Figure S1 (related to Figure 1)

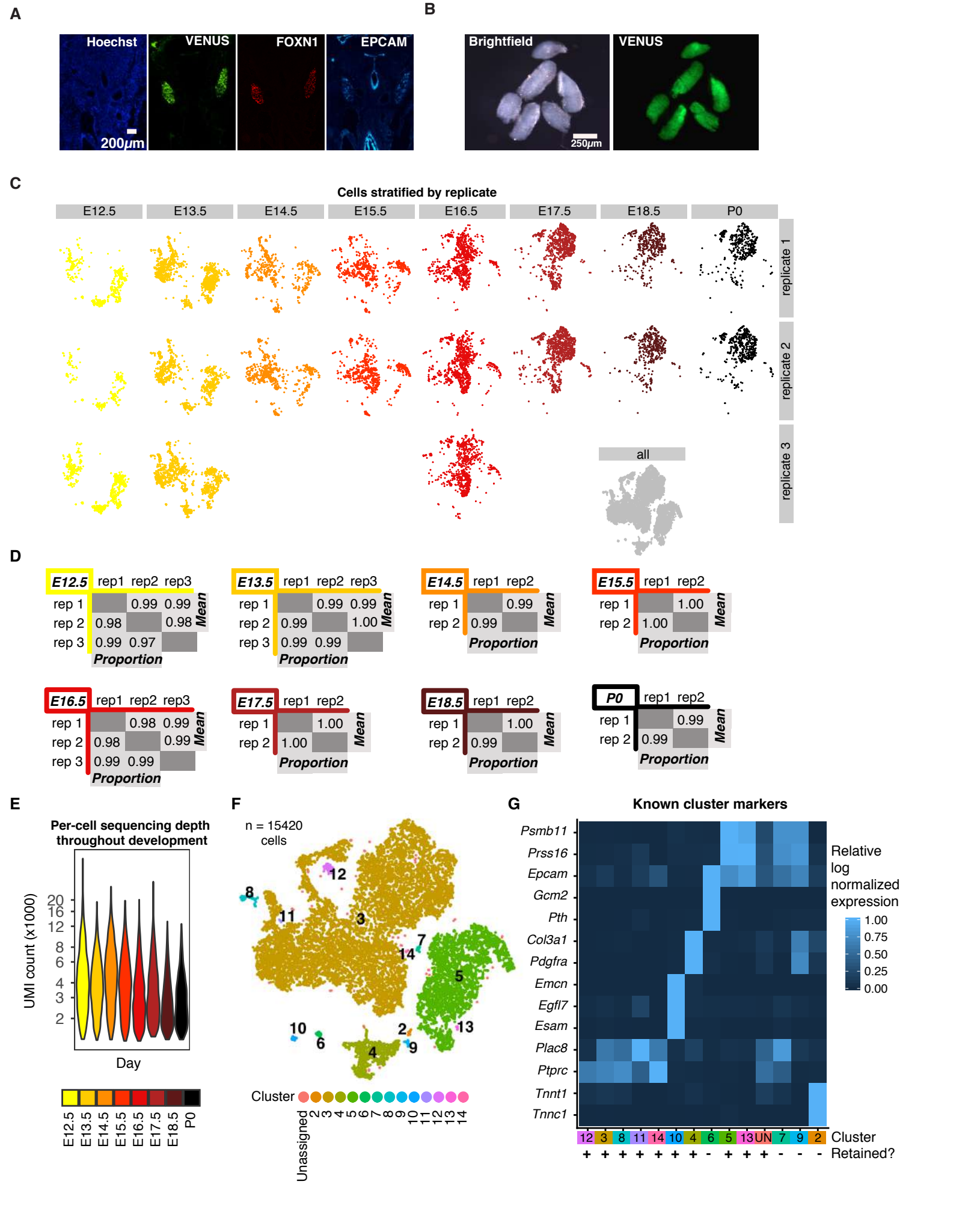
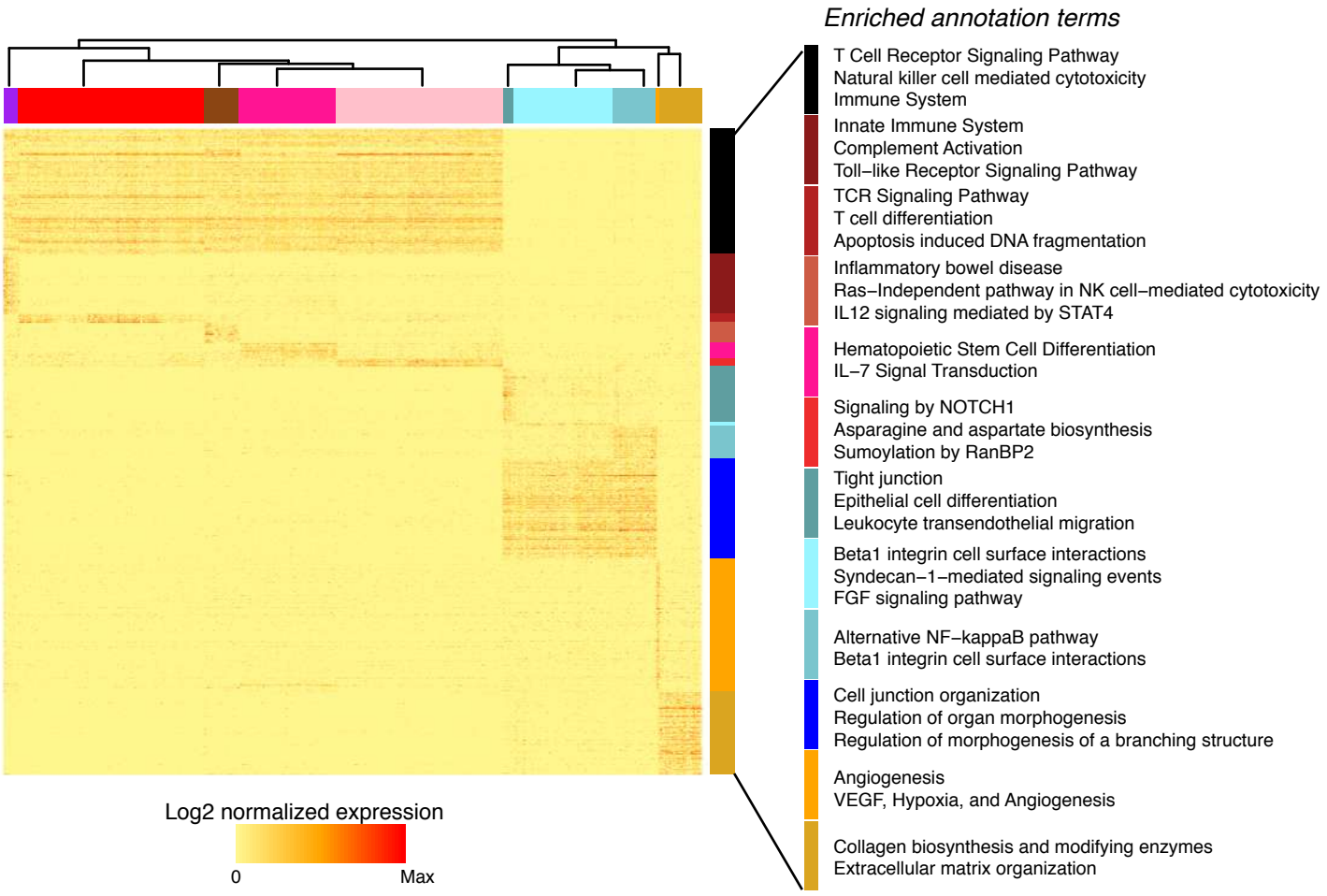


Figure S2 (related to Figure 1)

A



B

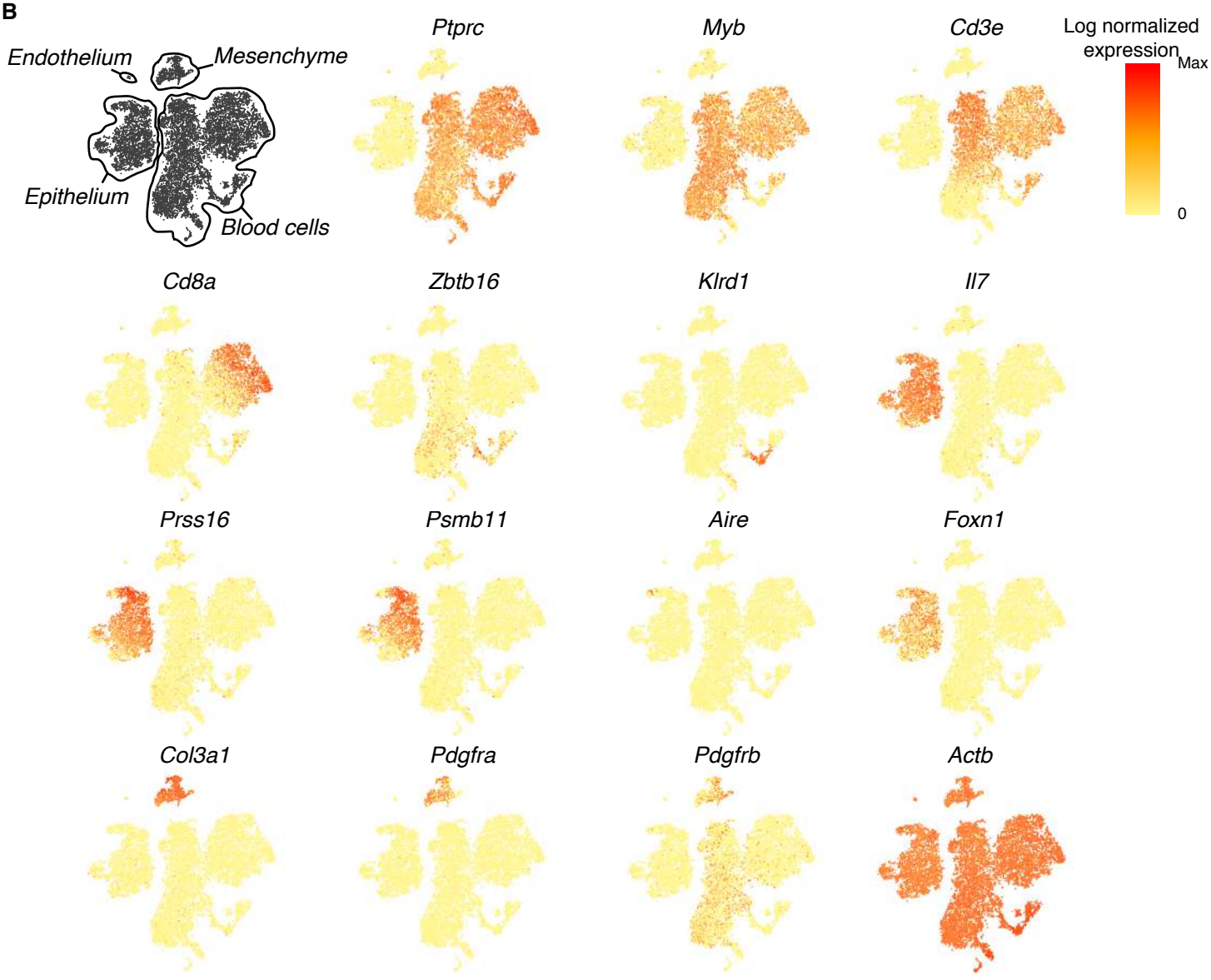


Figure S3 (related to Figure 2)

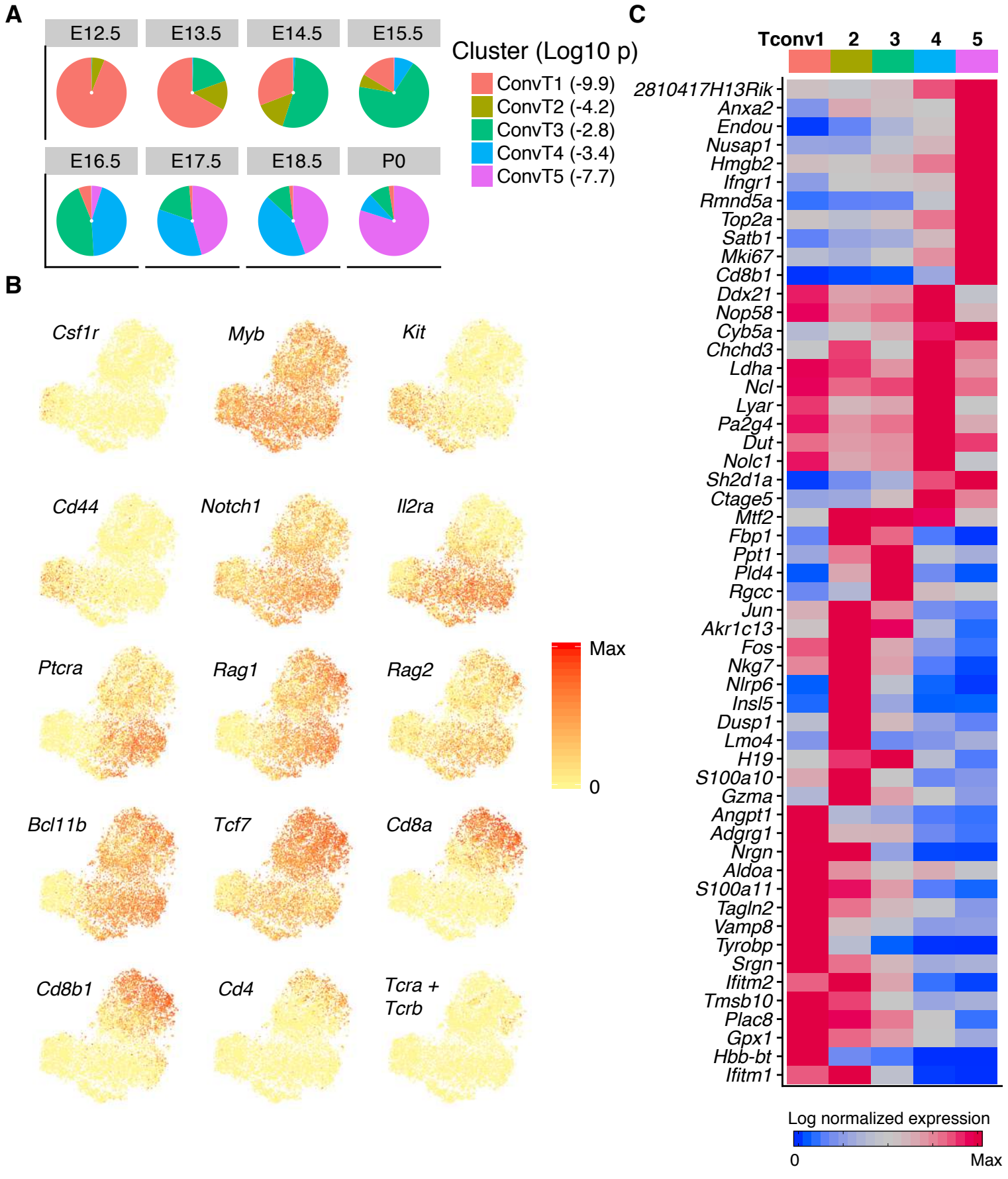
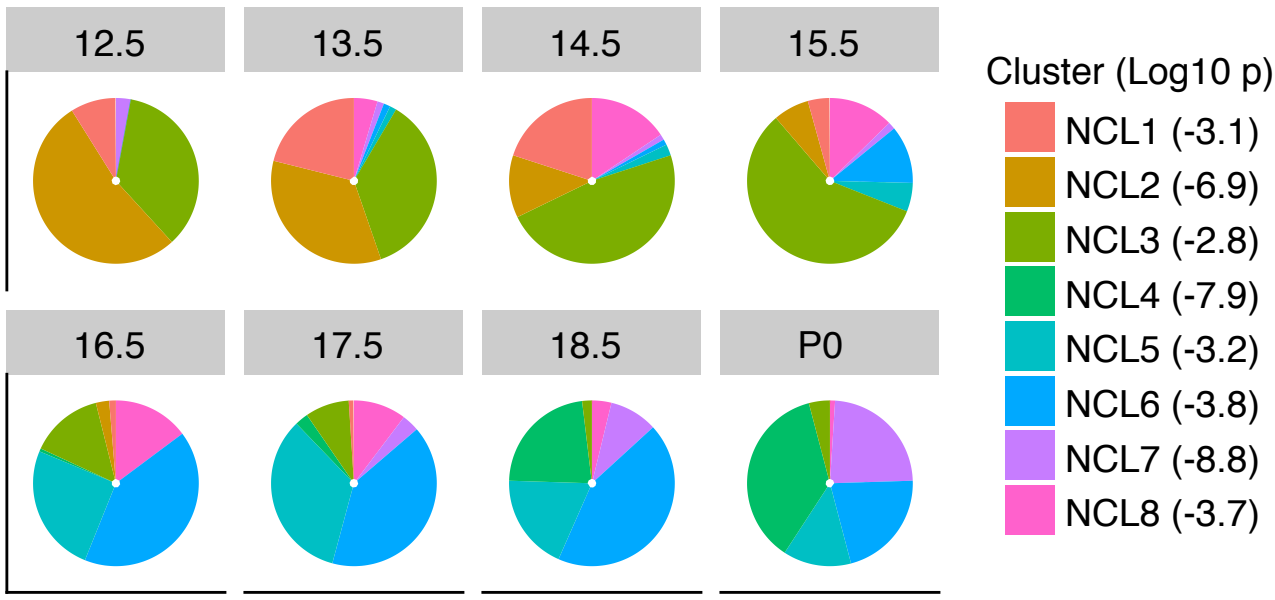


Figure S4 (related to Figure 3)

A



B

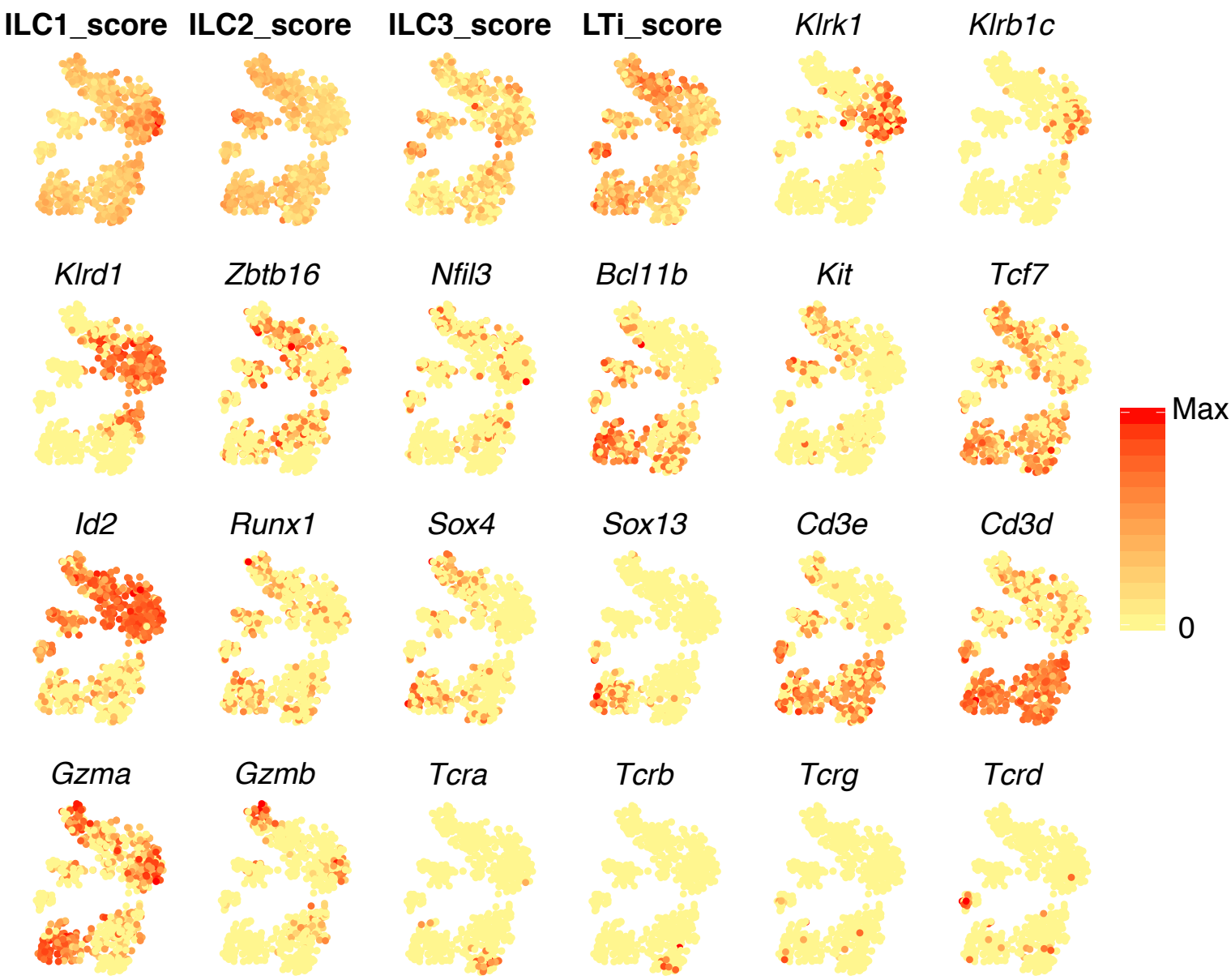


Figure S5 (related to Figure 5)

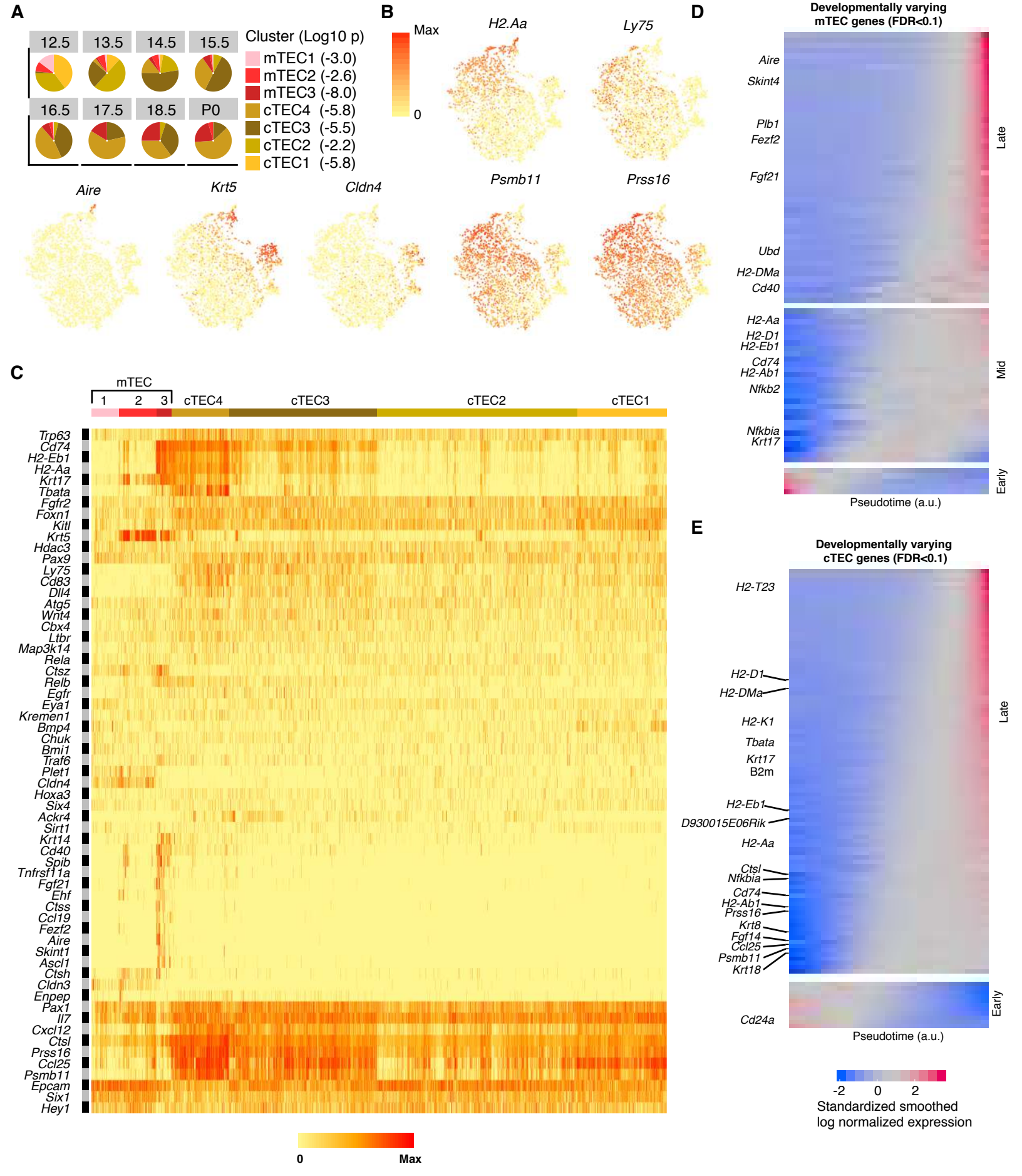
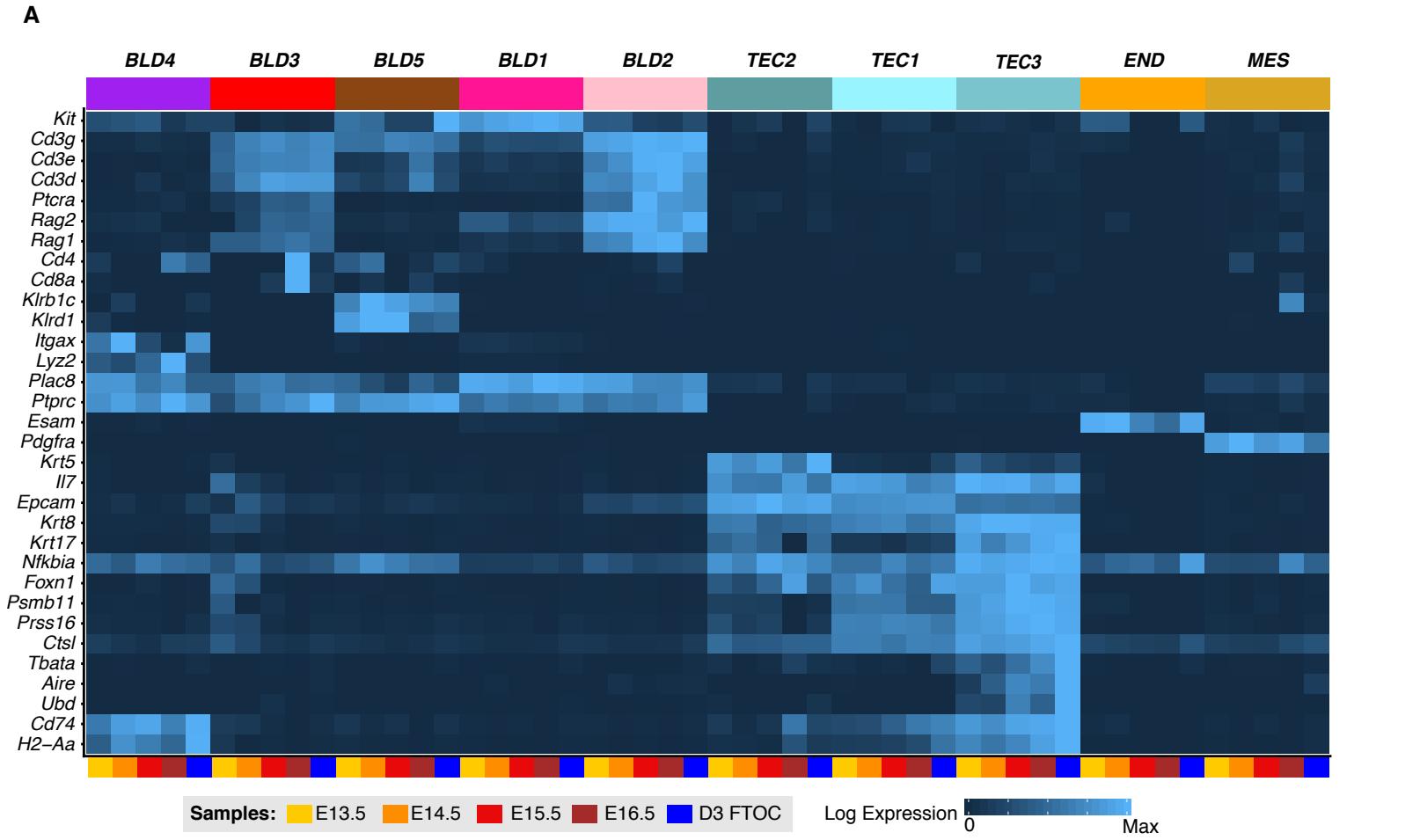
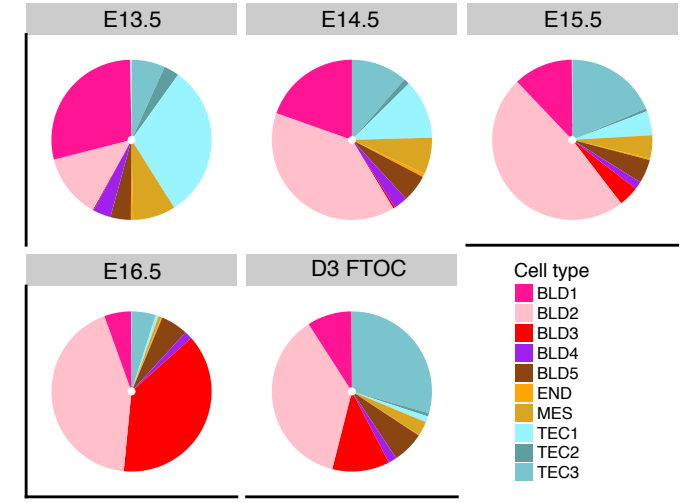


Figure S6 (related to Figure 6)





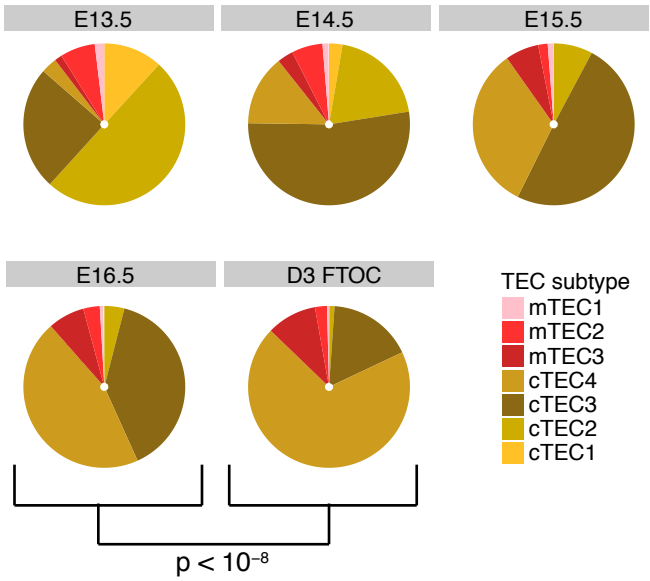


Figure S7 (related to Figures 6 and 7)

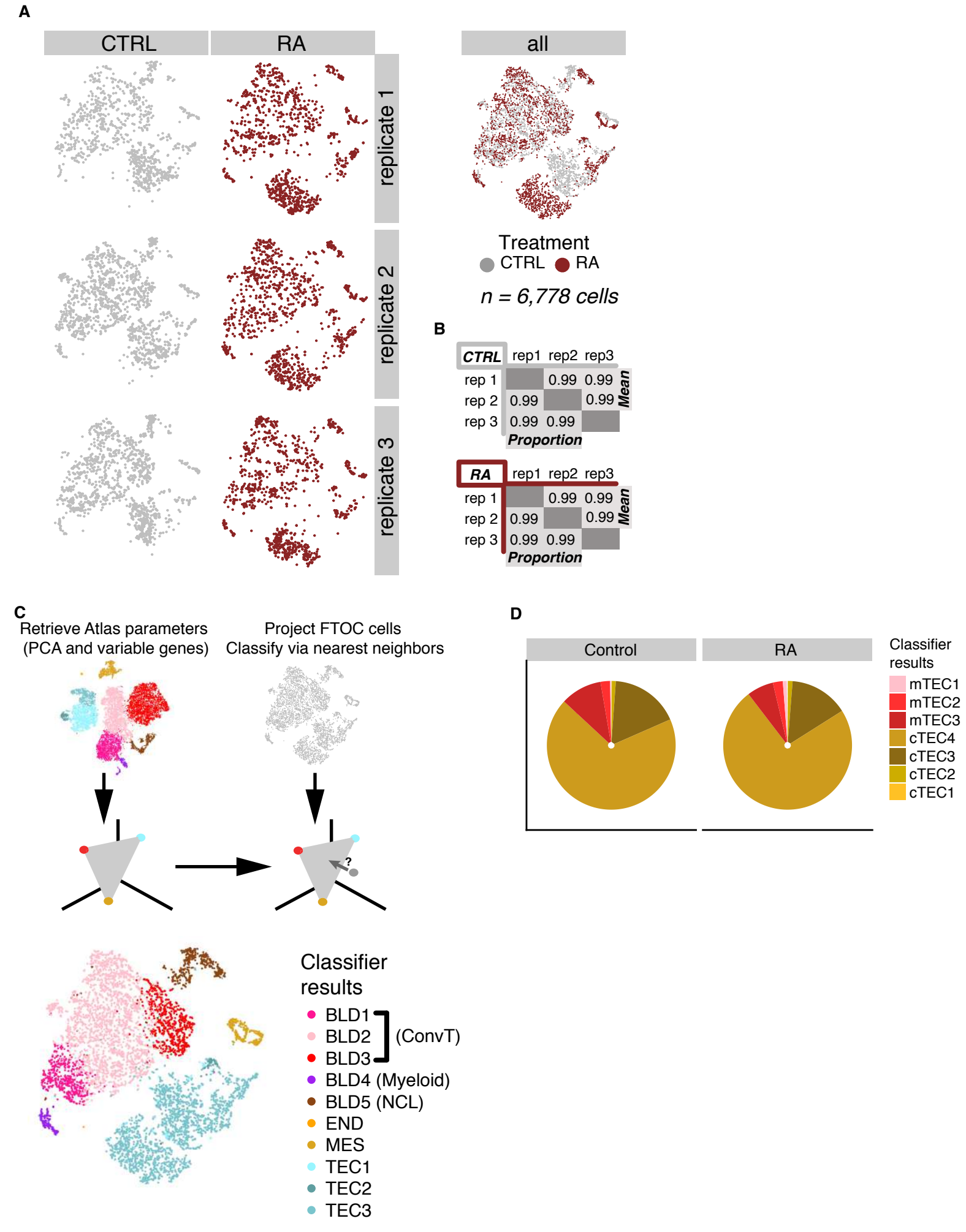


Figure S1 (Related to Figure 1): Data production and quality control for thymus cell atlas. (**A**) A coronal section of an E12.5 B6/*Pax9*^{VENUS} embryo showing VENUS fluorescence, and co-stained with FOXN1 and EpCAM, demonstrates how VENUS expression allows for identification of the E12.5 thymus; scale bar 200μm. (**B**) Thymi isolated from E12.5 VENUS+ B6/*Pax9*^{VENUS} embryos; scale bar 250μm. (**C**) tSNE with one dot per cell. Plot is faceted by replicate, and color corresponds to developmental day. (**D**) Pearson correlations across replicates. Replicates are compared gene by gene, with the expression of each gene measured as the average across cells (superdiagonal) or the proportion of cells expressing the gene (subdiagonal). All calculations begin with the log normalized data. (**E**) UMI counts resolved by developmental day. (**F**) Initial tSNE as shown in panel C and clustering results. Each dot is a cell. (**G**) Cluster-wise mean expression of selected markers. Each gene is scaled to the interval [0, 1]. Clusters are manually arranged to highlight major cell types, doublets, and muscle or parathyroid contaminants. Clusters are retained (+) or discarded (-) as indicated below the heatmap.

Figure S2 (Related to Figure 1): Transcriptomic diversity in the thymus. (**A**) Cluster-specific genes are shown at single-cell resolution, and clusters are ordered as in Fig. 1. Annotation terms are derived from top markers. Column and row colors match Fig. 1C, with black and blue row colors indicating marker annotation for BLD and TEC superclusters respectively. (**B**) tSNE with feature plots highlighting blood (*Ptprc*); T cells (*Myb*, *Cd3e, Cd8a*); innate lymphoid cells (*Zbtb16, Klrd1*); thymic epithelium (*Il7, Prss16*, *Psmb11*, *Aire*, *Foxn1*); and mesenchyme (*Col3a1*, *Pdgfra*, *Pdgfrb*). Housekeeping gene *Actb* is included for comparison. Color shows log normalized expression.

Figure S3 (Related to Figure 2): T cell-specific gene expression and cluster proportions. (**A**) Pie charts showing cluster proportions (as shown in Fig. 2C) by developmental day. P-values are computed via ANOVA with n=19 replicates (methods). (**B**) Selected T cell maturation genes, displayed via tSNE of developing thymocytes subsetted as in Fig. 2A. Color shows log normalized expression. (**C**) Row-standardized heatmap of genes showing the strongest expression differences across clusters. Entries display average log-normalized expression.

Figure S4 (Related to Figure 3): Cluster-specific expression and cluster proportions over time in NCLs. (**A**) Pie charts showing NCL cluster proportions, derived in Fig. 3C, by developmental day. P-values are computed via ANOVA with n=19 replicates (methods). (**B**) Selected genes displayed on tSNE of atlas NCL cells subsetted as in Fig. 3A. Color shows log normalized expression. ImmGen ILC and LTi scores are calculated based on Robinette et al (methods).

Figure S5 (Related to Figure 5): Additional information on TEC cellular heterogeneity. (**A**) Pie charts showing TEC cluster proportions, derived in Fig. 5C, by developmental day. P-values are computed via ANOVA with n=19 replicates (methods). (B) Selected genes shown on tSNE of the atlas TECs subsetted as in Fig. 5A. Color shows log normalized expression. (**C**) Heatmap showing curated transcripts at cellular resolution. Cells are ordered via hierarchical clustering within TEC subtypes. (**D,E**) Heatmap of pseudotime-ordered genes displayed by mTEC (**D**) and cTEC (**E**). After smoothing, genes are clustered via K-means. For display, each gene is standardized to have mean 0 and standard deviation 1. A full list of pseudotime-varying genes in mTEC and cTEC is given in table S5C and D.

Figure S6 (Related to Figure 6): FTOC cellular heterogeneity. (**A**) Cluster-specific expression of handpicked markers during natural development (E13.5 to E16.5) and in the day 3 (D3) FTOC. Heatmap shows average log normalized expression, and each row is scaled to a maximum of 1. (**B,C**) Cluster composition of *in vivo* samples and of the FTOC for the whole thymus (**B**) and TEC subsets (**C**). P-value in (**C**) compares E16.5 samples to the FTOC, corresponding to a null hypothesis that the FTOC cells do not differ from *in vivo* counterparts of the same age. P-values are computed by Fisher's exact test with n=1250 TECs.

Figure S7 (Related to Figures 6 and 7): Quality control measures for FTOC data and effects of retinoic acid on FTOC cellular heterogeneity. (**A**) tSNE showing treatments and replicates. (**B**) Pearson correlations between replicates. (**C**) Schematic of methodology for

applying atlas cluster labels to FTOC cells, and resulting tSNE showing classifier based labels (methods). (**D**) Composition of TEC subsets in control and RA-treated D3 FTOCs.

Supplemental table titles and legends

Table S1: Cell-type markers. Related to Fig. 1.

(**A**): Genes in Figure S2 large heatmap. Genes and cluster assignments as displayed in Fig. 1d. To be assigned to a cluster, the fold change in the percent expressing (FCPE) had to exceed 2. To be assigned to the blood or TEC superclusters, the FCPE had to exceed 4.

(**B**): Data-driven markers for clusters displayed in Fig. 1c, filtered as described in the methods.

Table S2: T-cell subset markers. Related to Fig. 2.

Data-driven markers for clusters displayed in Fig. 2c, filtered as described in the methods. Related to Fig. 2.

Table S3: NCL subset markers. Related to Fig. 3.

Data-driven markers for clusters displayed in Fig. 3c, filtered as described in the methods. Related to Fig. 3.

Table S4: Differential expression between Rag-/- and wt thymi. Related to Fig. 4.

(**A**): Differentially expressed genes between Rag1 KO (high) and wild-type (low) mature T-cells, filtered as described in the methods.

(**B**): Differentially expressed genes between Rag1 KO (high) and wild-type (low) mature T-cells, filtered as described in the methods.

(**C**): Geneset annotations for genes higher and lower expressed in Rag-/- NCLs.

Table S5: TEC subset markers and genes varying over pseudotime. Related to Fig. 5.

(**A**): a: Differentially expressed genes for TEC clusters, filtered as described in the methods.

(**B,C**): developmentally regulated genes in mTEC and cTEC respectively. Genes are filtered for FDR < 0.1 with p-values based on ANOVA.

(**D**): Subset of table S5a containing only surface receptors, re-ranked by FCPE.

Table S7: Differential expression and cellularity between RA and control FTOCs. Related to Fig. 7.

(**A**): Cell counts for RA and control FTOC replicates.

(**B**): Differentially expressed genes between control and RA FTOC for TEC compartment, filtered as described in the methods.