**Stem Cell Reports, Volume *13***

## Supplemental Information

# Single-Cell RNA Sequencing of hESC-Derived 3D Retinal Organoids Reveals Novel Genes Regulating RPC Commitment in Early Human Retinogenesis

**Xiying Mao, Qin An, Huiyu Xi, Xian-Jie Yang, Xiangmei Zhang, Songtao Yuan, Jinmei Wang, Youjin Hu, Qinghuai Liu, and Guoping Fan**

## Supplemental figures

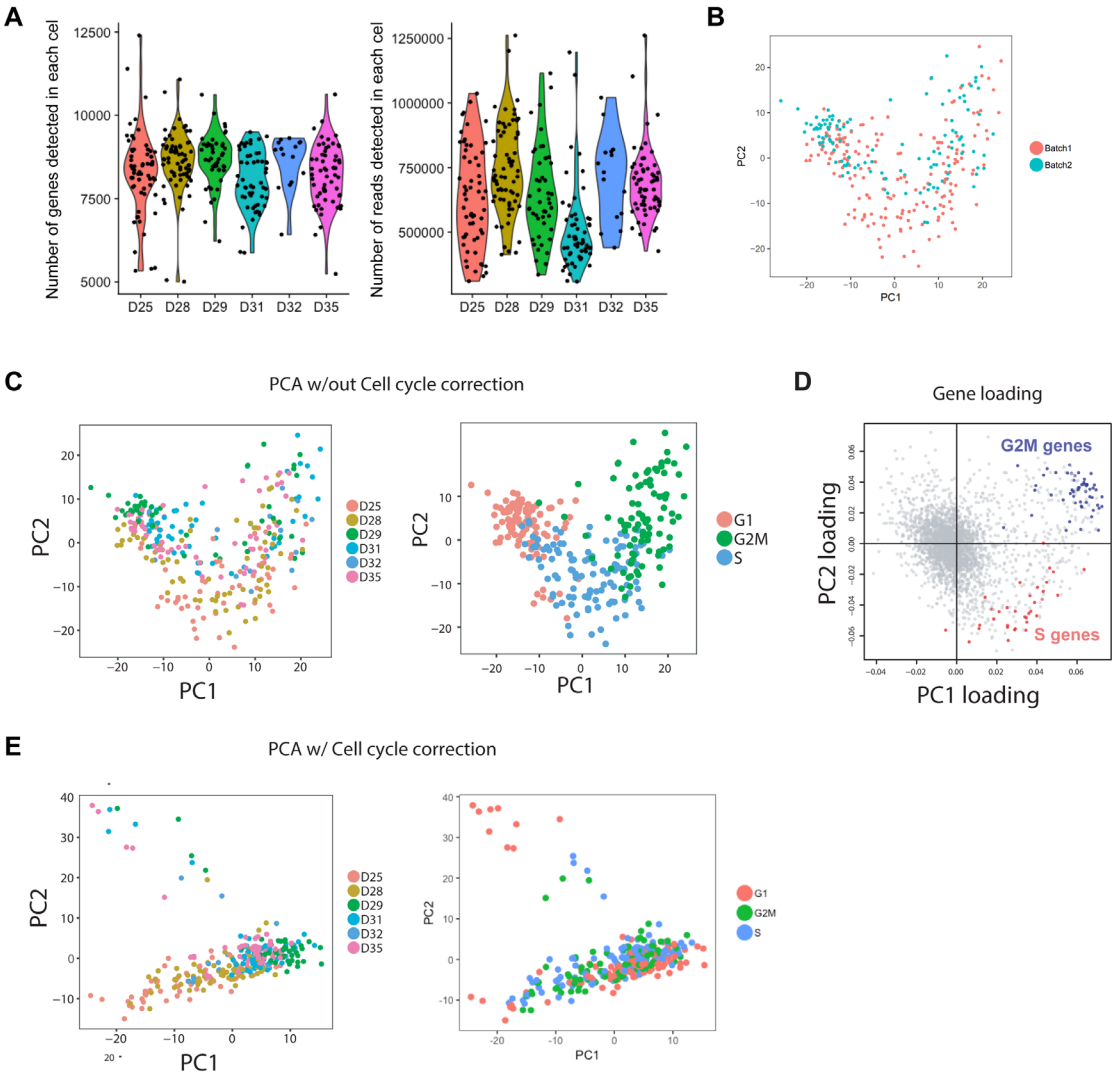


**Figure S1:** Sequencing and mapping of scRNA-seq data of 3D hESC-retinoids. **A:** Distribution of the genes and reads per single cell from different time points of differentiation. **B:** Scatterplot showing no obvious batch effect between the two batches of differentiation. **C:** PCA plot using transcriptomic data without cell-cycle bias correction. Cells were clustered by cell cycle phase. **D:** Scatter plot showing the PC loading score for each gene. Genes associated with PC loading showed that cell-cycle associated genes were highly variable and contributed most of the variation between single cells. **E:** PCA plot using the transcriptomic data after correcting cell-cycle bias. Cells at different cell cycle stages were presented in different colors. **Related to Figure 1.**

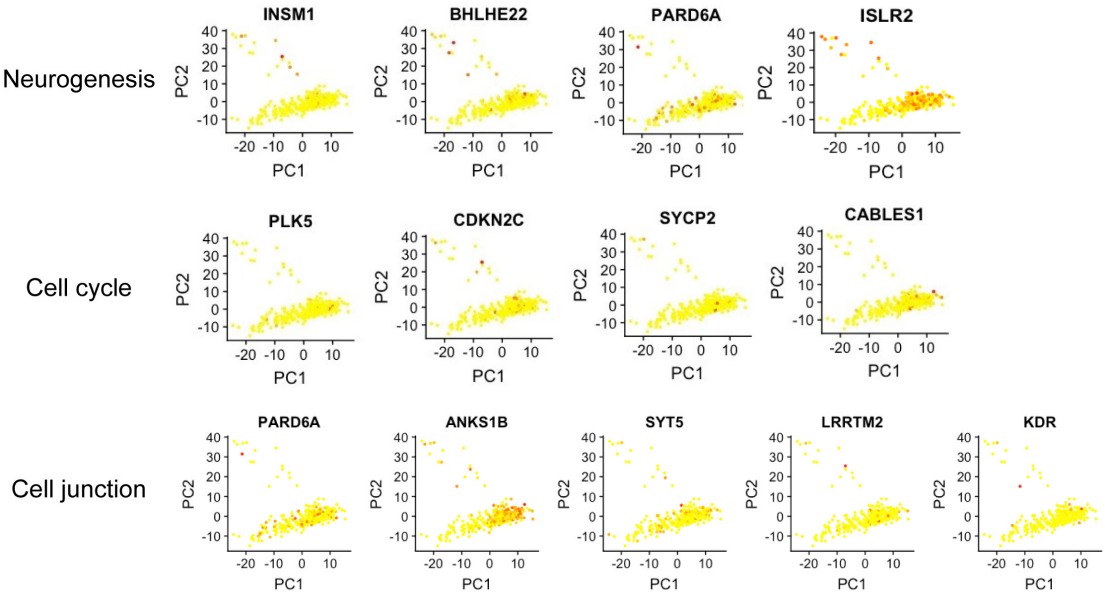


**Figure S2:** Expression pattern of additional ASCL1-correlated genes related to Neurogenesis (*INSM1*, *BHLHE22*, *PARD6A* and *ISLR2*), Cell-cycle (*PLK5*, *CDKN2C*, *SYCP2* and *CABLES1*) and Cell-junction (*PARD6A*, *ANKS1B*, *SYT5*, *LRRTM2* and *KDR*). **Related to Figure 2.**

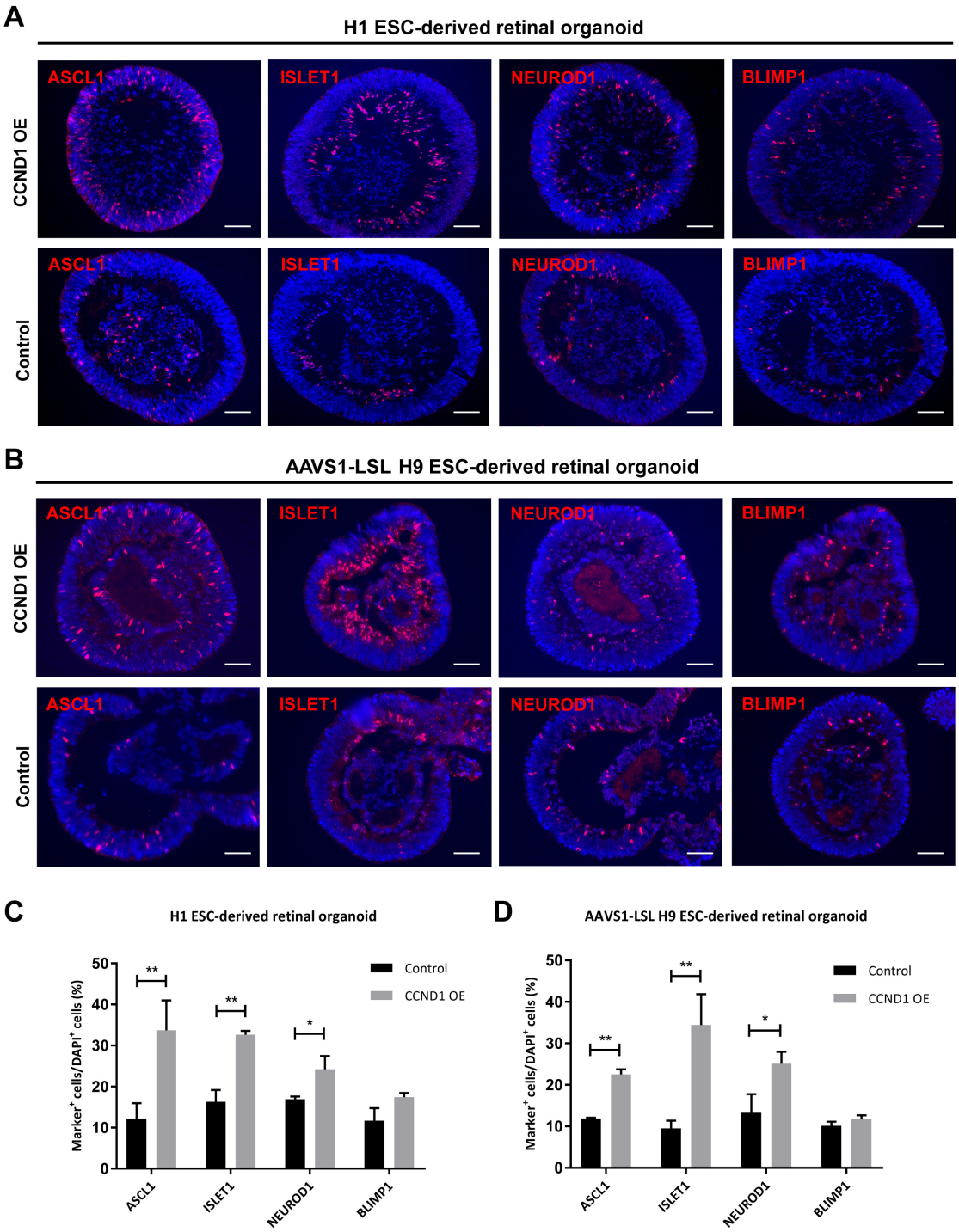


**Figure S3:** Similar results of CCND1 overexpression experiment were obtained in other hESC lines. **A-B:** Expression patterns of ASCL1, ISLET1, NEUROD1 and BLIMP1 in D32 CCND1-overexpressing (CCND1 OE) retinal organoids and control organoids derived from H1 ESC and AAVS1-loxp-3Stop-loxp-tdTomato (AAVS1-LSL) H9 ESC. Scale bars, 50 μm. **C-D:** Quantification of Marker$^+$ cells in CCND1 OE retinal organoids and control organoids (mean ± SD; n ⩾ 5 organoids from 2 differentiations) at D32. ∗p < 0.05, ∗∗p < 0.001, unpaired t test. **Related to Figure 3.**

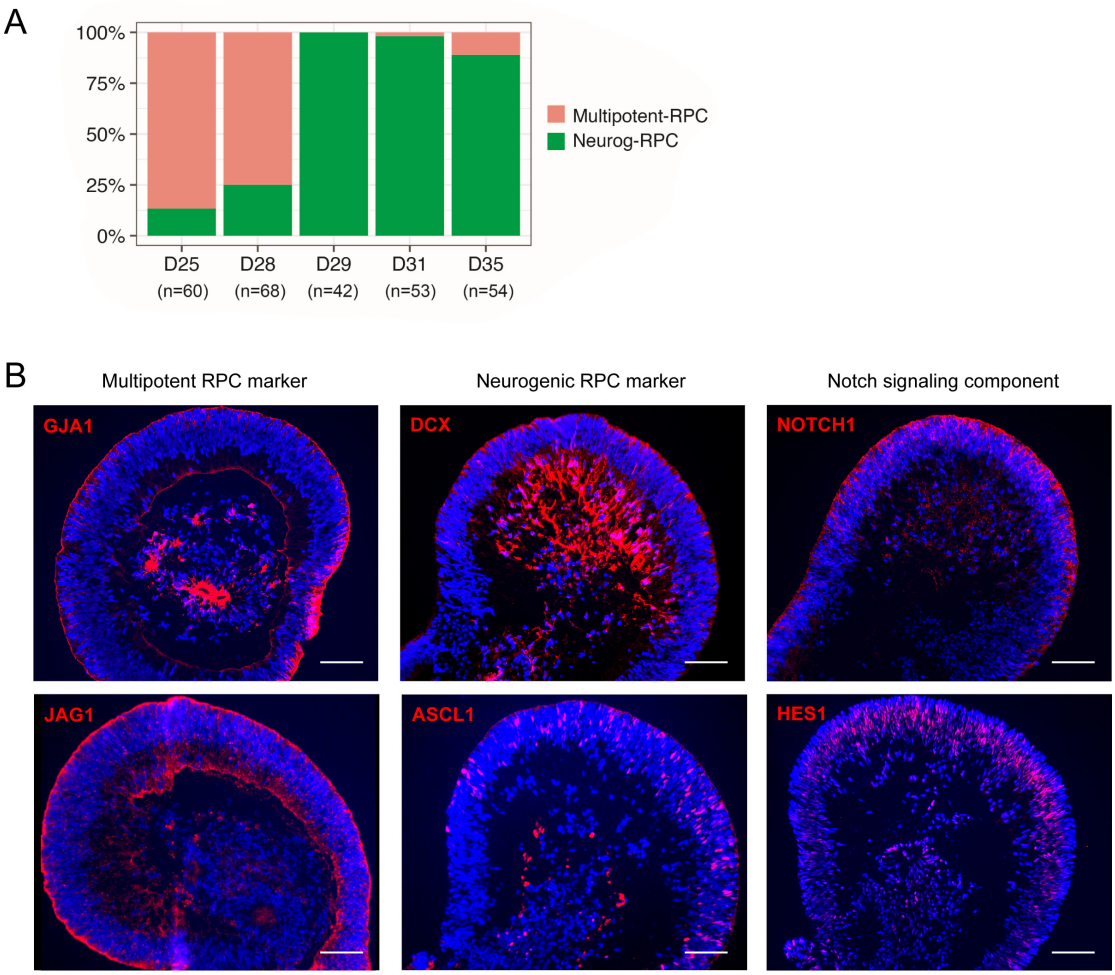


**Figure S4:** Molecular transition from multipotent RPC to neurogenic RPC during early retinogenesis. **A**: Stacked barplot showing the percentage of pluripotent RPCs and neurogenic RPCs at each time points. Most pluripotent RPCs were collected before D28 whereas most neurogenic RPCs were collected after D28. **B:** Immunostaining of multipotent RPC markers (JAG1 and GJA1), neurogenic RPC markers (DCX and ASCL1) and notch signaling components (NOTCH1 and HES1) showing a central-to-peripheral graded pattern at D40. Scale bars, 50 µm. **Related to Figure 4.**

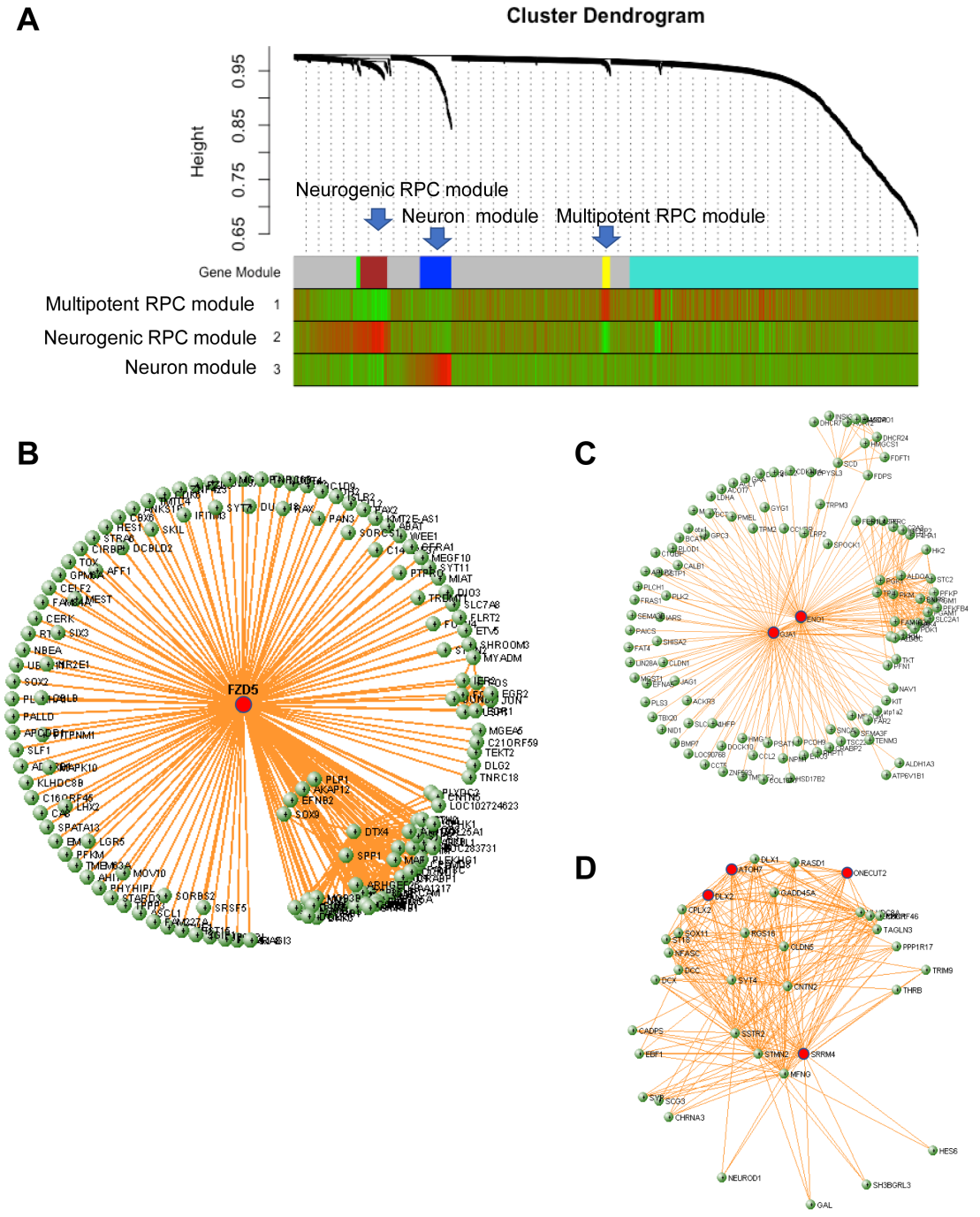


**Figure S5:** WGCNA Network analysis of RPC progression in early retinogenesis. **A:** On top, hierarchical cluster tree showed co-expression relationship between each gene. The "Gene Module" color bar shows the assignment of each gene to modules. The grey means the gene cannot be assigned into any significant co-expression module. The bottom 3 bars show the correlation between each gene and the cell types. Red means a gene is positively correlated with a cell type label, therefore trend to be highly expressed in this cell type. Green means a gene is negatively correlated with a cell type label, therefore trend to be repressed in this cell type. **(B-D):** Visualization of network topological structure in neurogenic module (B), multipotent RPC module (C) and neuronal module (D).

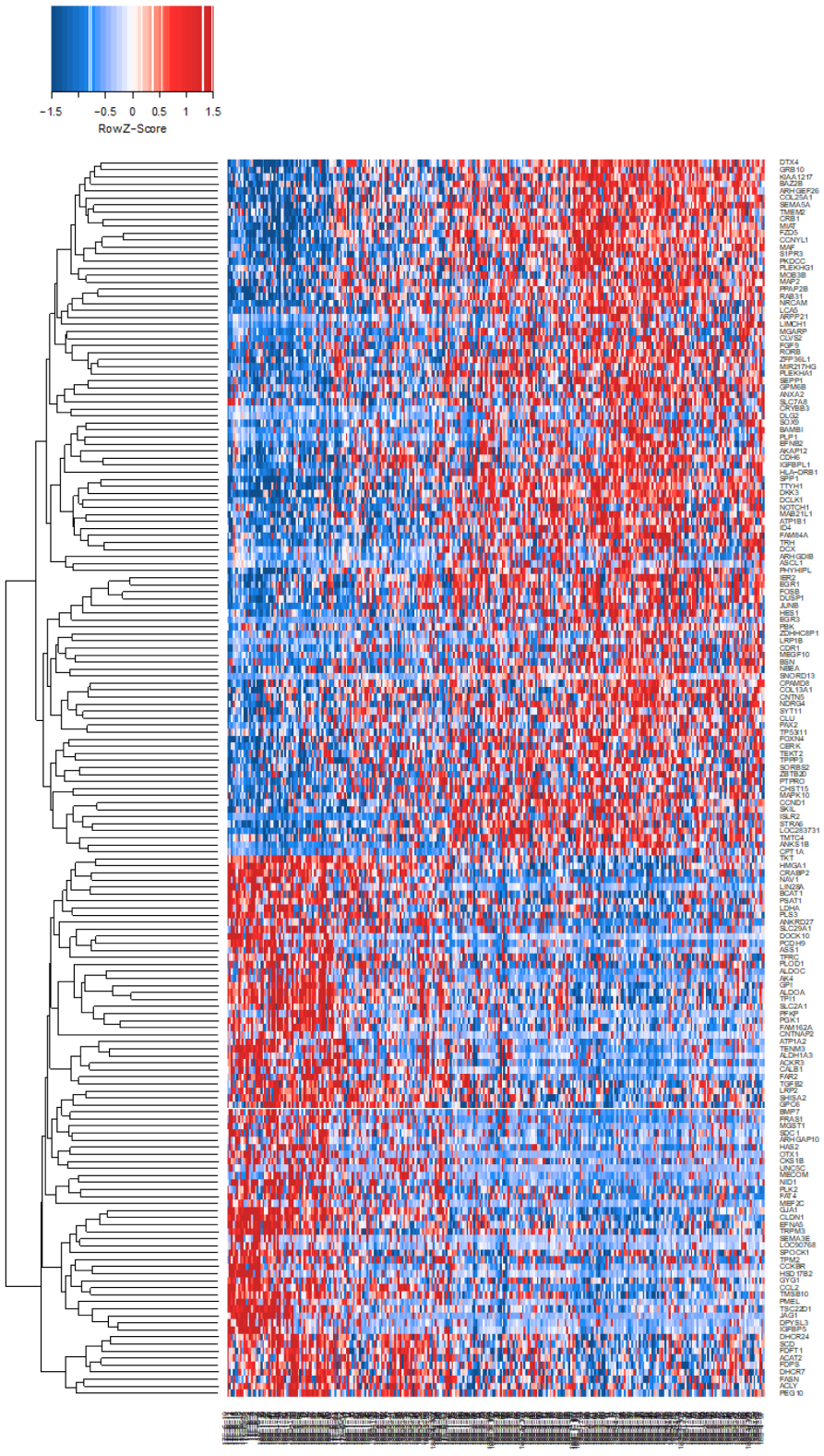


**Figure S6:** Pseudotime analysis of RPC progression using principle curve method. Heatmap of differential expression genes positively or negatively correlated with the pseudotime of RPC progression. **Related to Figure 5.**

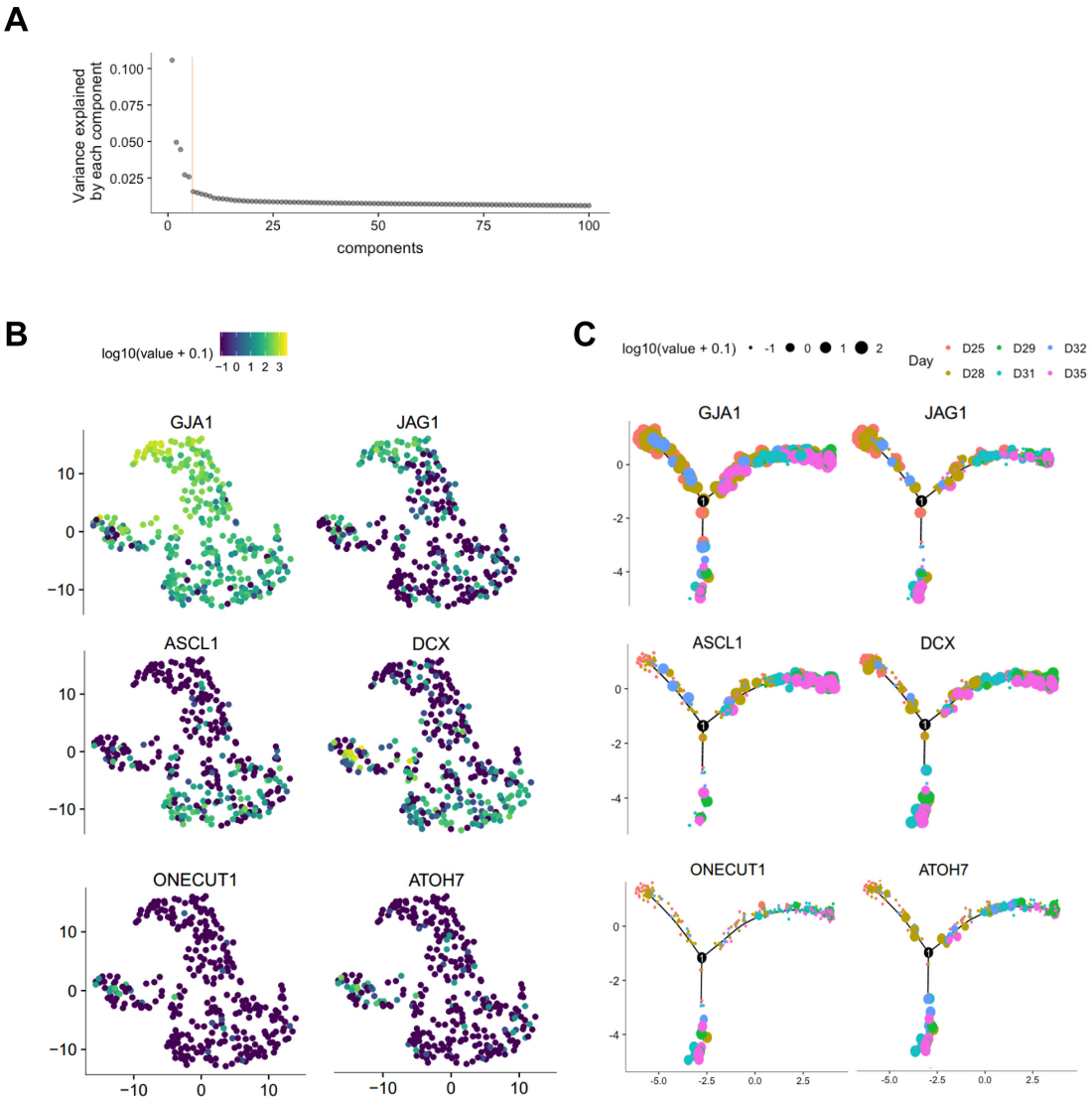


**Figure S7:** Pseudotime analysis of all single cells using Monocle 2. **A:** Contribution of each PCs to explain the total variance. The first 5 PCs were used for t-SNE dimensional reduction when running Monocle2. **B:** t-SNE plots generated by Monocle2 using top 5 most significant PC. Colors of each cell shows the normalized expression levels (log10[value+0.1]) of *GJA1*, *JAG1*, *ASCL1*, *DCX*, *ONECUT1* and *ATOH7*. **C:** Gene expression pattern on the trajectory generated by Monocle 2. The size of each cell shows the normalized expression levels (log10[value+0.1]) of GJA1, JAG1, ASCL1, DCX, ONECUT1 and ATOH7. **Related to Figure 6.**

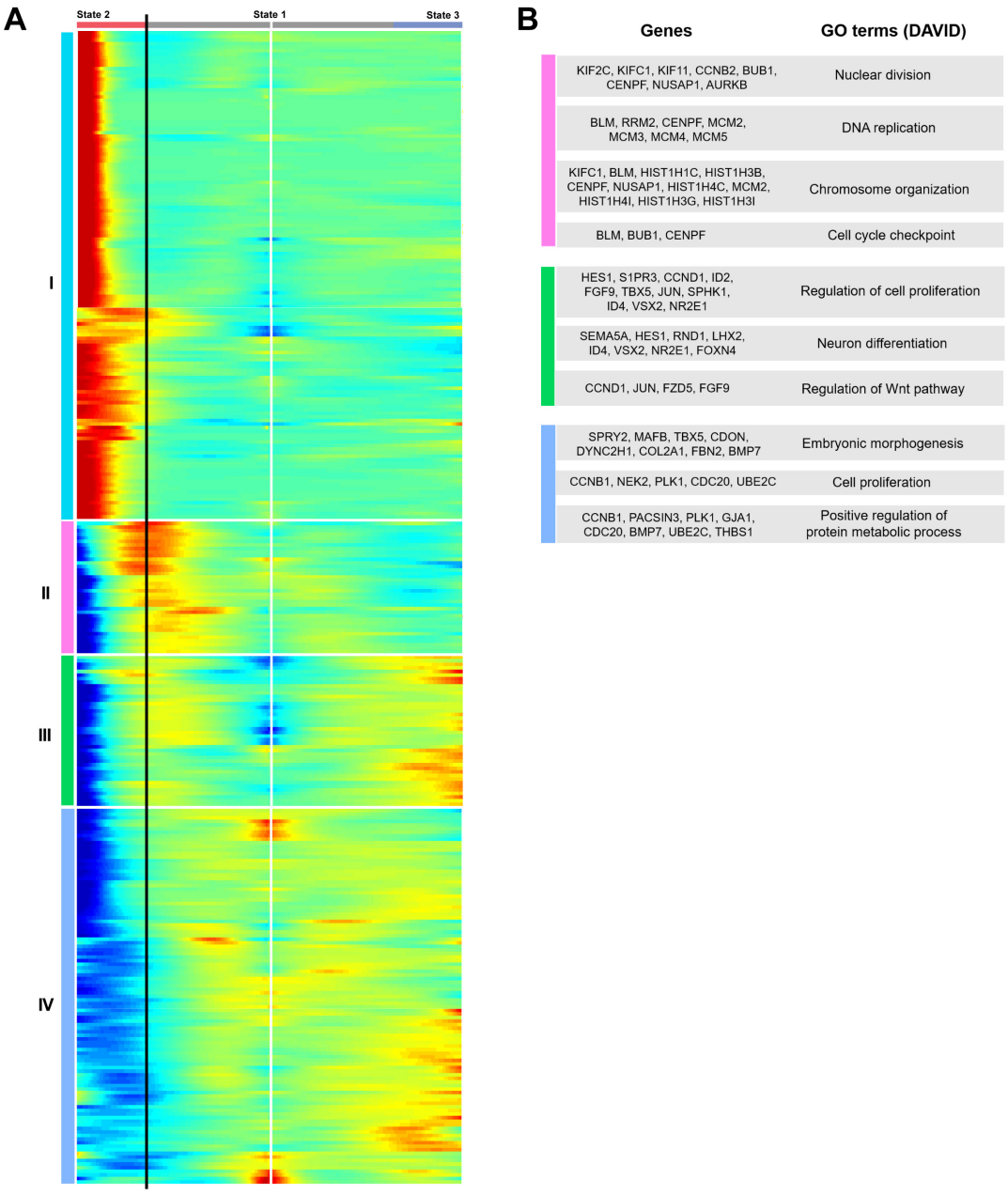


**Figure S8:** Differential expression analysis along pseudotime generated by Monocle 2. **A:** Heatmap showing the expression pattern of 384 genes that were dynamically expressed along the pseudotime, in neuronal or RPC lineage. Blue means low expression, green means moderate expression and red means high expression. 384 genes were unbiasedly classified into 4 clusters based on their expression pattern. **B:** Gene ontology enrichment of genes in each of the 4 clusters and the corresponding representative genes. **Related to Figure 6.**

**Supplemental Tables**

Table S1: Top 100 genes with highest Spearman correlation with ASCL1 in RPCs. Related to Figure 2.

Table S2: Differential expressed genes in the 3 RPC clusters (Cluster 1, 3 and 5) identified in the initial unbiased clustering. Related to Figure 4.