Science Advances

AAAS

Supplementary Materials for

Social reprogramming in ants induces longevity-associated glia remodeling

Lihong Sheng, Emily J. Shields, Janko Gospocic, Karl M. Glastad, Puttachai Ratchasanmuang, Shelley L. Berger, Arjun Raj, Shawn Little, Roberto Bonasio*

*Corresponding author. Email: roberto@bonasiolab.org

The PDF file includes:

Figs. S1 to S8

Other Supplementary Material for this manuscript includes the following:

(available at advances.sciencemag.org/cgi/content/full/6/34/eaba9869/DC1)
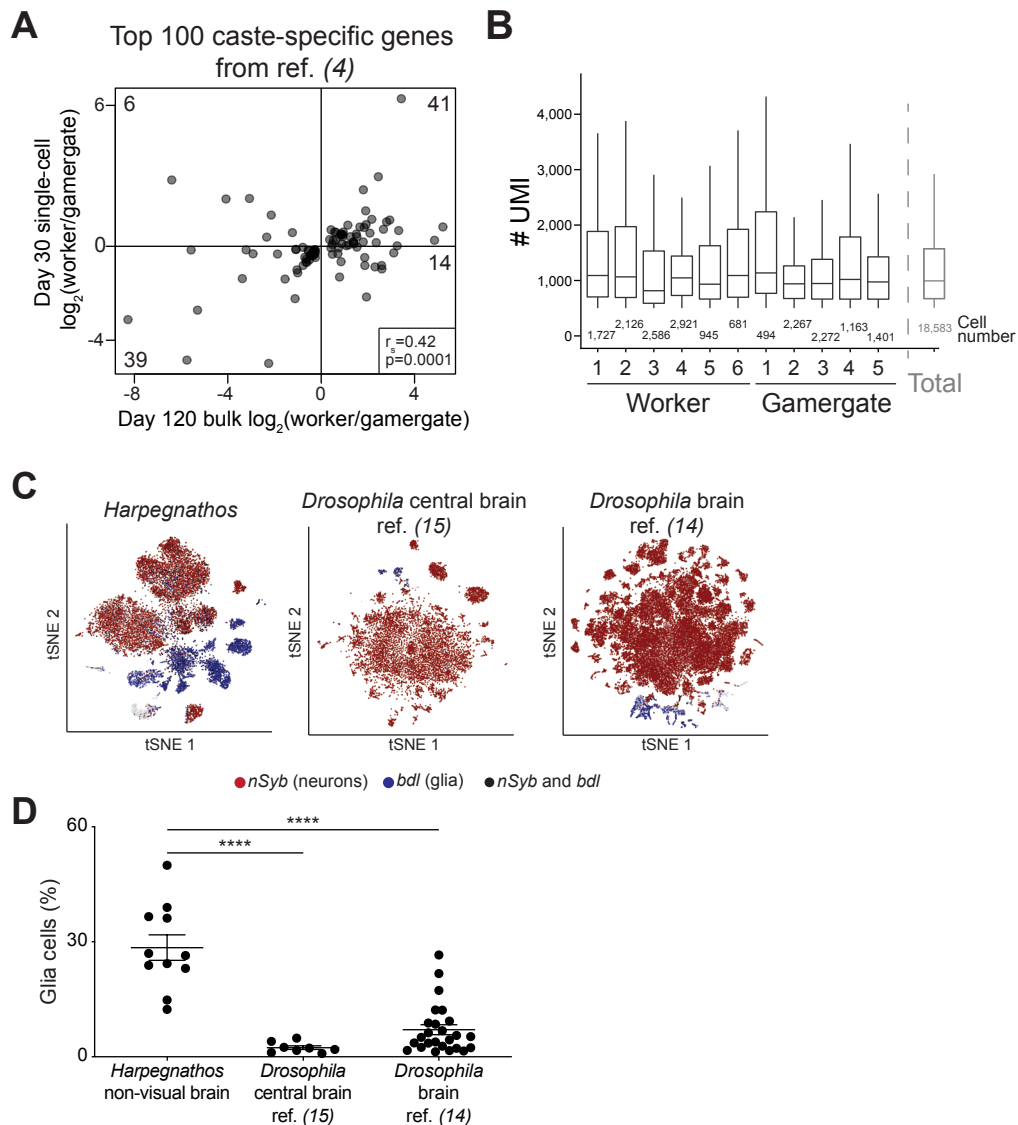
Tables S1 to S6

Figure S1. Single-cell dataset features; frequency of neurons and glia

(A) Scatter plot showing the correlation between differential expression of caste-specific genes at day 120 and day 30. The genes show are the top 100 most differentially expressed genes (by *P*-value) in brains from *Harpegnathos* workers *vs.* gamergates as identified in our previous bulk RNA-seq at day 120 *(4)*. The x-axis shows the log-converted fold-change in bulk RNA-seq from gamergate vs. worker brains at day 120 *(4)*; the y-axis shows the log-converted fold-change between gamergate and worker brains in the 10x Genomics single-cell RNA-seq dataset obtained at day 30 for this study. Numbers of genes in each quadrant are shown.

(B) Boxplots showing the distribution of UMIs in all cells included in clustering for the separate replicates (black) and the pooled data (gray) used for clustering and tSNE visualization in Fig. 1. The number of cells per sample are indicated.

(C) Expression patterns plotted over tSNE for *nSyb* (neurons, red) and *bdl* (glia, blue) expression in cells from the current study in *Harpegnathos* (left), the *Drosophila* central brain (middle; ref. *(15)*), and the *Drosophila* whole brain (right; ref. *(14)*). Cells with expression of both markers are depicted in black.

(D) Relative frequency (as % of total cells) glia cells in *Harpegnathos*, *Drosophila* central brain *(15)*, and *Drosophila* whole brain *(14)*. Horizontal bars indicate means ± SEM. Each point represents one biological replicate. ****, P < 0.0001 from one-way ANOVA and Holm-Sidak test.
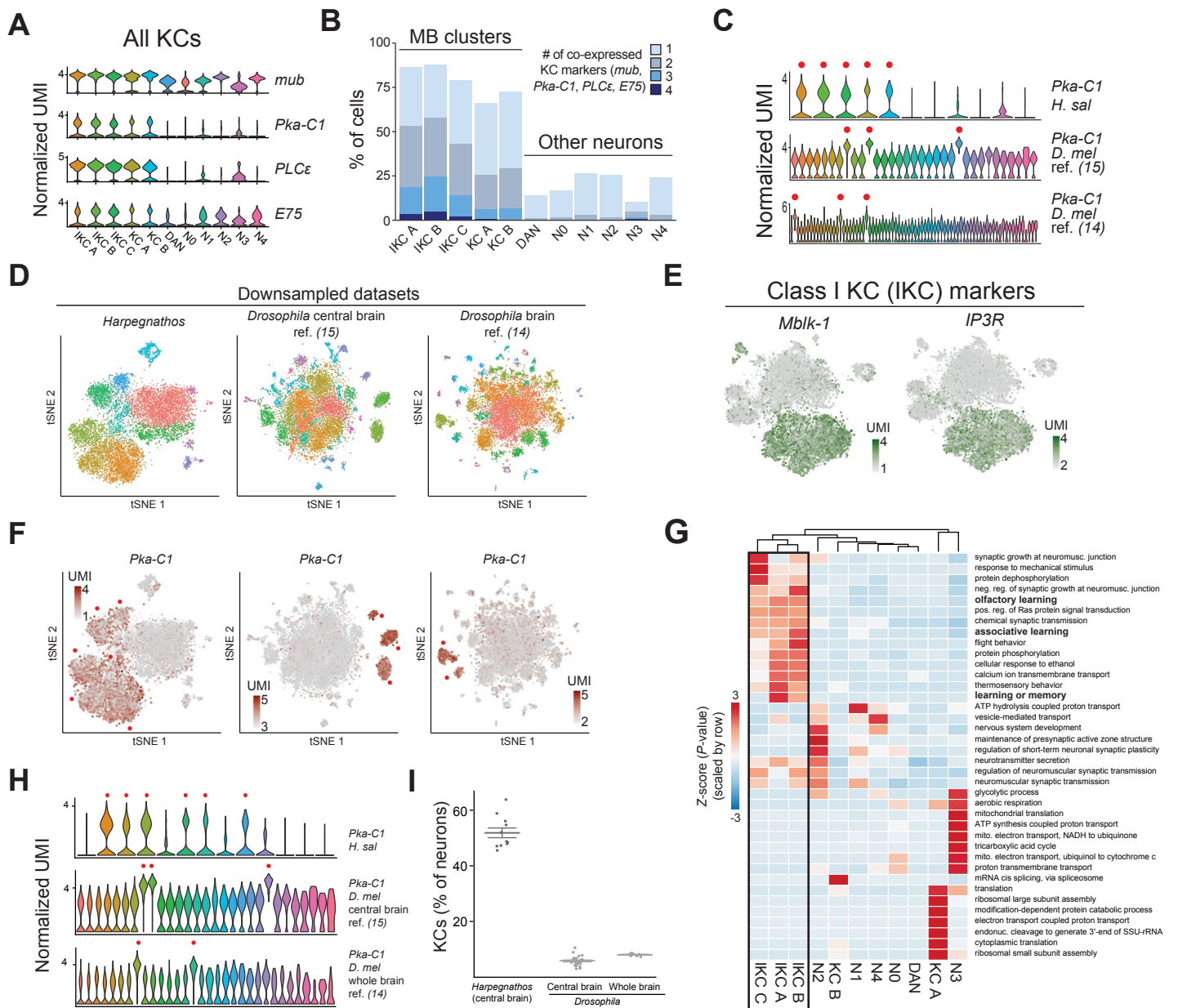
Figure S2. Mushroom body neurons in *Harpegnathos* and *Drosophila*

(A) Violin plots showing expression (normalized UMIs) of mushroom body and KC markers in all neuronal clusters.

(B) Bar plot showing the co-expression at the single-cell level of the four KC markers from (A) in the different neuronal clusters. Expression was defined as one standard deviation above the mean across all clusters.

(C) Violin plots showing expression (normalized UMIs) for *Pka-C1* in *Harpegnathos* and *Drosophila* central brain and whole brain for all neuronal clusters.

(D) Visualization of the indicated datasets by tSNE after downsampleing to the same number of cells and UMI/cell distribution.

(E) Heatmap plotted over neuronal tSNE for two IKC markers previously described in the honey bee.

(F) Heatmap plotted over downsampled tSNE showing the expression (normalized UMIs) of *Pka-C1* in *Harpegnathos* and *Drosophila* neurons.

(G) Clustered heatmap of the *P*-values for the enrichment of GO terms (rows) associated with genes specifically expressed in each cluster (columns). The heatmap contains all biological process GO terms with an adjusted *P*-value < 0.1 for at least one cluster. The color scale represents $-\log_{10}(P\text{-value})$ scaled by row. Terms enriched in genes expressed specifically in the IKC A, B, and C clusters are indicated with a black box. Terms mentioned in the text are bolded.

(H) Violin plots showing expression (normalized UMIs) for *Pka-C1* in *Harpegnathos* and *Drosophila* central brain and whole brain for all neuronal clusters after downsampling. Red dots indicate clusters with elevated expression included in the quantifications.

(I) Relative frequency of KCs as determined by percentage of neurons in clusters that express *Pka-C1* in *Harpegnathos* brains and in the two *Drosophila* single-cell RNA-seq datasets after downsampling. Horizontal bars indicate mean ± SEM.
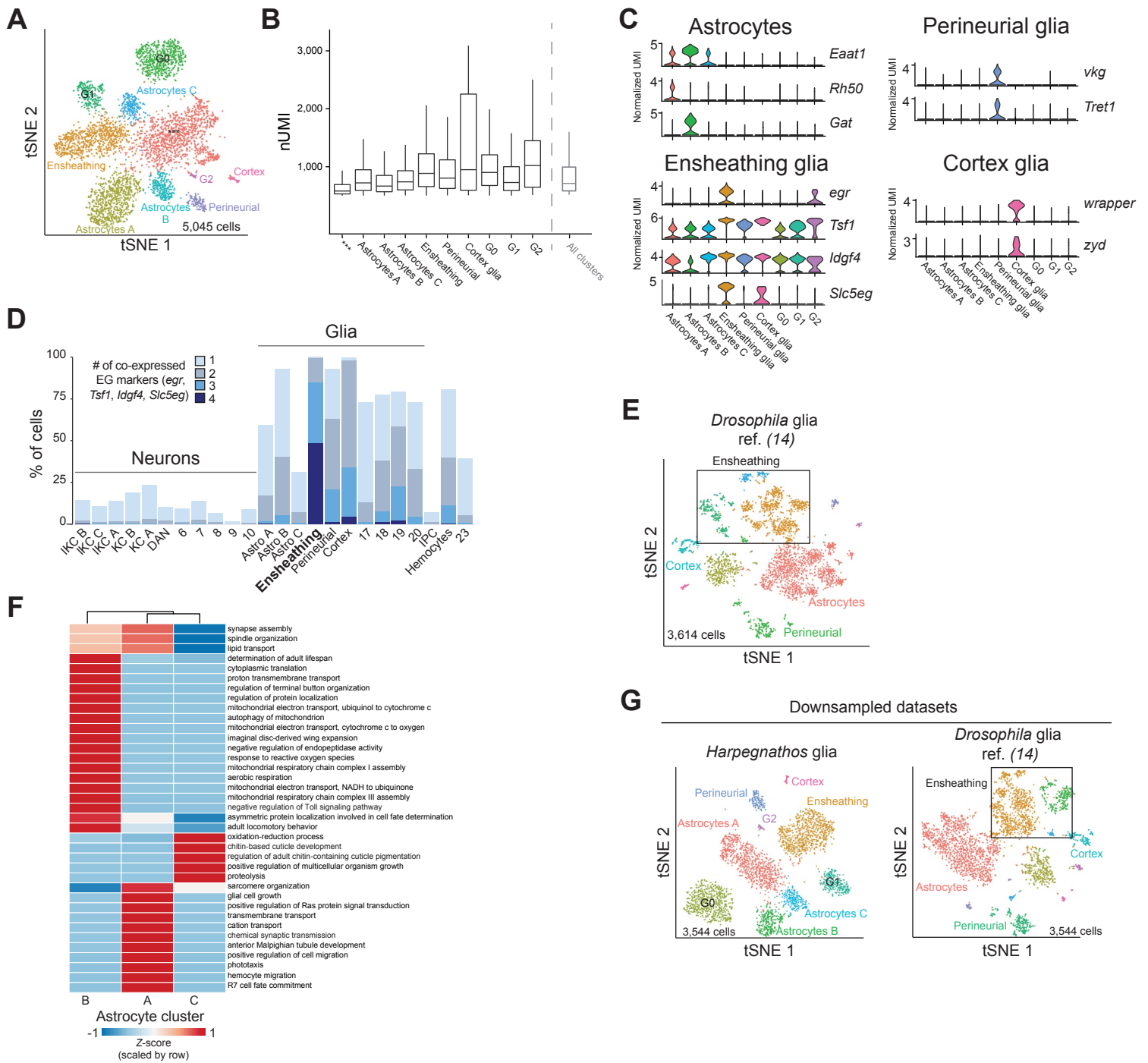
Figure S3. Glia in *Harpegnathos* and *Drosophila*

(A) Annotated tSNE visualization of glia-only reclustering including the low quality cluster, labeled "***".

(B) Boxplot showing the distribution of UMIs in all clusters from (A) including the low quality ("***") cluster.

(C) Violin plots showing expression (normalized UMIs) in all glia clusters for marker genes of the indicated glia subtypes.

(D) Bar plot showing the co-expression at the single-cell level of the four ensheathing glia markers from (C) in all clusters. Expression was defined as one standard deviation above the mean across all clusters.

(E) Annotated tSNE visualization of glia-only reclustering of single-cell transcriptomes from the *Drosophila* brain *(14)*.

(F) Clustered heatmap of the *P*-values for the enrichment of GO terms (rows) associated with genes specifically expressed in each of the three *Harpegnathos* astrocyte clusters (columns). The heatmap contains all biological process GO terms with an adjusted *P*-value < 0.1 for at least one cluster. Colors represent -log$_{10}$(*P*-value) scaled by row.

(G) Annotated tSNE visualization of *Harpegnathos* and *Drosophila (14)* glia after downsampling to obtain the same number of cells and UMI/cell distributions.
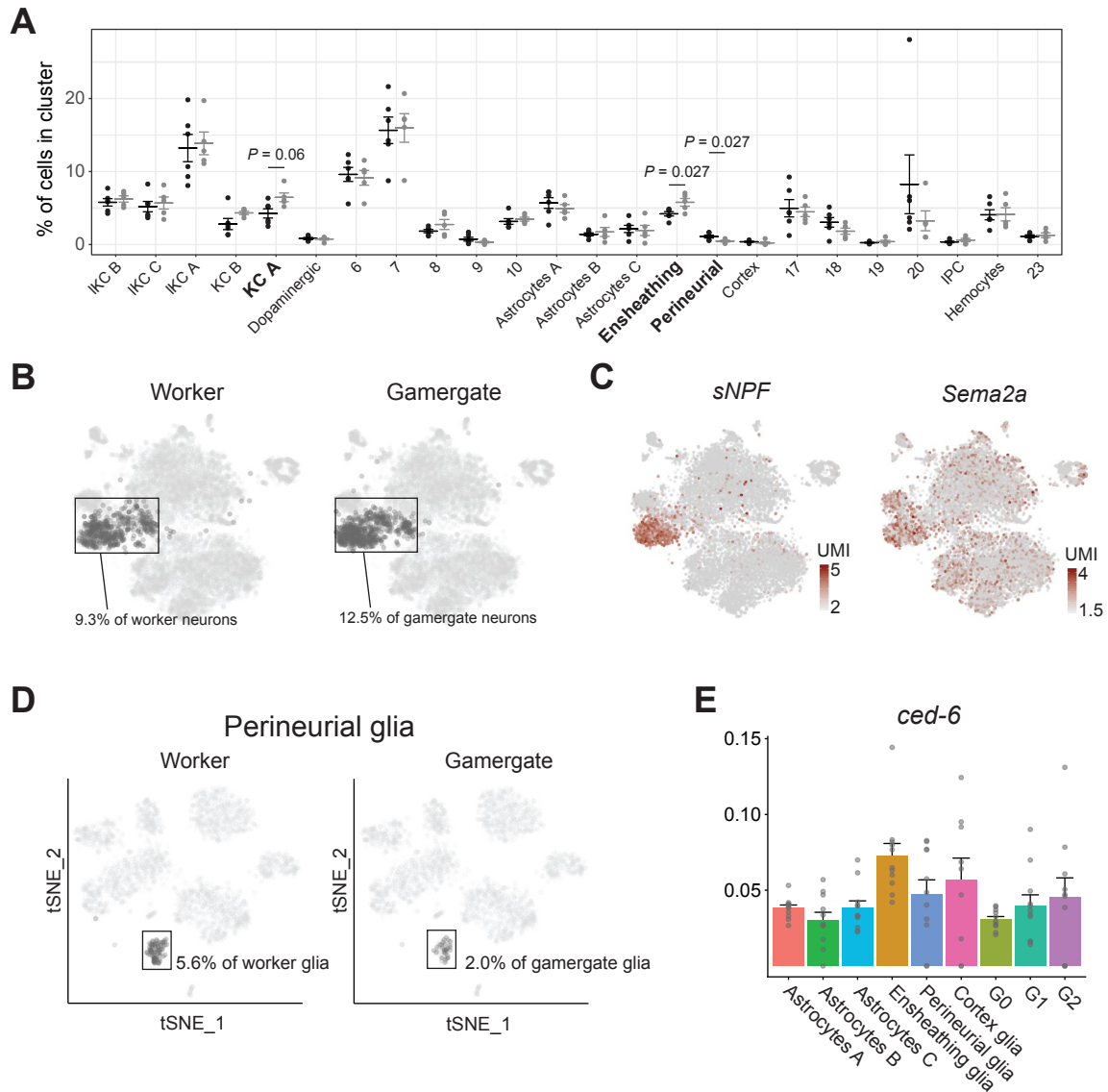
Figure S4. Additional analyses on caste-specific single-cell RNA-seq

(A) Relative frequency (as % of total cells) of cells in all clusters identified in Fig. 1, separated in workers (black circles) and gamergates (gray cyrcles). Each point is a biological replicate. Horizontal bars indicate mean ± SEM. *P*-values are from Wald tests with Benjamini-Hochberg multiple test correction. *P*-values > 1 are not shown but are reported in Table S2.

(B) Visualization and quantification of worker (left) and gamergate (right) contributions to the KC A cluster in the reclustered tSNE for neurons only. Worker and gamergate datasets were downsampled to include the same number of total cells for comparison.

(C) Heatmaps plotted over neuronal tSNE indicating the single-cell expression levels for the indicated genes as normalized UMIs.
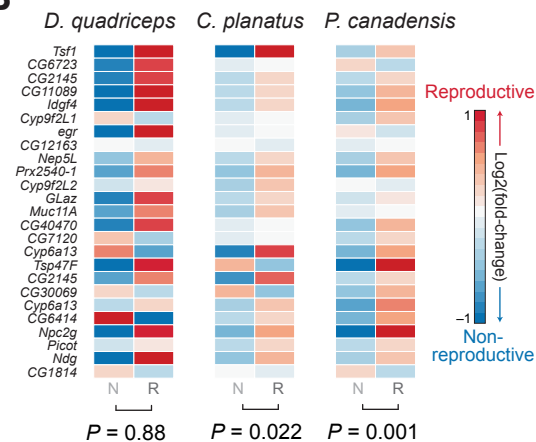
(D) Visualization and quantification of worker (left) and gamergate (right) contributions to the perineurial glia cluster in the reclustered tSNE for glia only. Worker and gamergate datasets were downsampled to include the same number of total cells for comparison.

(E) Expression levels (% of cluster UMIs) across glia subsets for *ced6*. Bars indicate means + SEM.
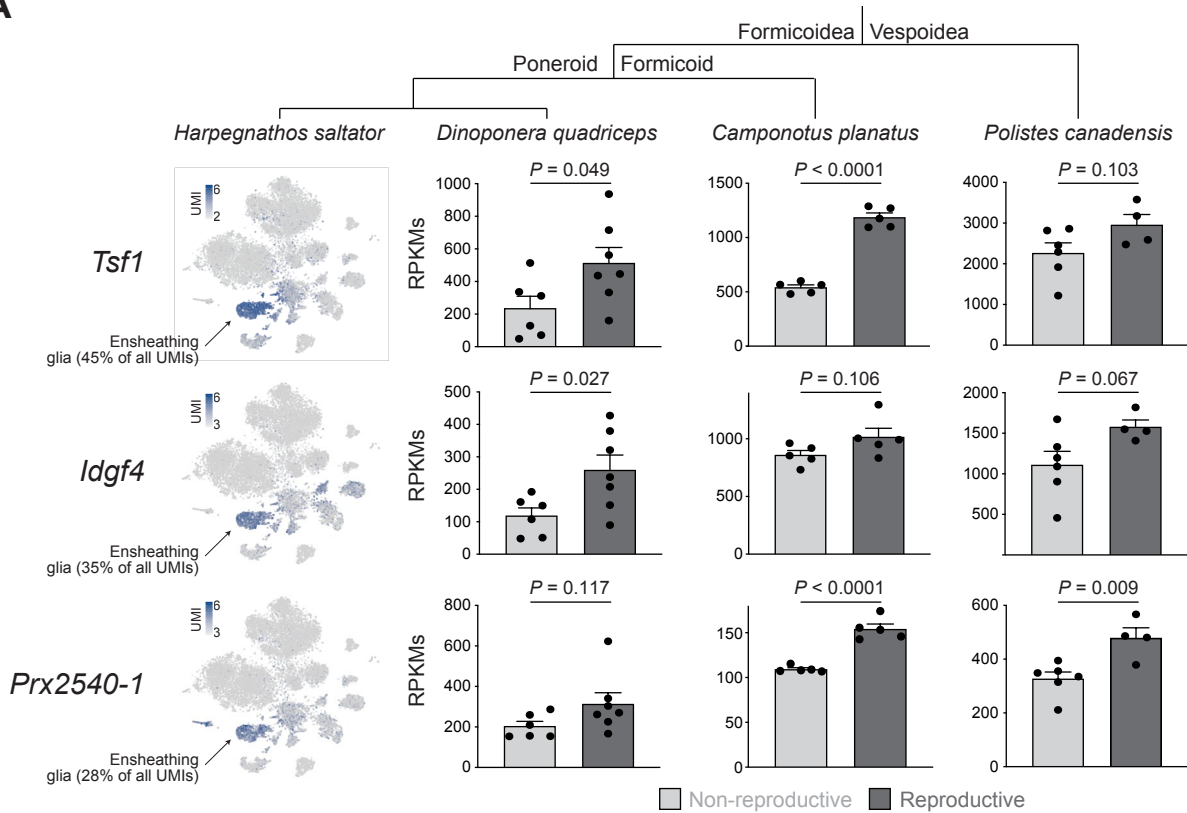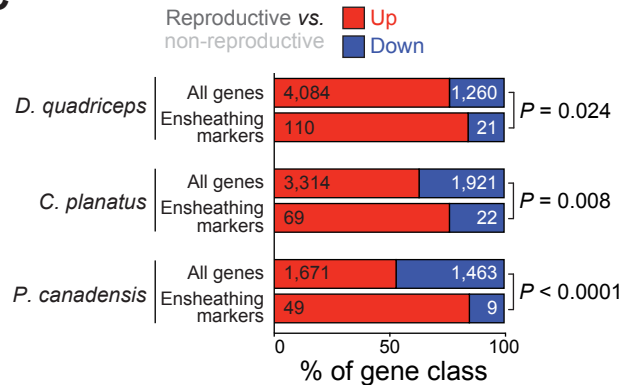
Figure S5. Caste-specific expression of ensheathing glia markers in other Hymenoptera

(A) Expression patterns for three top ensheathing glia markers in the single-cell RNA-seq from *Harpegna-thos* brain (left) and three published RNA-seq datasets from the brains of reproductive and non-reproductive individuals from the ponerine ant *Dinoponera quadriceps*, the ant *Camponotus planatus*, and the red paper wasp *Polistes canadensis (6, 16)*. Expression levels are plotted as reads per kilobase per million (RPKMs). Bars show means + SEM. *P*-values are from Student's *t*-tests.

(B) Heatmap of relative expression changes in the same RNA-seq datasets utilized in (A) for the homologs of the top 25 marker genes for ensheathing glia in *Harpegnathos*. Data and is expressed as the relative $\log_2$(fold-change) in RPKMs for reproductive (left) and non-reproductive (right) compared to the other caste. Genes with higher expression in brains of reproductive individuals are in red and genes with higher expression in brains of non-reproductive individuals are in blue. *P*-values are from Fisher's tests.

(C) Bars show the ratio between the number of genes upregulated or downregulated (by at least 25%) in reproductive vs. non-reproductive individuals when considering all genes (gray bars) or only the homologs to the 213 genes that mark ensheathing glia in *Harpegnathos* (see Table S2). *P*-values are from Fisher's tests.
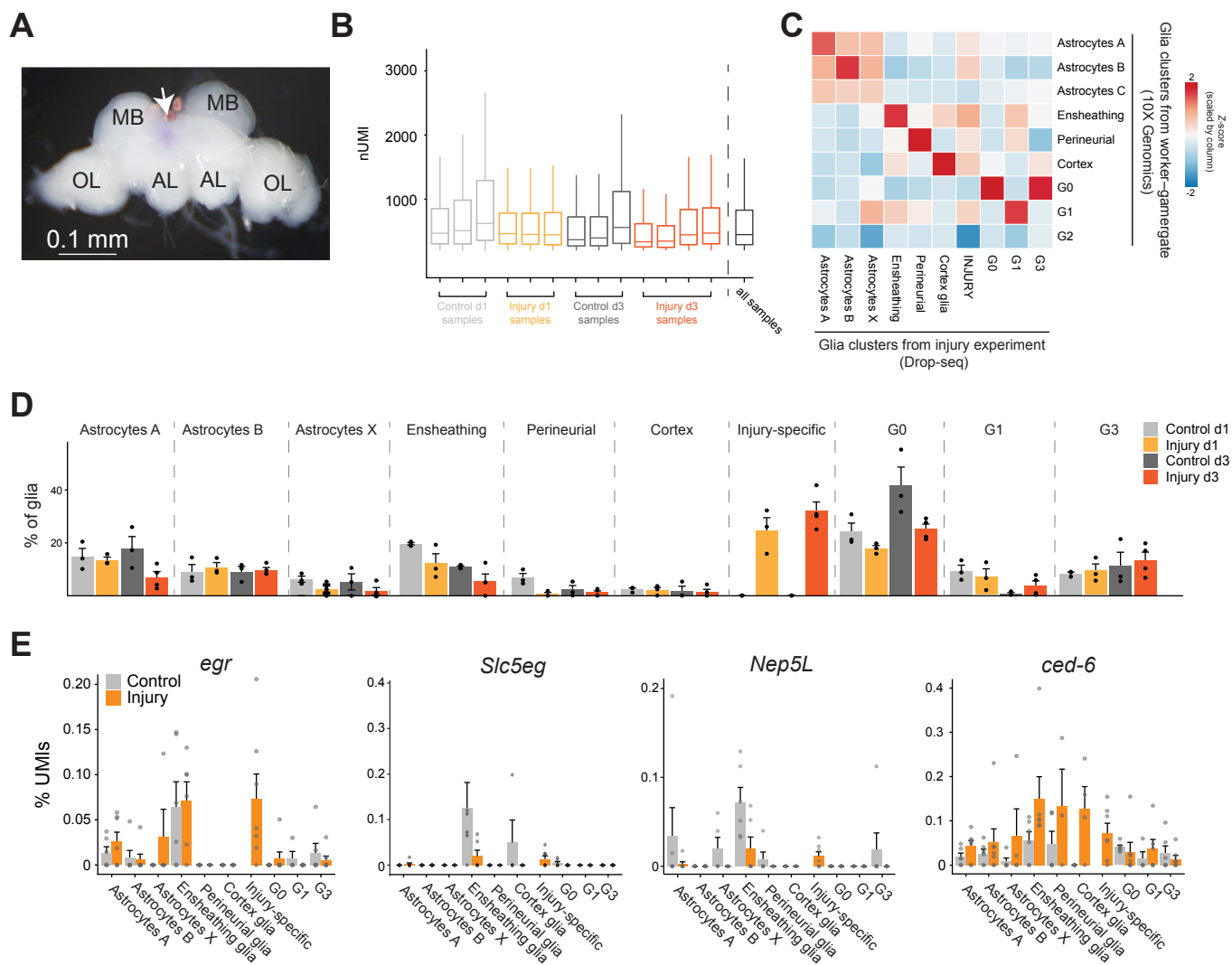
Figure S6. Cellular response to brain injury

(A) Microphotograph of a dissected *Harpegnathos* brain after a stabbing injury delivered by puncturing with a needle. The site of stabbing injury was visualized with a blue dye (arrow). AL, antennal lobe; MB, mushroom body; OL, optic lobe. Photo Credit: Lihong Sheng, University of Pennsylvania.

(B) Boxplots showing the distribution of UMIs in all clusters for the injury Drop-seq experiment.

(C) Heatmap showing the *z*-score for the pairwise Pearson correlation between collapsed transcriptomes (pseudo-bulk analysis) of glia clusters from the day 30 worker vs. gamergates comparison (10x Genomics, Fig. 2G) and from the injury experiment (Drop-seq, Fig. 4B), considering only variable genes that were utilized to define the clusters by Seurat.

(D) Relative frequency (as % of all glia) of each cluster in control and injury day 1 and day 3. Bars represent means + SEM.

(E) Expression levels (as % of cluster UMIs) for the *Harpegnathos* ensheathing glia markers *egr*, *Slc5eg*, and *Nep5L,* and for the phagocytosis gene *ced-6*. Bars indicate means + SEM.
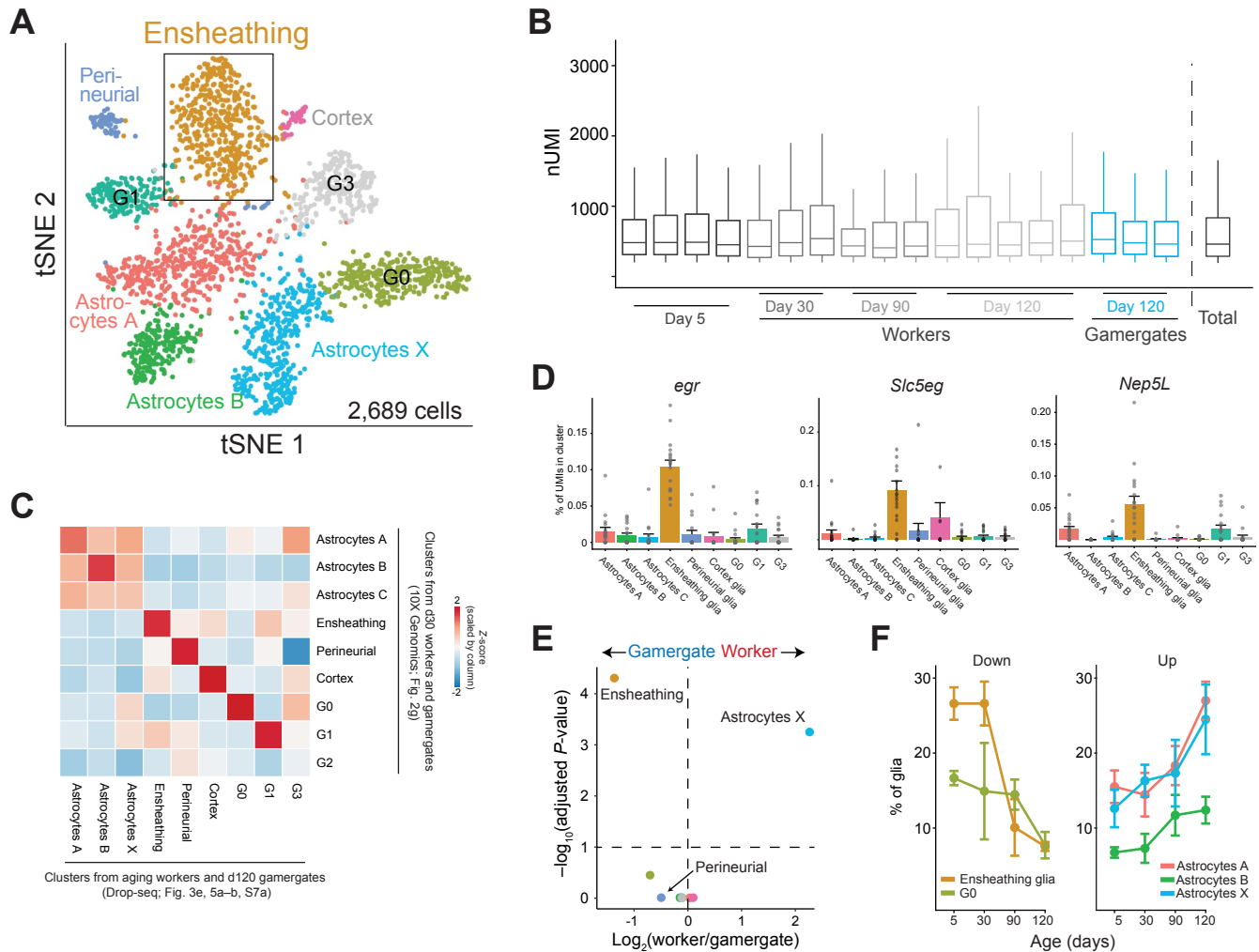
Figure S7. Aging-related changes in cellular composition in *Harpegnathos* brains

(A) Annotated tSNE visualization of glia-only reclustering from Drop-seq performed on brains from workers of different ages (day 5, *n* = 4; day 30, *n* = 3; day 90, *n* = 3, day 120, *n* = 5).The position of ensheathing glia in the tSNE is indicated by the box.

(B) Boxplots showing the distribution of UMIs for the brain Drop-seq datasets from aging workers and day 120 gamergates

(C) Heatmap showing the *z*-score for the pairwise Pearson correlation between collapsed transcriptomes (pseudo-bulk analysis) of glia clusters from the day 30 worker *vs.* gamergates comparison (10x Genomics, Fig. 2G) and from the aging experiment (Drop-seq, Fig. S7A), considering only variable genes that were utilized to define the clusters by Seurat in both datasets (intersection).

(D) Expression levels (% of cluster UMIs) across glia subsets for *egr*, *Slc5eg*, and *Nep5L*. Bars indicate means + SEM.

(E) Volcano plot of $\log_2$(ratio) and $-\log_{10}$(*P*-value from Wald test with Benjiamini-Hochberg multiple test correction) for the relative frequency of each cluster in gamergates *vs.* workers at day 120.

(F) Relative frequencies (as % of total glia) of select glia clusters during worker aging. Circles indicate the means, error bars represent ± SEM.

A *CG6723*

UMI

B

P = 0.013

CG6723 (% of *Rpl32*)

0.02

0.01

0

Young (d5)   Old (d70)

C

All ensheathing glia markers

Young

Old

Z-score

2

−2

SLC22A CG12355 CG31706 CG1552 CG7888 CG7084 hoe1 CG42613 CG6723 CG42235 CG3940 CG4409 CG10513 qgre CG9657 trol Ugt35b Gs2 Rh50 bdl CG9394 CG149445 CG9507 nrv2 CG12177 Sod3 Irk2 CG58995 CG8369 bnb List Inx3 CG31676 CG9686 CG30197 DNaseII Spn43Ab Cyp28d1 MRE16
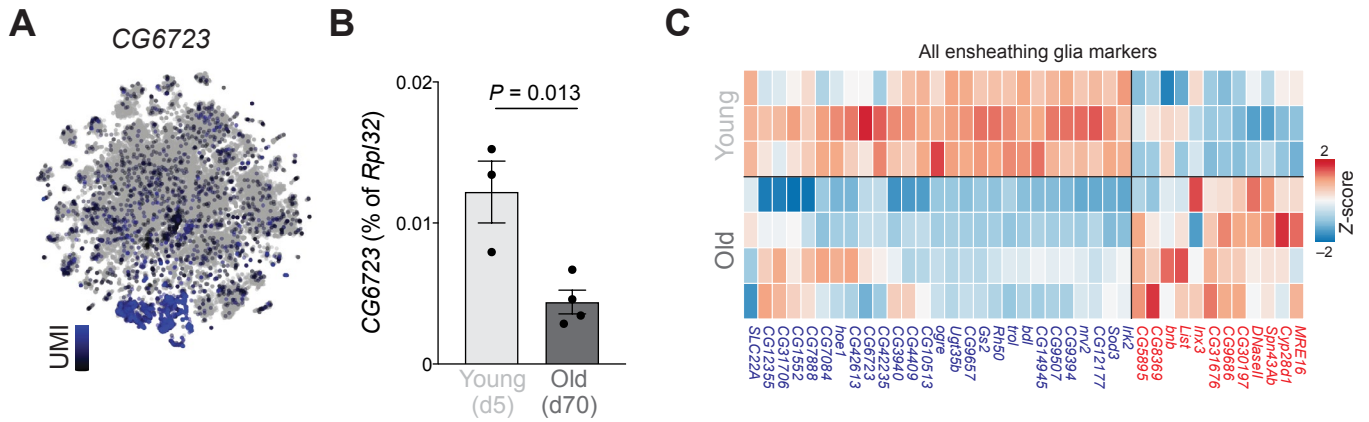
Figure S8. Additional evidence of age-associated decline of ensheathing glia in *Drosophila*

(A) Heatmap of UMI levels per cell plotted over tSNE for *CG6723* in *Drosophila* brains obtained at the SCope website (http://scope.aertslab.org) with single-cell RNA-seq data from ref. *(14)*.

(B) Abundance of *CG6723* mRNA as determined by RT-qPCR (normalized to *Rpl32*) from dissected brains of day 5 (young) and day 70 (old) *Drosophila* females. Bars represent the mean ± SEM. *P*-value is from a Student's *t*-test.

(C) Heatmap of the expression levels (RPKMs) converted to *Z*-scores for all ensheathing glia marker genes in individual RNA-seq replicates from young (day 5) and old (day 70) *Drosophila* brains.