Cell Genomics, Volume *3*

Supplemental information

A multi-omics atlas of the human retina at single-cell resolution

Qingnan Liang, Xuesen Cheng, Jun Wang, Leah Owen, Akbar Shakoor, John L. Lillvis, Charles Zhang, Michael Farkas, Ivana K. Kim, Yumei Li, Margaret DeAngelis, and Rui Chen

Supplemental Information

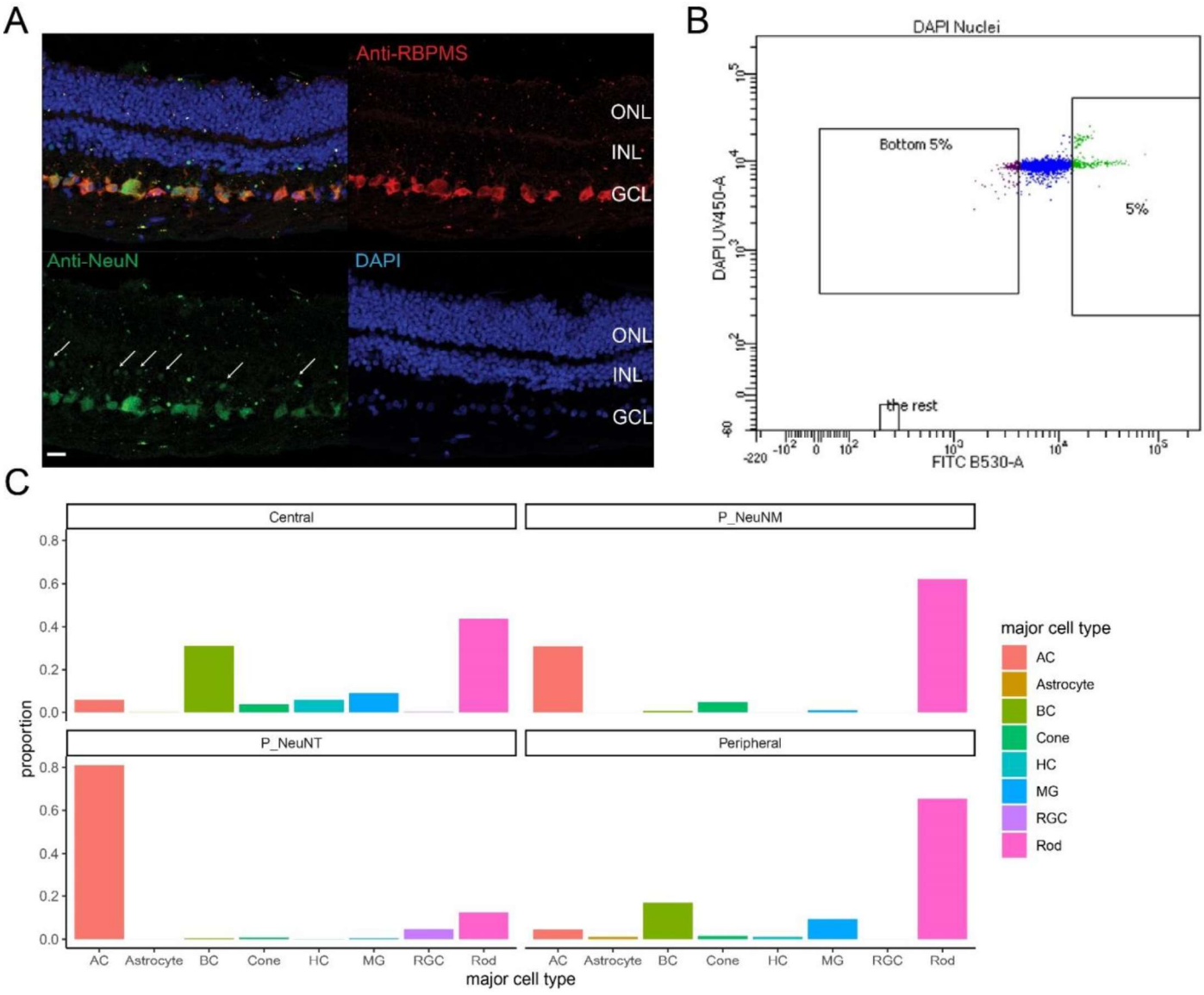

Figure S1, related to Figure 1 and STAR Methods. NeuN enrichment of rare retinal cell classes.

A. NeuN is highly expressed in the RGCs and amacrine cells in the human retina. Immunofluorescent staining with anti-NeuN (green) and anti-RBPMS (red), a pan RGC marker, was performed on human adult retina sections. The top left showed the merged-channel image with NeuN, RBPMS, and DAPI signals. White arrows highlight NeuN-positive cells in the INL, likely being amacrine cells based on their nuclei location. The scale bar (bottom left) indicates 20 microns.

B. Representative FACS sorting plot for separating DAPI positive nuclei with the highest 5% anti-NeuN signal (FITC channel).

C. The proportion of major cell classes in four groups of the snRNA-seq data: central retina, peripheral retina with medium-high NeuN signal (P_NeuM), peripheral retina with top NeuN signal (P_NeuNT), peripheral retina with no enrichment.

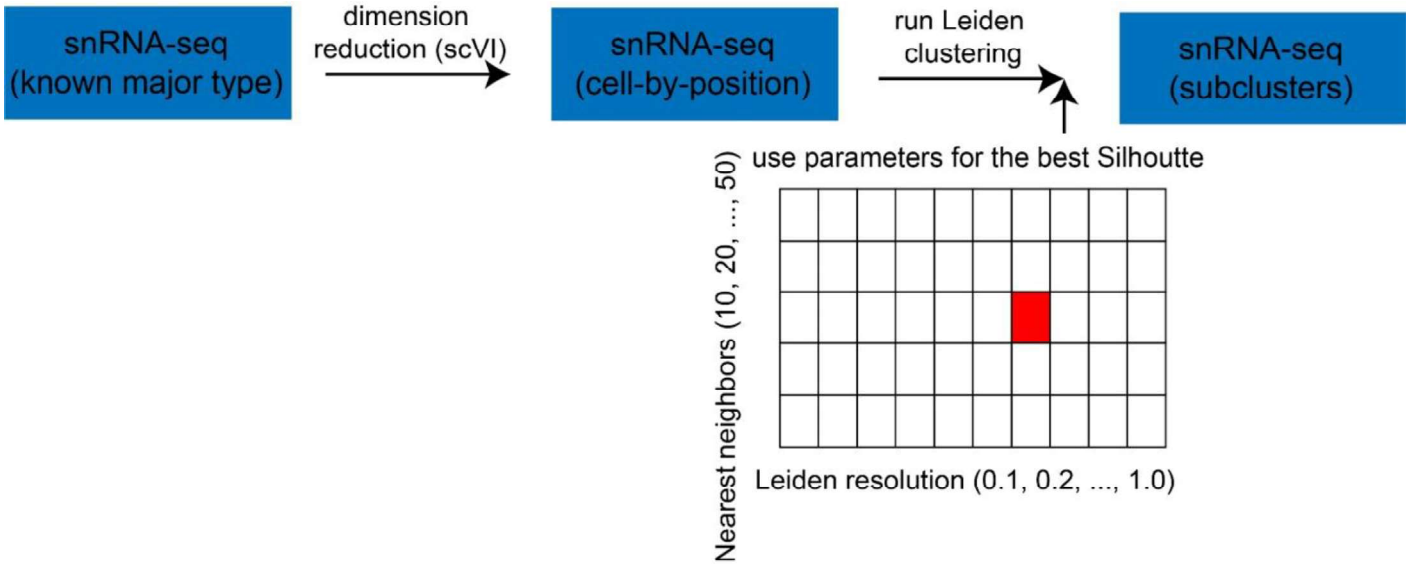

Figure S2, related to STAR methods. The strategy of detecting cell types. For each complex class (BC, AC, HC, RGC, cone), dimension reduction was performed with the scVI framework. In the latent space, Leiden clustering was performed 50 times with a combination of number of nearest neighbors (10, 20, …, 50) and Leiden resolution (0.1, 0.2, …, 1.0) to find the combination giving the best Silhouette score. The clustering result from that very combination was then used as the optimal clustering used for downstream analysis.

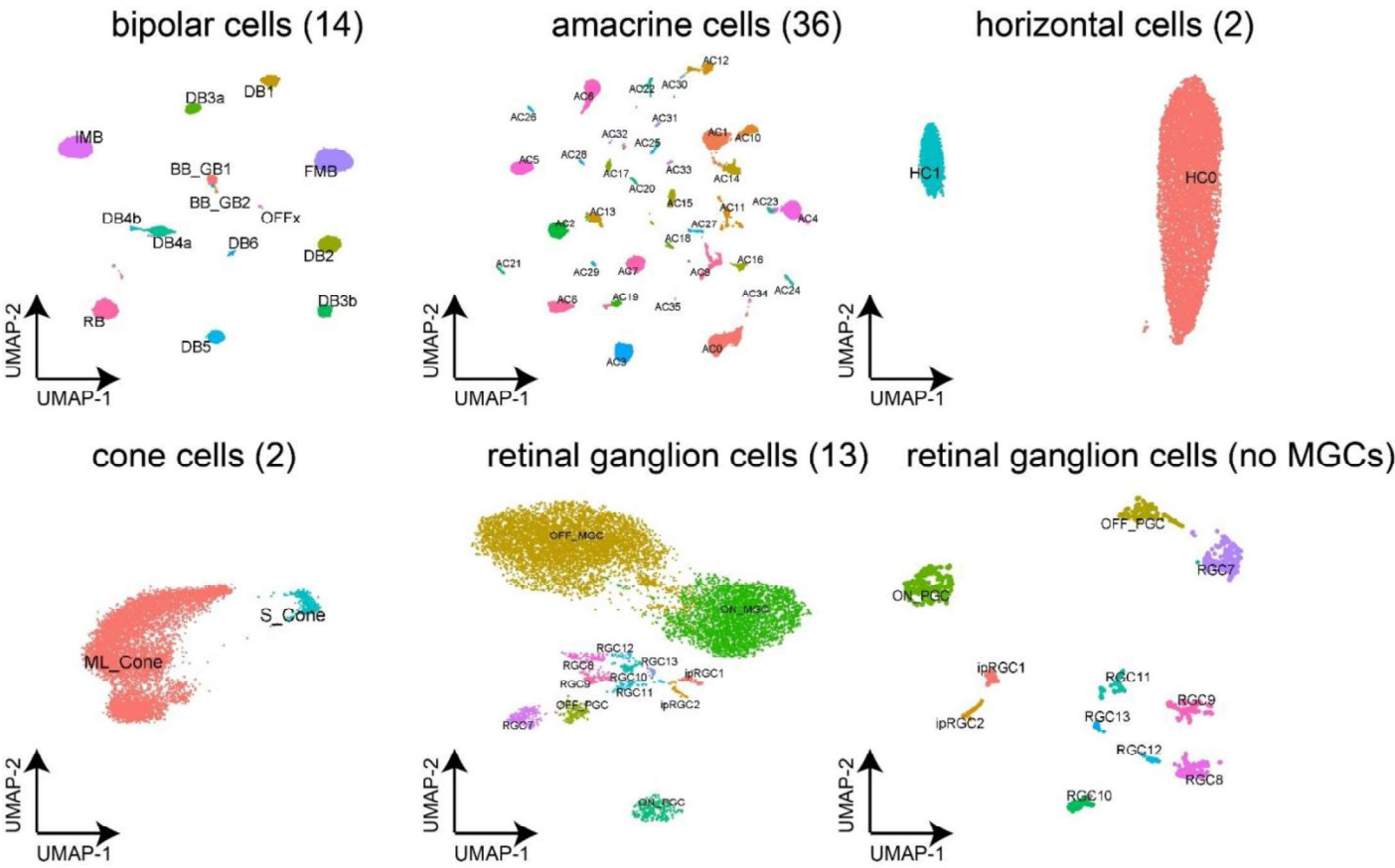

Figure S3, related to Figure 1. Cell types detected from snRNA-seq data. Two-dimensional embeddings of cell types from BCs, ACs, HCs, cones, and RGCs were visualized. For RGCs, given that most of the cell types other than MGCs and ON-PGC looked squeezed, another UMAP embedding without MGCs were also visualized (bottom right).

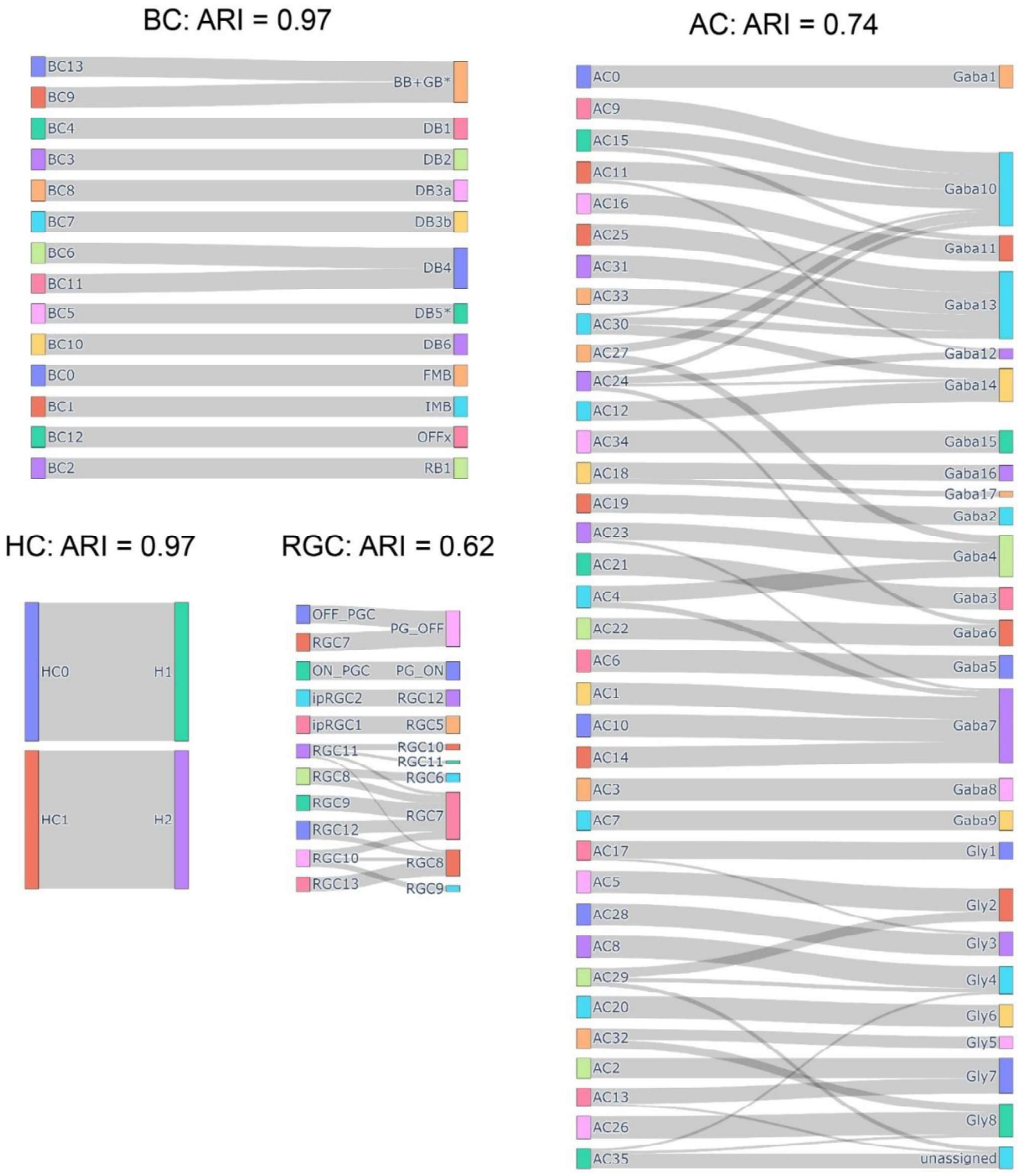

Figure S4, related to Figure 1 and STAR Methods. Comparison of the cell type transcriptome profiles between the current study and the previous report by Yan et al. For each cell type, random forest classifier was trained in the Yan dataset and applied to the current dataset for label

prediction. The portion of each cluster (in the current work) being predicted as a cell type (in the Yan dataset) were then visualized as Sankey diagrams. Adjusted Rand Index (ARI) were computed between the two label sets (current clustering vs label prediction) to evaluate the consistency of the two datasets.

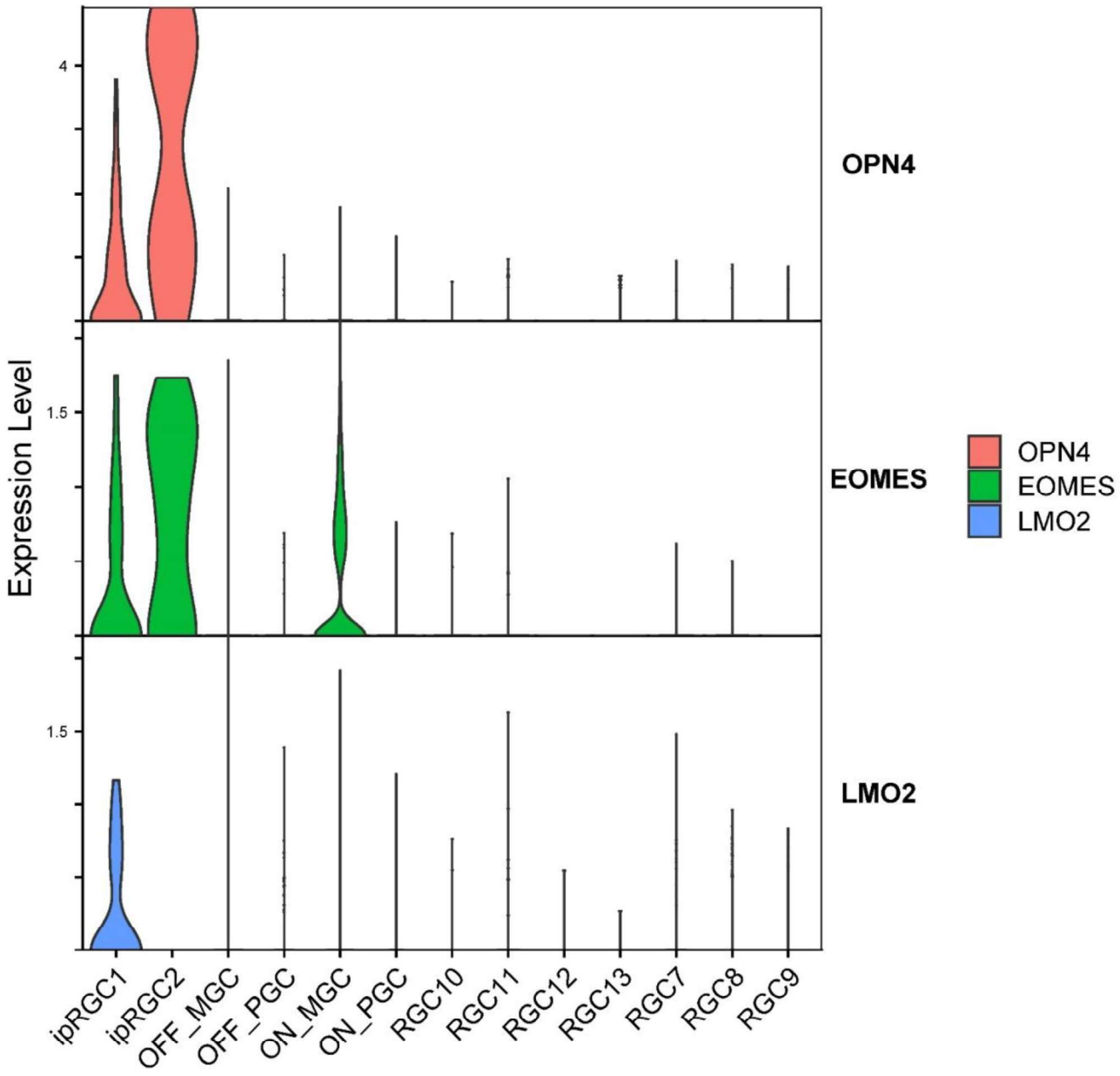

Figure S5, related to Figure 1. The expression of previously reported OPN4 positive RGC markers in the current dataset.

Figure S6, related to Figure 1. Relative cell proportion of AC (A), BC (B), HC (C), and RGC (D) types between central and peripheral retina. The relative cell proportion for each type was calculated within the cell class (e.g., AC, BC, etc.) and within the region (e.g., central), and were visualized by stacking the relative proportion for each cell type.

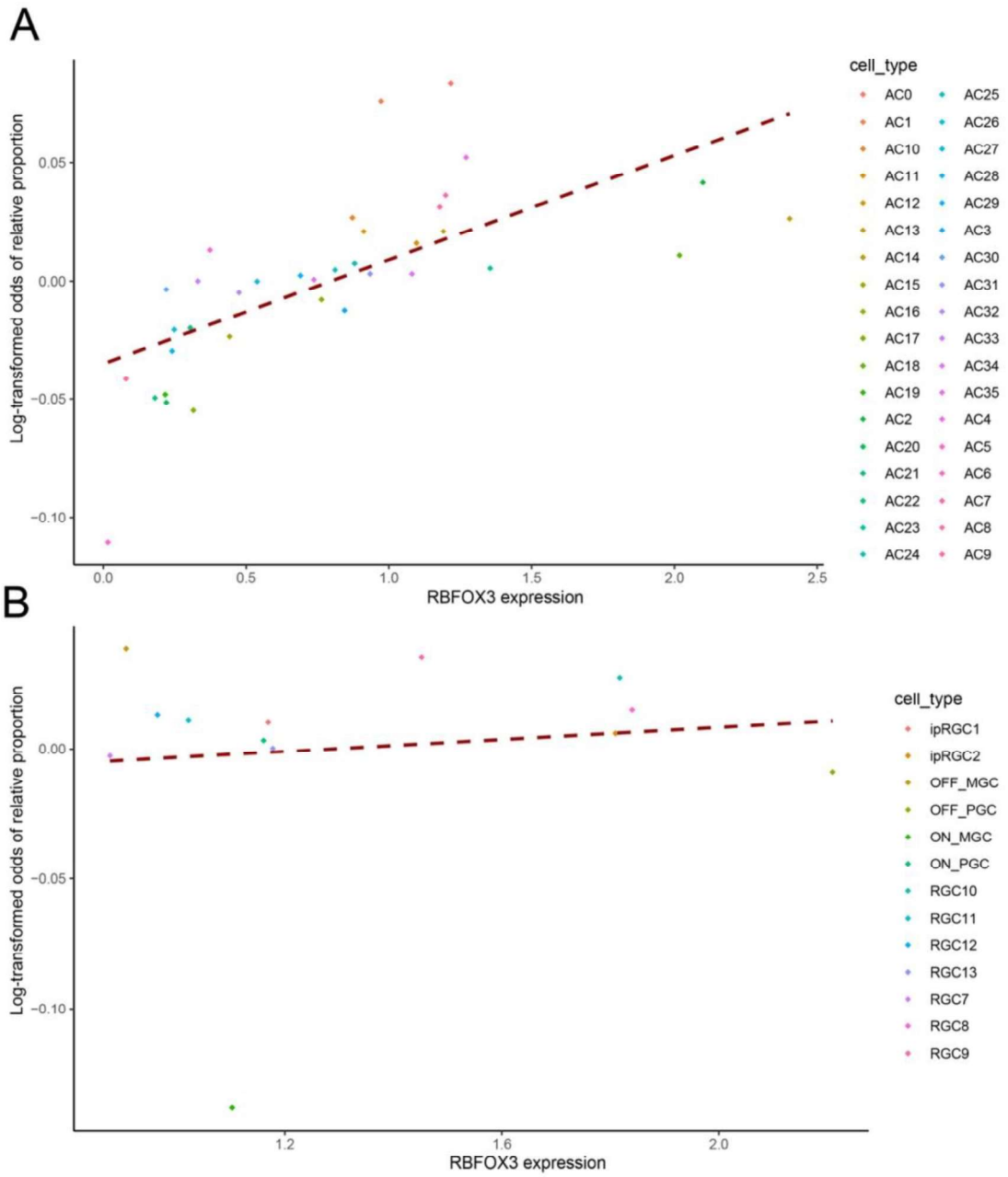

Figure S7, related to Figure 1. Correlation between RBFOX3 expression and relative cell abundance in peripheral retina of ACs (A) and RGCs (B). The relative cell abundance is measured by log-transformed quotient of peripheral relative proportion and central relative proportion for each cell type.

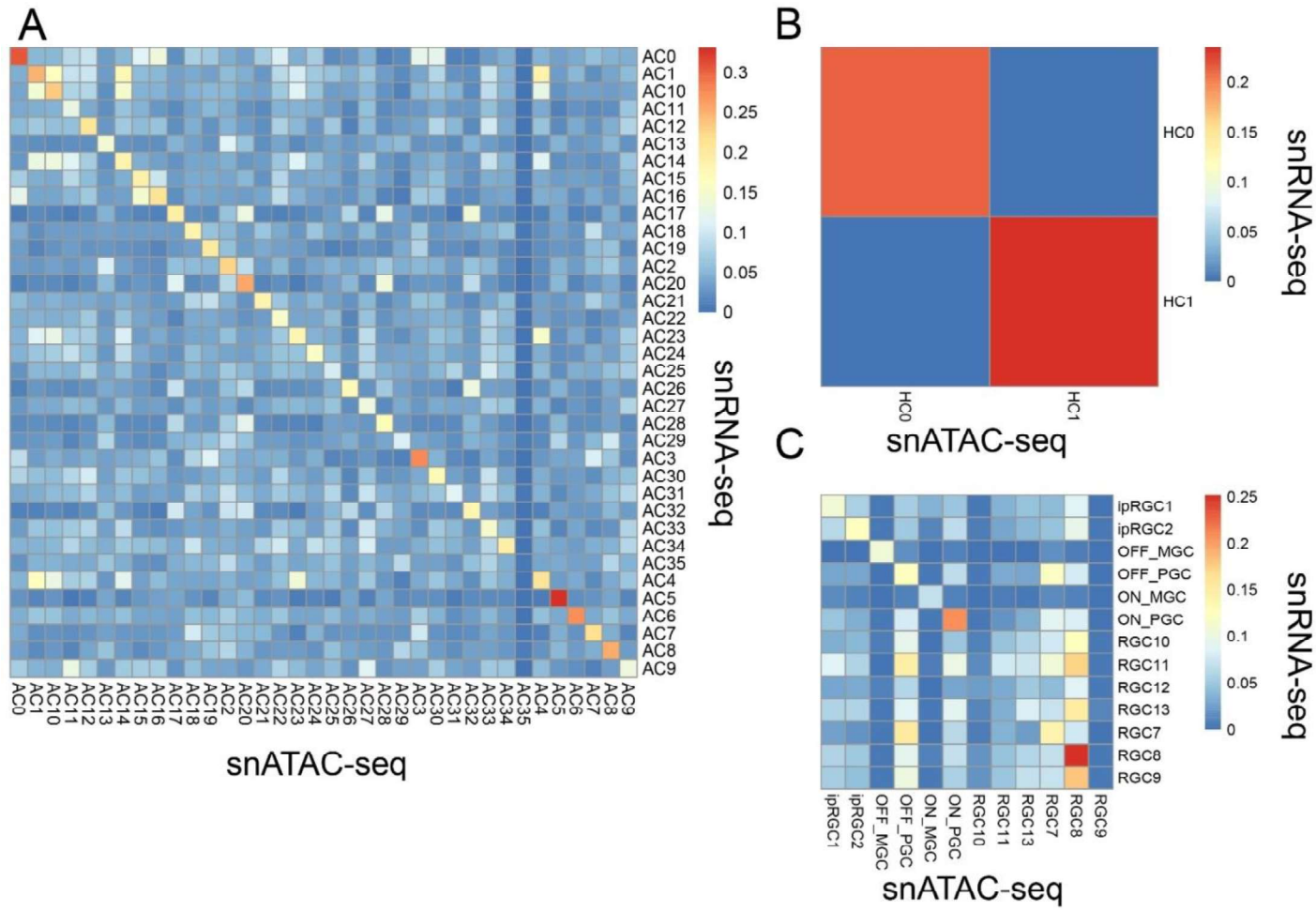

Figure S8, related to Figure 2. Evaluation of the reliability of the integration of snRNA-seq and snATAC-seq. Heatmap representing the similarity between the differentially expressed genes and the differentially accessible genes from each AC (A), HC (B), and RGC (C) types.

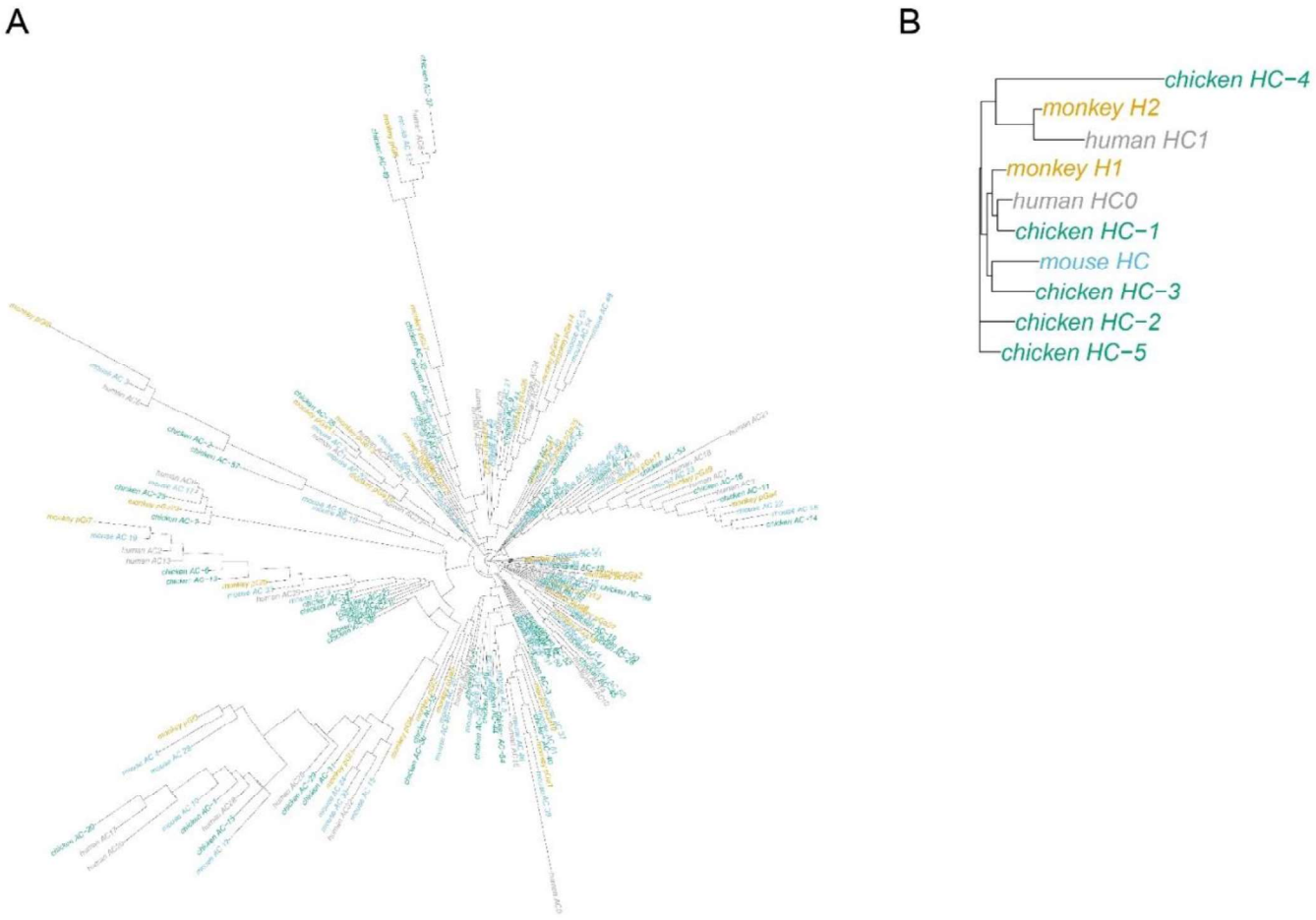

Figure S9, related to Figure 2. Phylogenetic tree representing the overall similarity of AC (A) and HC (B) types among four species: human (grey), monkey (yellow), mouse (blue), and chicken (green).

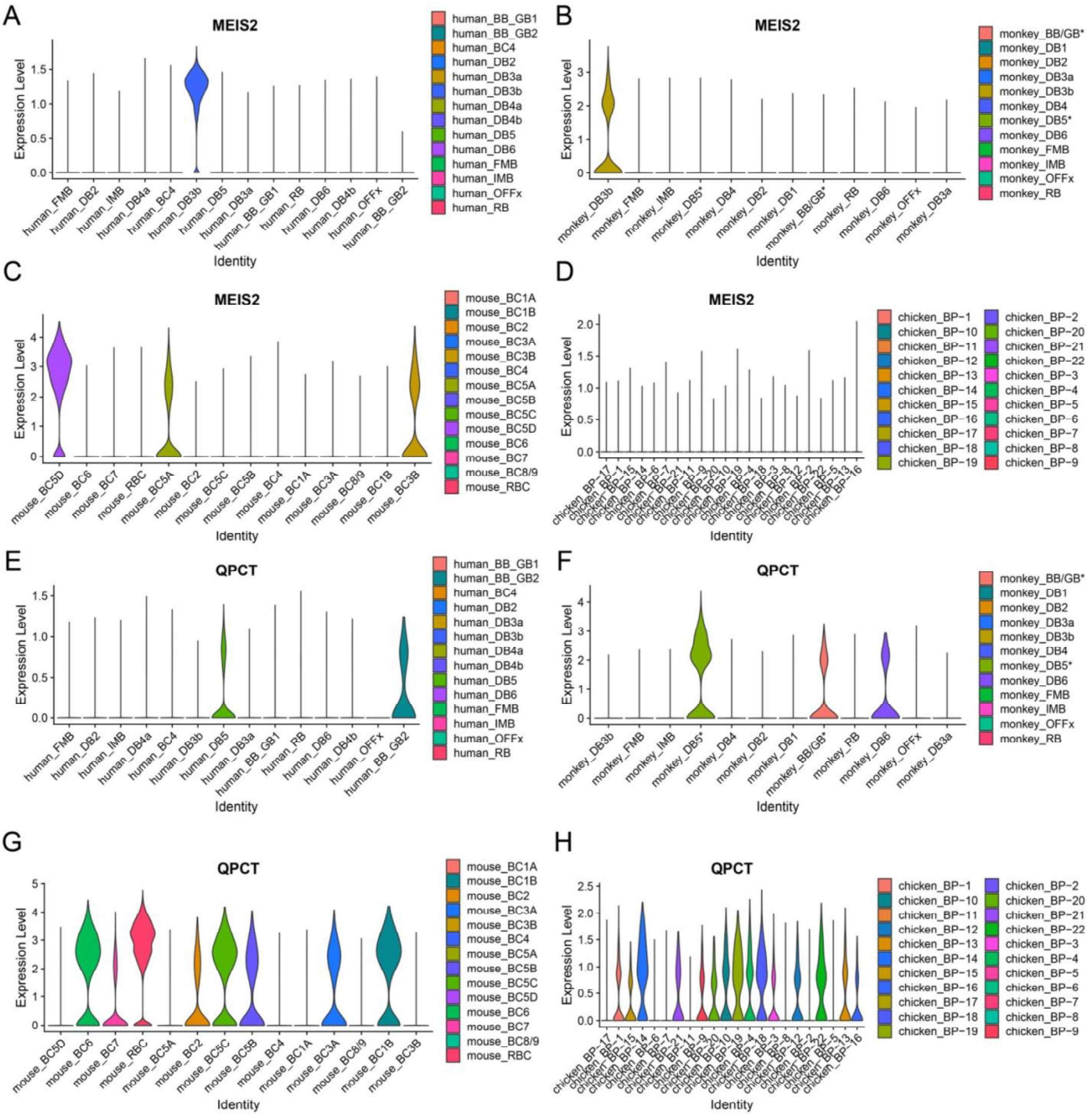

Figure S10, related to Figure 2. Marker genes conserved for DB3b and DB5 of primate retina.

A-D. Violin plot showing *MEIS2* expression in human (A), monkey (B), mouse (C), and chick (D) bipolar cells

E-H. Violin plot showing *QPCT* expression in human (E), monkey (F), mouse (G), and chick (H) bipolar cells

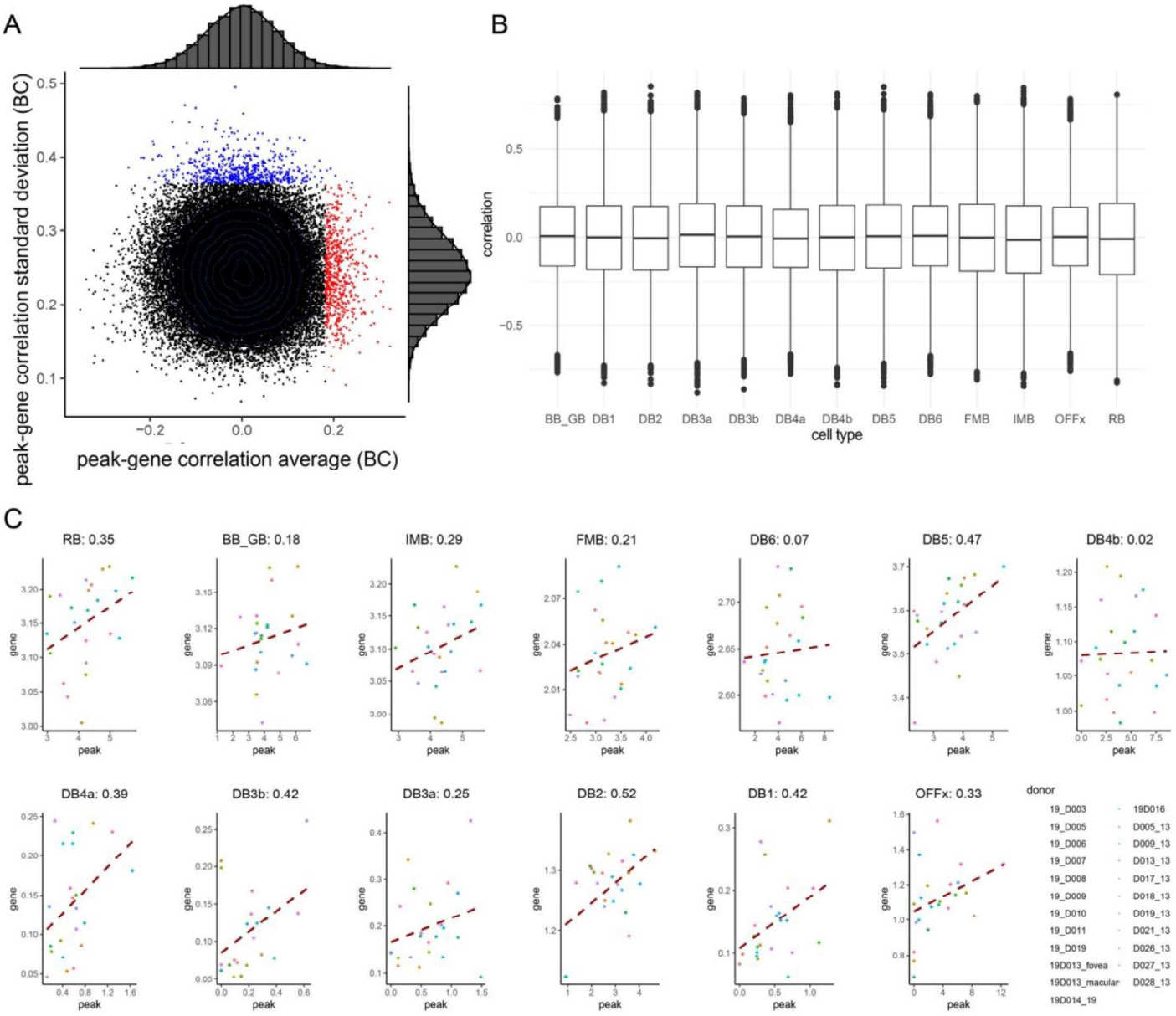

Figure S11, related to Figure 3. The consistency of CRE functions across types within a major cell class (BC)

A. Scatter plot showing the distribution of mean and standard deviation of each peak-gene links among the BC types. The data points being highlighted as blue and red were considered as 'heterogenous' and 'constant' peak-gene links, respectively.

B. Boxplots showing the overall distribution of gene-peak correlations for each BC type.

C. The peak gene correlation between chr3:61548584-61550601 and PTPRG in all bipolar cell types. For each panel, each data point represents a sample and the red dashed-line represents

the linear fit of the correlation between the peak and the gene. The Pearson's correlation coefficients were labeled on each panel.

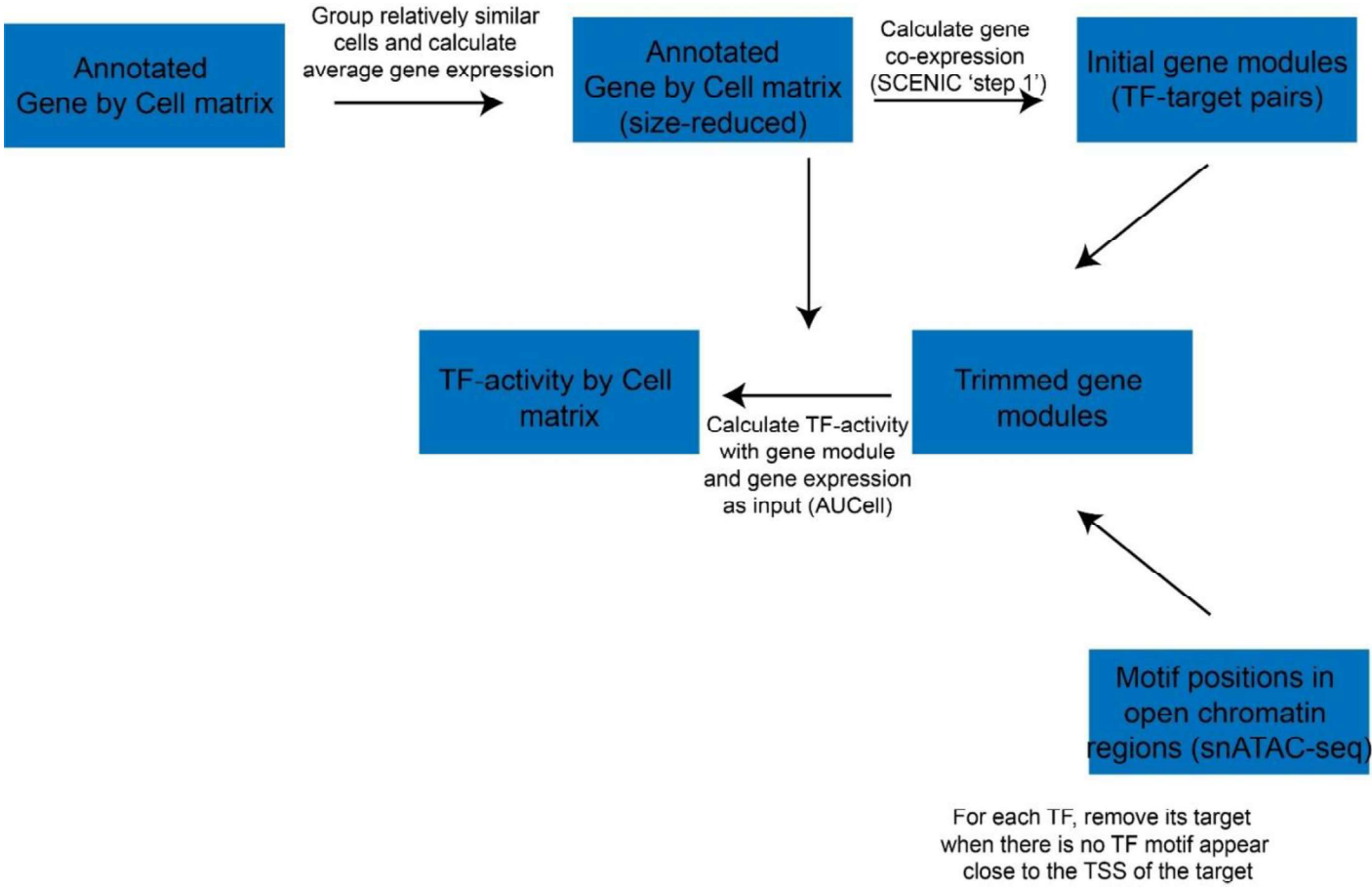

Figure S12, related to Figure 4 and STAR Methods. A schematic showing the workflow of the calculation of the TF activity.

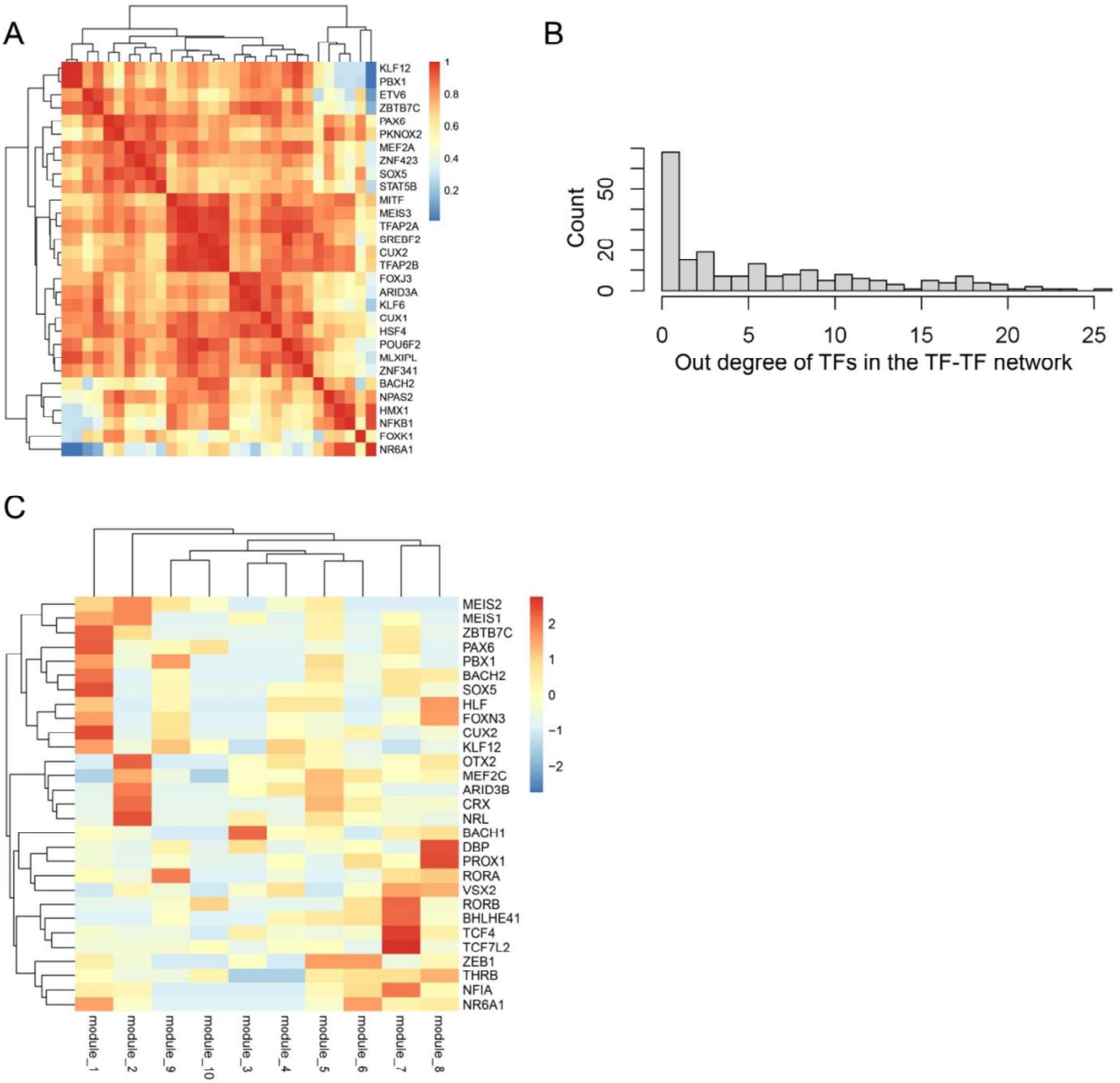

Figure S13, related to Figure 4 and STAR Methods. Detailed view and potential hub regulators of the TF-TF network.

A. A zoom-in view of the TF module 1 for enhanced demonstrations of underlying substructures.

B. Histogram showing the distribution of the out degree of TFs in the network.

C. Heatmap depicting the overlapping between TF targets and TFs in the pre-defined modules. The TFs were selected based on the out degree (larger than 15). The overlapping was evaluated by computing the Jaccard Similarity. The data was scaled by row before plotting.

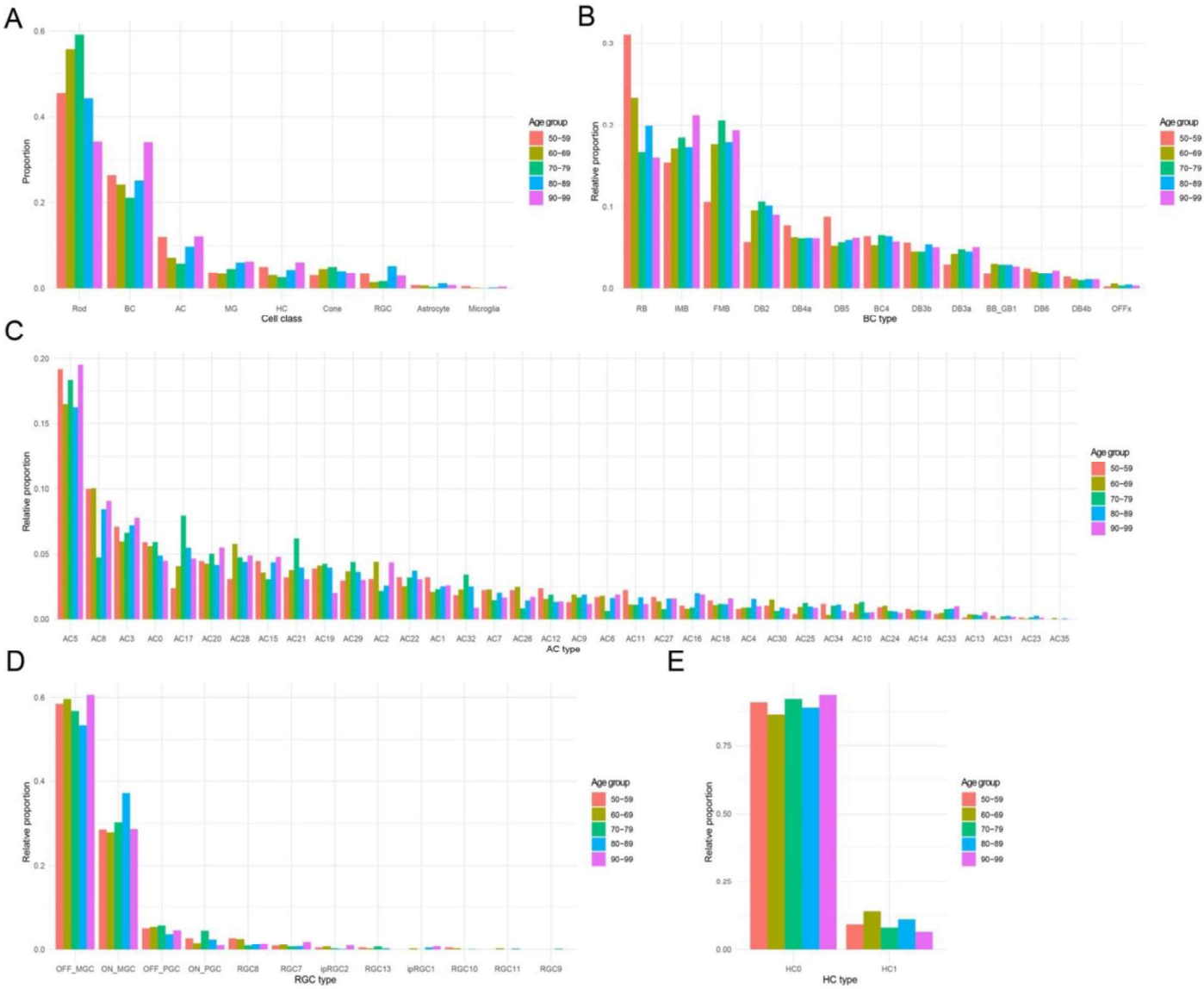

Figure S14, related to Discussion. Proportion and relative proportion of major cell classes (A) and BC (B), AC (C), RGC (D), and HC (E) from different age groups of the snATAC-seq data