick P, Tang H. 2018. GOATOOLS: A Python library for Gene Ontology analyses. Sci Rep 8:10872. doi:10.1038/s41598-018-28948-z
29. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, Baglaenko Y, Brenner M, Loh P-R, Raychaudhuri S. 2019. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 16:1289–1296. doi:10.1038/s41592-019-0619-0
30. Kudryavtsev BN, Kudryavtseva MV, Sakuta GA, Stein GI. 1993. Human hepatocyte polyploidization kinetics in the course of life cycle. Virchows Arch, B, Cell Pathol 64:387–393. doi:10.1007/BF02915139
31. La Manno G, Siletti K, Furlan A, Gyllborg D, Vinsland E, Mossi Albiach A, Mattsson Langseth C, Khven I, Lederer AR, Dratva LM, Johnsson A, Nilsson M, Lönnerberg P, Linnarsson S. 2021. Molecular architecture of the developing mouse brain. Nature 596:92–96. doi:10.1038/s41586-021-03775-x
32. Lee S, Chen Y-C, FCA Consortium, Gillen AE, Taliaferro JM, Deplancke B, Li H, Lai EC. 2022. Diverse cell-specific patterns of alternative polyadenylation in Drosophila. Nat Commun 13:5372. doi:10.1038/s41467-022-32305-0

33. Li H, Janssens J, De Waegeneer M, Kolluru SS, Davie K, Gardeux V, Saelens W, David FPA, Brbić M, Spanier K, Leskovec J, McLaughlin CN, Xie Q, Jones RC, Brueckner K, Shim J, Tattikota SG, Schnorrer F, Rust K, Nystul TG, Zinzen RP. 2022. Fly Cell Atlas: A single-nucleus transcriptomic atlas of the adult fruit fly. Science 375:eabk2432. doi:10.1126/science.abk2432
34. Li H, Jasper H. 2016. Gastrointestinal stem cells in health and disease: from flies to humans. Dis Model Mech 9:487–499. doi:10.1242/dmm.024232
35. Li H, Qi Y, Jasper H. 2016. Preventing Age-Related Decline of Gut Compartmentalization Limits Microbiota Dysbiosis and Extends Lifespan. Cell Host Microbe 19:240–253. doi:10.1016/j.chom.2016.01.008
36. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. 2013. The hallmarks of aging. Cell 153:1194–1217. doi:10.1016/j.cell.2013.05.039
37. Losick VP, Fox DT, Spradling AC. 2013. Polyploidization and cell fusion contribute to wound healing in the adult Drosophila epithelium. Curr Biol 23:2224–2232. doi:10.1016/j.cub.2013.09.029
38. Lucchetta EM, Ohlstein B. 2017. Amitosis of polyploid cells regenerates functional stem cells in the drosophila intestine. Cell Stem Cell 20:609-620.e6. doi:10.1016/j.stem.2017.02.012
39. McLaughlin CN, Qi Y, Quake SR, Luo L, Li H. 2022. Isolation and RNA sequencing of single nuclei from Drosophila tissues. STAR Protocols 3:101417. doi:10.1016/j.xpro.2022.101417
40. Micchelli CA, Perrimon N. 2006. Evidence that stem cells reside in the adult Drosophila midgut epithelium. Nature 439:475–479. doi:10.1038/nature04371
41. Niccoli T, Partridge L. 2012. Ageing as a risk factor for disease. Curr Biol 22:R741-52. doi:10.1016/j.cub.2012.07.024
42. Ohlstein B, Spradling A. 2006. The adult Drosophila posterior midgut is maintained by pluripotent stem cells. Nature 439:470–474. doi:10.1038/nature04333
43. Regan JC, Khericha M, Dobson AJ, Bolukbasi E, Rattanavirotkul N, Partridge L. 2016. Sex difference in pathology of the ageing gut mediates the greater response of female lifespan to dietary restriction. eLife 5:e10956. doi:10.7554/eLife.10956
44. Rezaei A, Krishna MS, Santhosh HT. 2015. Male Age Affects Female Mate Preference, Quantity of Accessory Gland Proteins, and Sperm Traits and Female Fitness in D. melanogaster. Zool Sci 32:16–24. doi:10.2108/zs140121
45. Rodriguez-Fernandez IA, Tauc HM, Jasper H. 2020. Hallmarks of aging Drosophila intestinal stem cells. Mech Ageing Dev 190:111285. doi:10.1016/j.mad.2020.111285
46. Roux AE, Yuan H, Podshivalova K, Hendrickson D, Kerr R, Kenyon C, Kelley DR. 2022. The complete cell atlas of an aging multicellular organism. doi:10.1101/2022.06.15.496201
47. Ruhmann H, Koppik M, Wolfner MF, Fricke C. 2018. The impact of ageing on male reproductive success in Drosophila melanogaster. Exp Gerontol 103:1–10. doi:10.1016/j.exger.2017.12.013
48. Supek F, Bošnjak M, Škunca N, Šmuc T. 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS ONE 6:e21800. doi:10.1371/journal.pone.0021800
49. Sziraki A, Lu Z, Lee J, Banyai G, Anderson S, Abdulraouf A, Metzner E, Liao A, Epstein A, Xu Z, Zhang Z, Gan L, Nelson PT, Zhou W, Cao J. 2022. A global view of aging and Alzheimer's pathogenesis-associated cell population dynamics and molecular signatures in the human and mouse brains. doi:10.1101/2022.09.28.509825
50. Tabula Muris Consortium. 2020. A single-cell transcriptomic atlas characterizes ageing tissues in the mouse. Nature 583:590–595. doi:10.1038/s41586-020-2496-1

51. Unhavaithaya Y, Orr-Weaver TL. 2012. Polyploidization of glia in neural development links tissue growth to blood-brain barrier integrity. Genes Dev 26:31–36. doi:10.1101/gad.177436.111
52. von Aderkas P, Rouault G, Wagner R, Chiwocha S, Roques A. 2005. Multinucleate storage cells in Douglas-fir (Pseudotsuga menziesii (Mirbel) Franco) and the effect of seed parasitism by the chalcid Megastigmus spermotrophus Wachtl. Heredity 94:616–622. doi:10.1038/sj.hdy.6800670
53. Wang R, Zhang P, Wang J, Ma L, E W, Suo S, Jiang M, Li J, Chen H, Sun H, Fei L, Zhou Z, Zhou Y, Chen Y, Zhang W, Wang X, Mei Y, Sun Z, Yu C, Shao J, Han X. 2022. Construction of a cross-species cell landscape at single-cell level. Nucleic Acids Res. doi:10.1093/nar/gkac633
54. Wolf FA, Angerer P, Theis FJ. 2018. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol 19:15. doi:10.1186/s13059-017-1382-0
55. Wolf FA, Hamey FK, Plass M, Solana J, Dahlin JS, Göttgens B, Rajewsky N, Simon L, Theis FJ. 2019. PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. Genome Biol 20:59. doi:10.1186/s13059-019-1663-x
56. Zeng H. 2022. What is a cell type and how to define it? Cell 185:2739–2755. doi:10.1016/j.cell.2022.06.031
57. Zhao X, Karpac J. 2021. Glutamate metabolism directs energetic trade-offs to shape host-pathogen susceptibility in Drosophila. Cell Metab 33:2428-2444.e8. doi:10.1016/j.cmet.2021.10.003