iScience, Volume *27*

Supplemental information

Craniofacial chondrogenesis in organoids from human stem cell-derived neural crest cells

Lauren Foltz, Nagashree Avabhrath, Jean-Marc Lanchy, Tyler Levy, Anthony Possemato, Majd Ariss, Bradley Peterson, and Mark Grimes

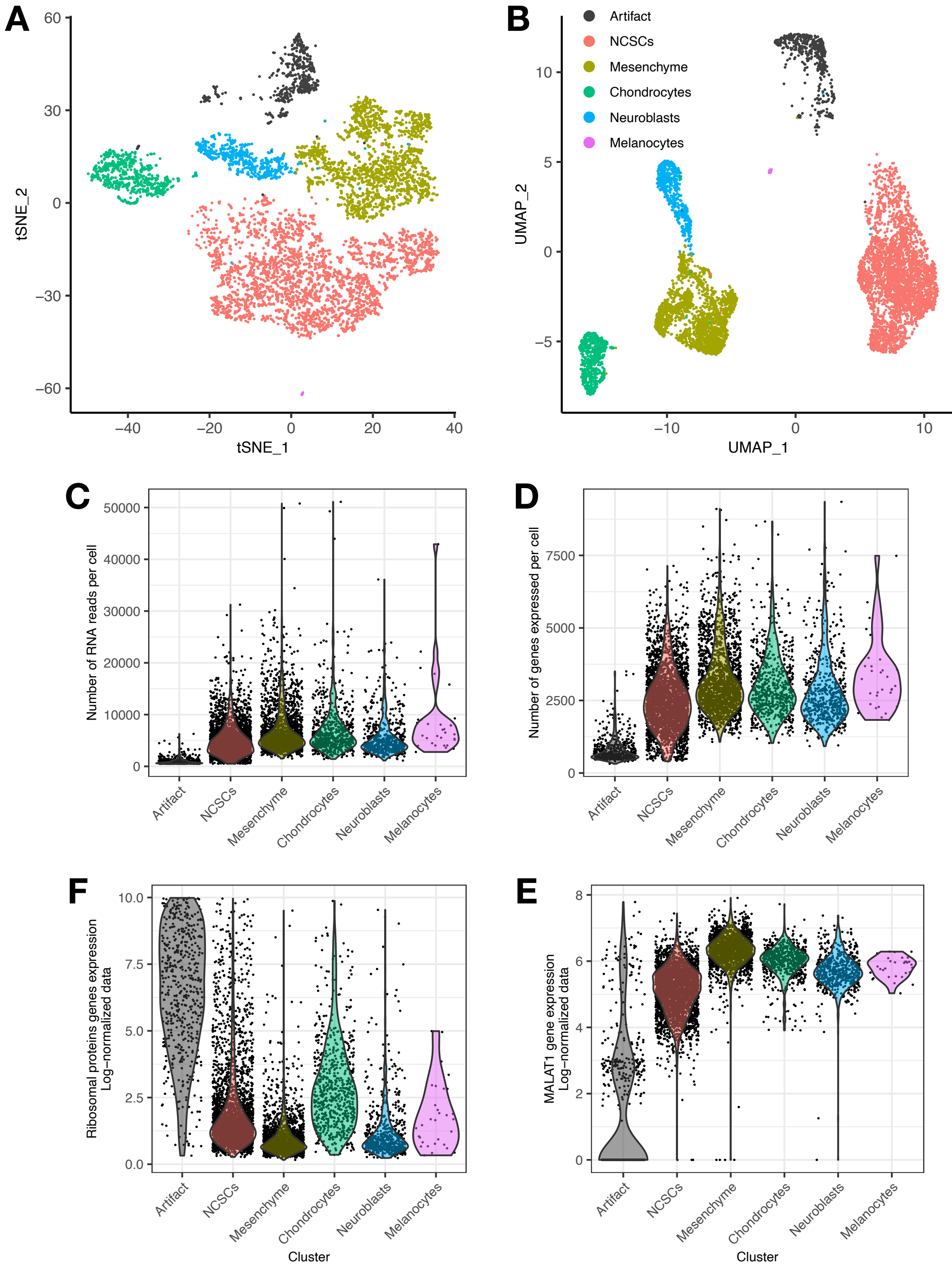
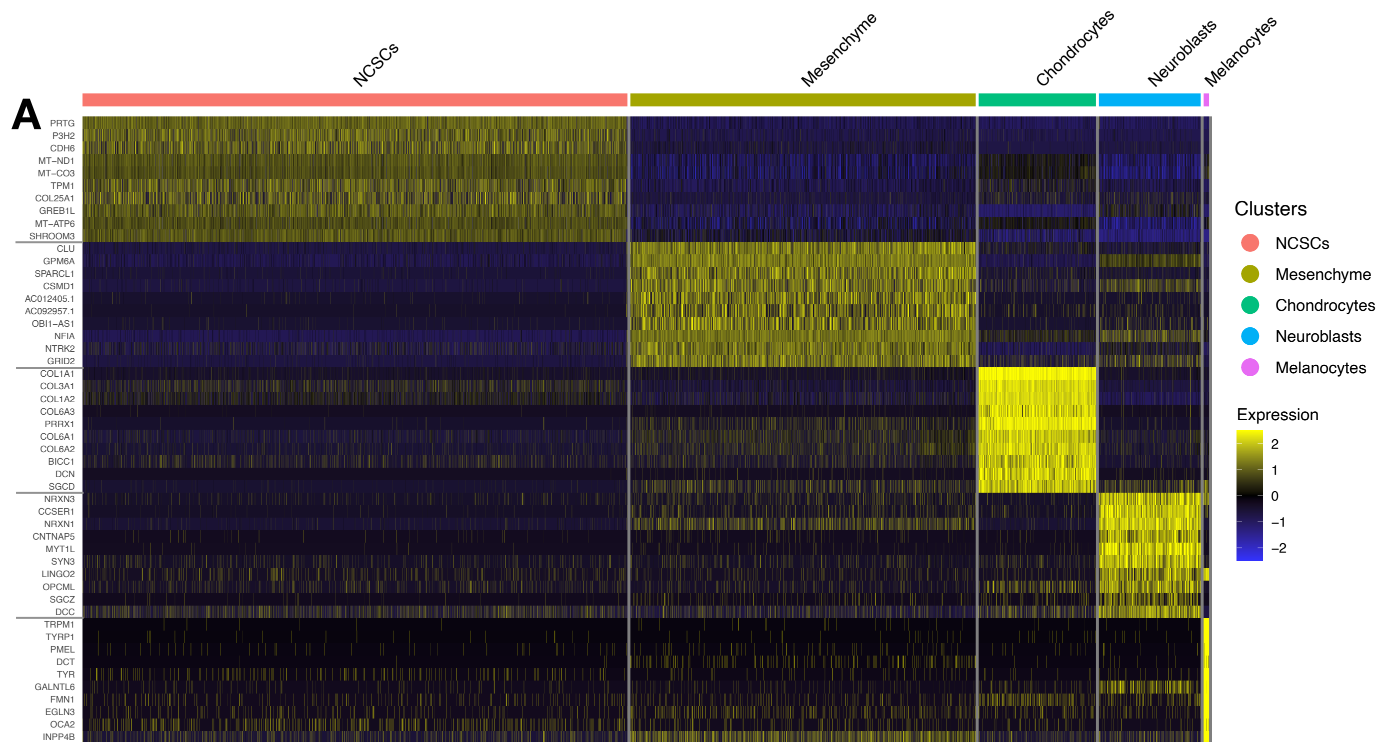


Figure S1: Initial clustering of the single nuclei RNA sequencing data.

The data from the single nuclei RNA sequencing of day1 and day 87 cartilage organoid samples were merged in R/Rstudio as a Seurat object using the Seurat v4.3 R package. The initial Seurat cluster analysis, tSNE (A) and UMAP (B) revealed one cluster with several abnormal features: the number of reads per cell (C), the number of expressed genes detected per cell (D), the percentile of ribosomal proteins genes expressed per cell (E), and the percentile of the nuclear marker MALAT-1 expressed per cell (F).



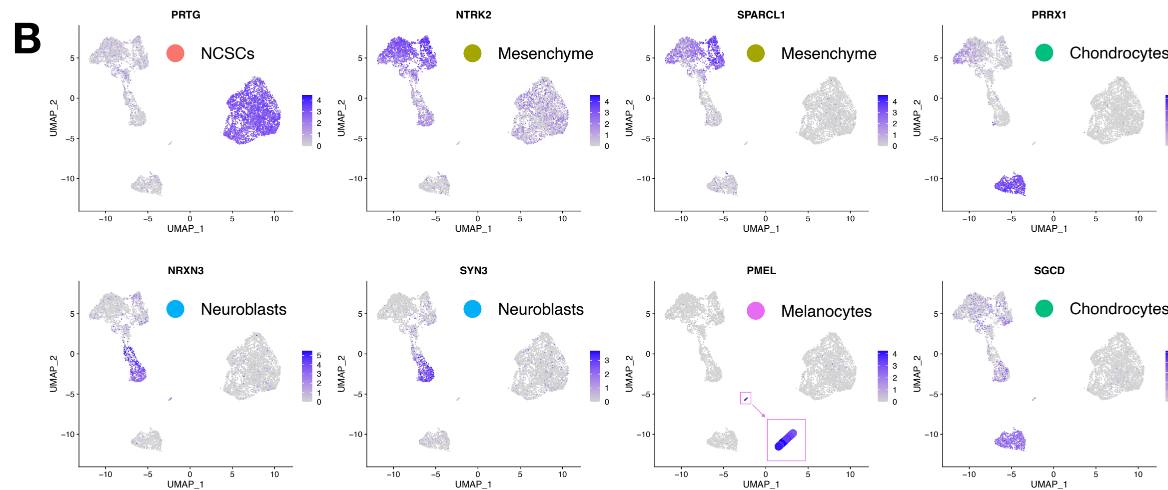


Figure S2. Top 10 markers in cells in differentiating cultures.

(A) Heatmap showing expression of top 10 markers in different cell populations from snRNASeq data. Colors (inset) indicate average $\log_2$ fold change.

(B) Gene expression patterns in individual cells from days 1 and 87 plotted together on UMAP projections as in Figure 2. Genes and clusters they mark are indicated.

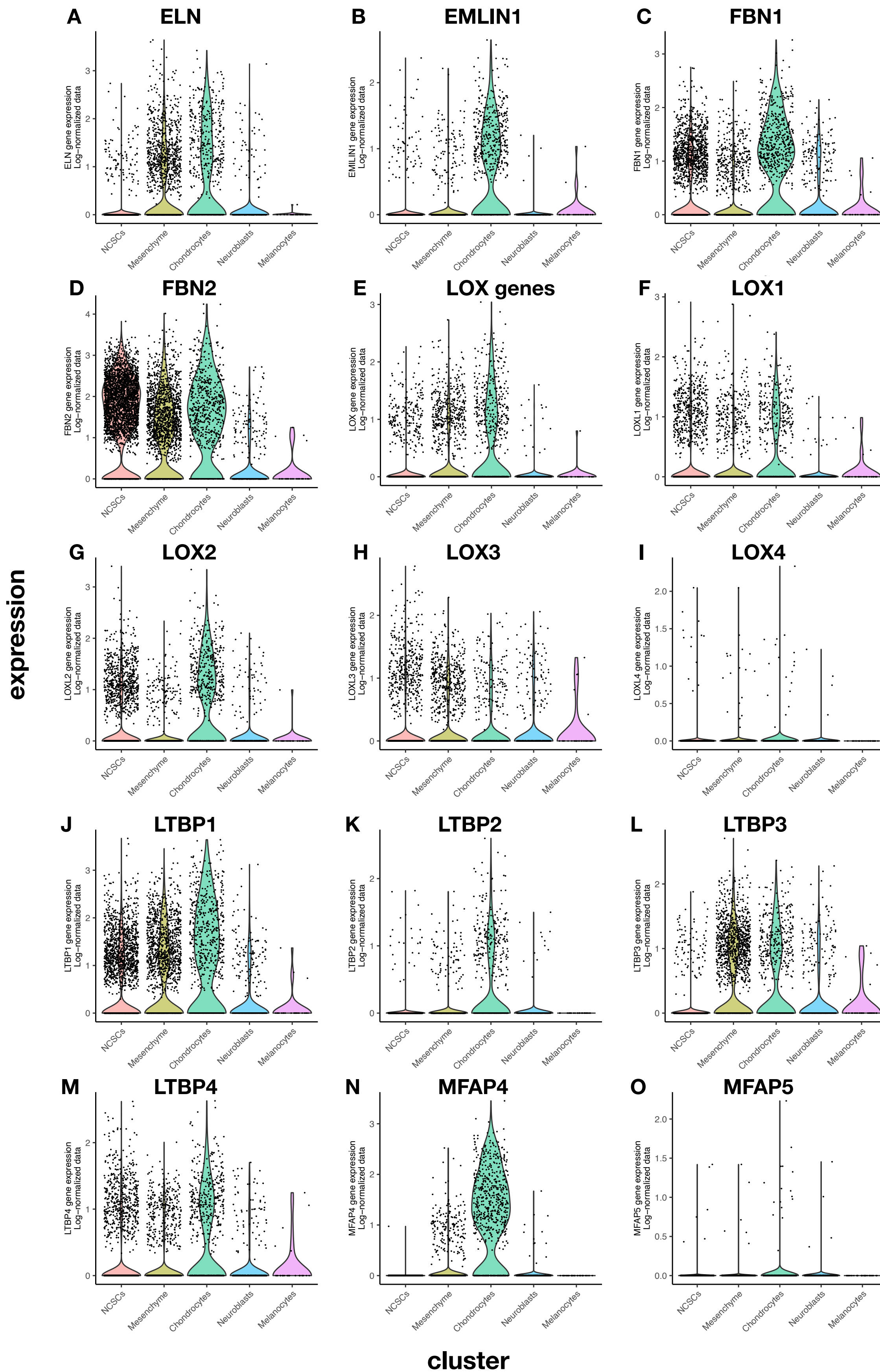


Figure S3. Elastic fiber marker expression in differentiating cultures.

Elastic fiber markers (ref. 17), graphed as violin plots from snRNASeq data. Clusters are assigned as in Figure 2A, B.

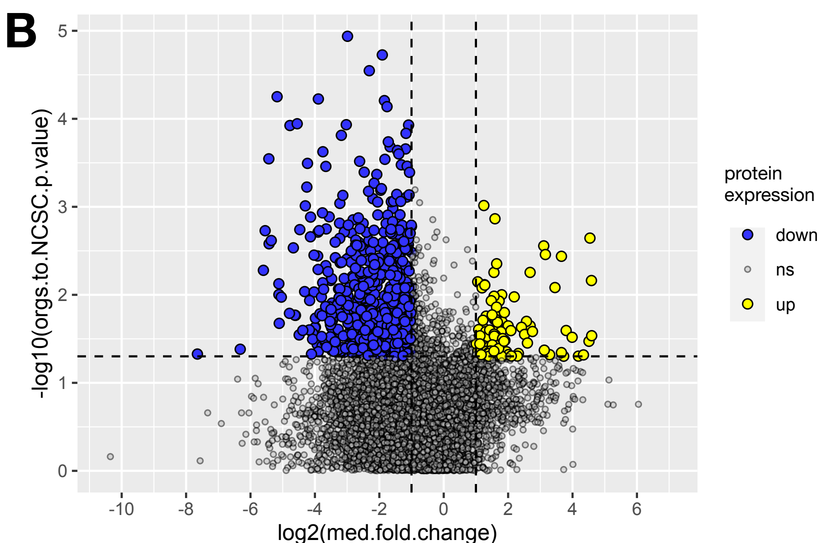


Figure S4. Chondrocyte markers in organoids vs. NCSCs.
(A) Heatmap showing expression of 131 chondrocyte markers in hESCs (WA09), neural stem cells (NSCs), neural crest stem cells (NCSCs) and organoids (ORGs). Median and maximum values were calculated from the row z-score values (yellow is up, blue is down, see inset). Data are ordered by hierarchical clustering (dendrogram). (B) Volcano plot comparing protein expression (TMT mass spectrometry data) in craniofacial cartilage organoids vs. NCSCs. (C) Heatmap of the 106 proteins highlighted in yellow in B with at least 2-fold increase and $p < 0.05$.

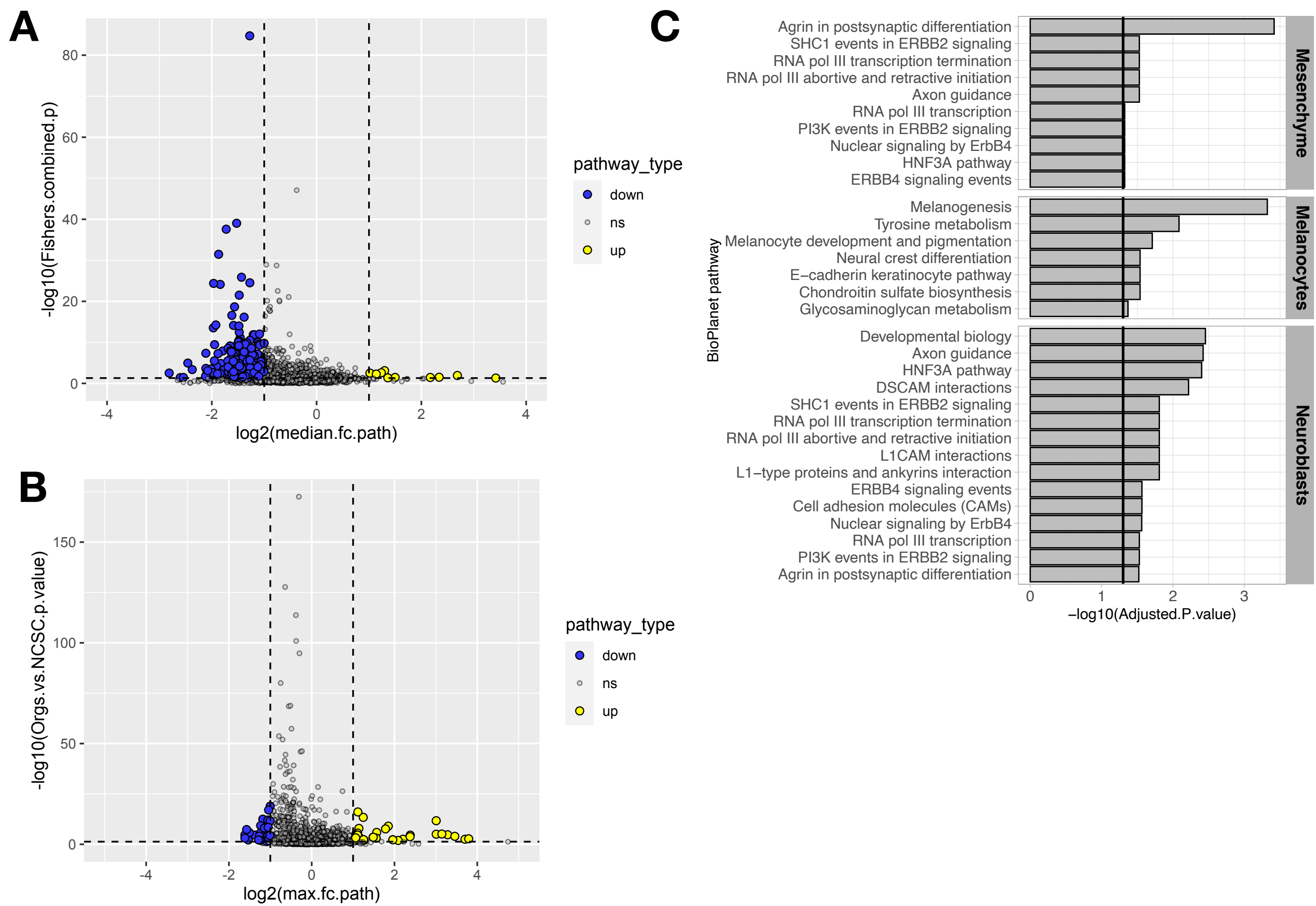


Figure S5. BioPlanet pathways enriched in craniofacial cartilage organoids.

(A, B) Using median (A) and maximum (B) total protein fold change expression in organoids compared to NCSCs, 11 (A) and 37 (B) pathways were significantly upregulated (yellow; Fishers truncated $p \leq 0.05$). Pathways identified in A are: *Inhibition of matrix metalloproteinases; Glycosphingolipid biosynthesis: ganglio series; Chondroitin sulfate biosynthesis; Dermatan sulfate biosynthesis; Matrix metalloproteinases; Ganglio sphingolipid metabolism; Small leucine-rich proteoglycan (SLRP) molecules; Gamma-carboxylation, transport, and amino-terminal cleavage of proteins; Neurotransmitter release cycle; AMPA receptor activation; Dopamine clearance from the synaptic cleft.*

(C) Pathways identified by EnrichR in mesenchyme, melanocyte, and neuroblast snRNASeq cell clusters as for chondrocytes in Figure 4B.

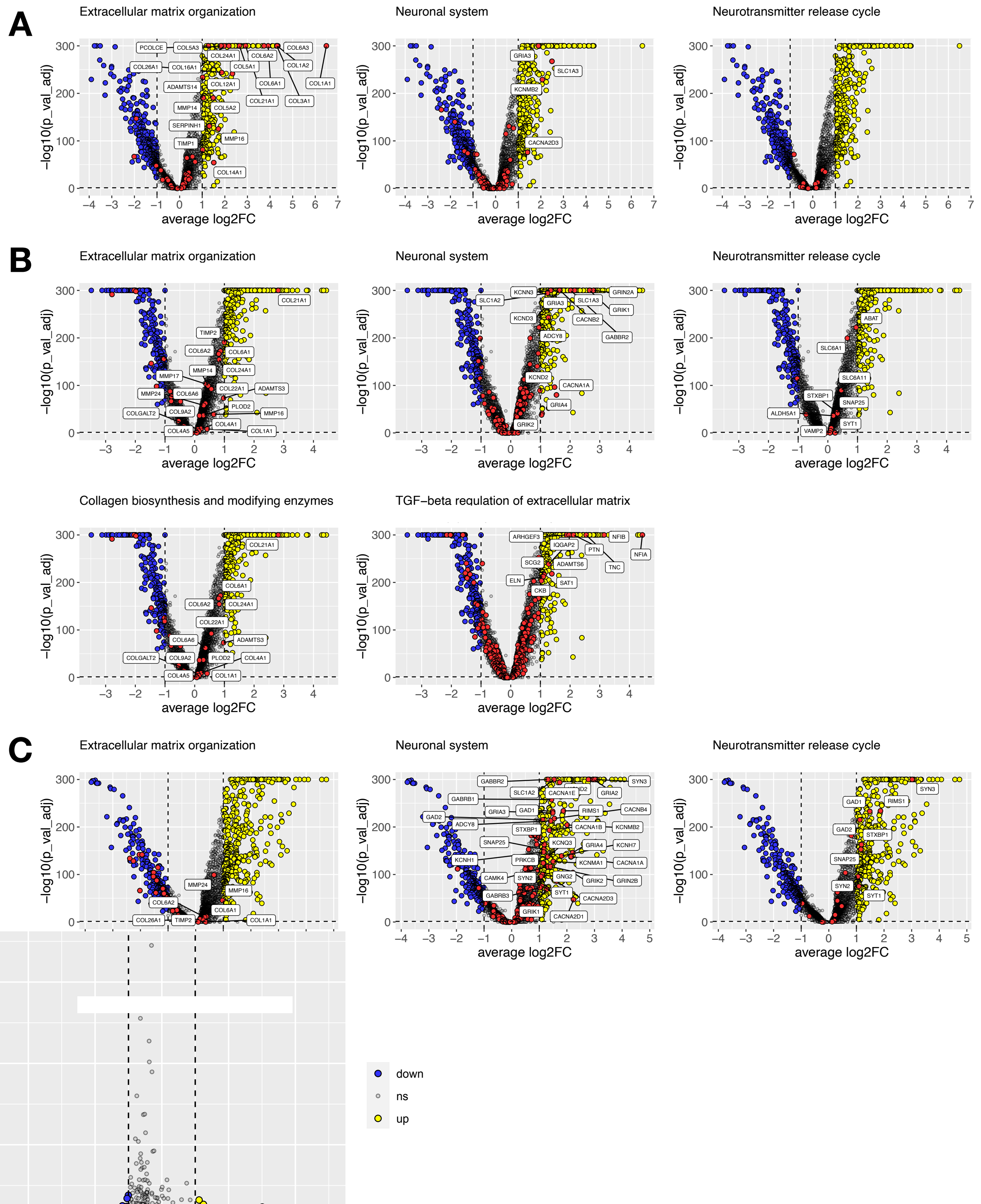


Figure S6. Volcano plots of BioPlanet pathway genes in snRNASeq cell clusters.

Volcano plots of genes from indicated pathways in chondrocytes (A), mesenchyme cells (B), and neuroblasts (C).

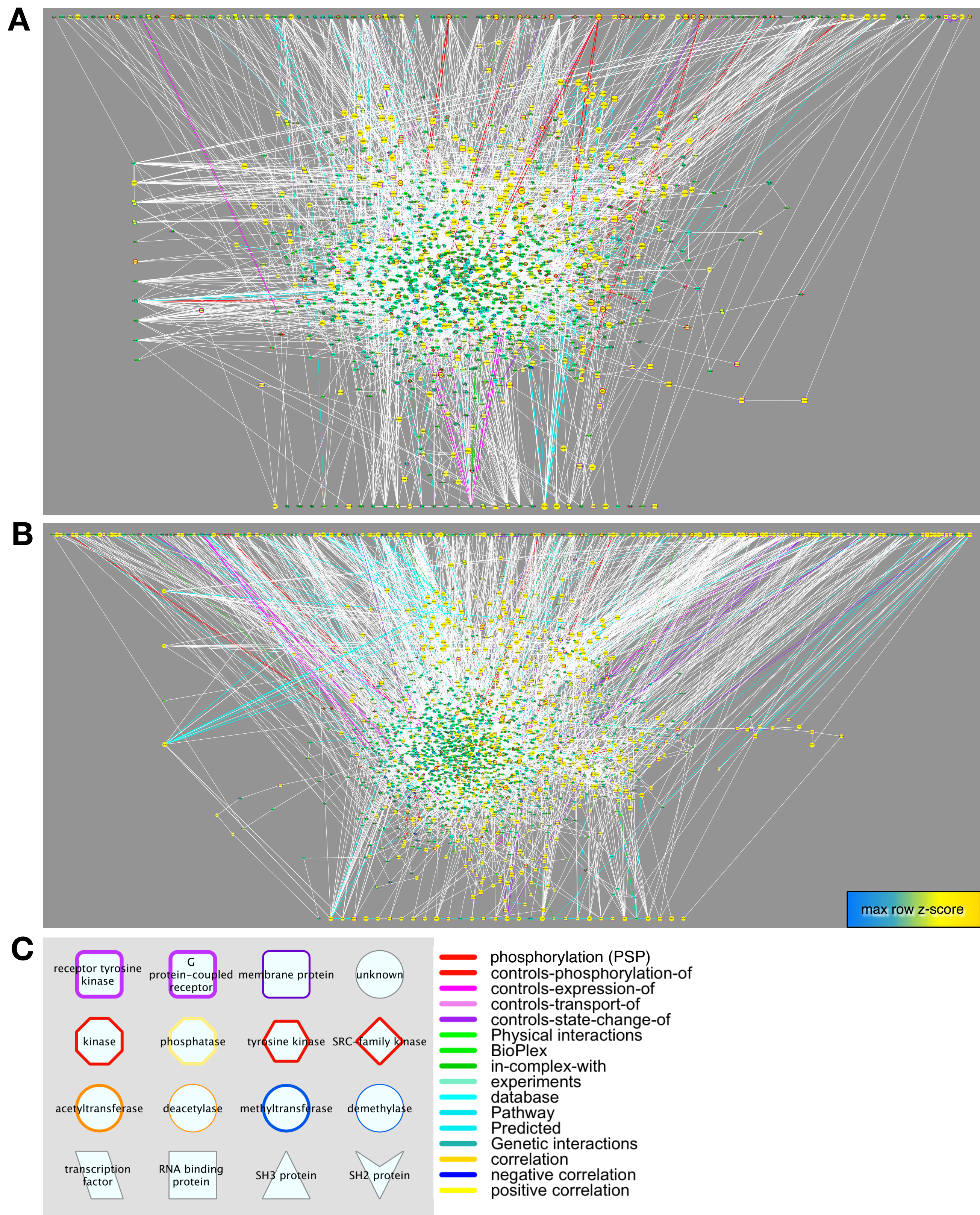


Figure S7. Cluster-filtered networks (CFNs) of proteins that interact with two pathways.

(A) All shortest paths between pathways *Wnt signaling pathway* (top row) and *Neural crest differentiation* (bottom row; proteins in both pathways are aligned vertically at the left). (B) All shortest paths between pathways *TGF-beta regulation of extracellular matrix* (top row) and *Glycosaminoglycan metabolism* (bottom row; common nodes at left). Node size and color represents the max $\log_2$ fold change in organoids over NCSCs (color scale bar shown in B). Node border and shape and edge colors are defined in C; merged edges from different databases are colored white (ref. 9).

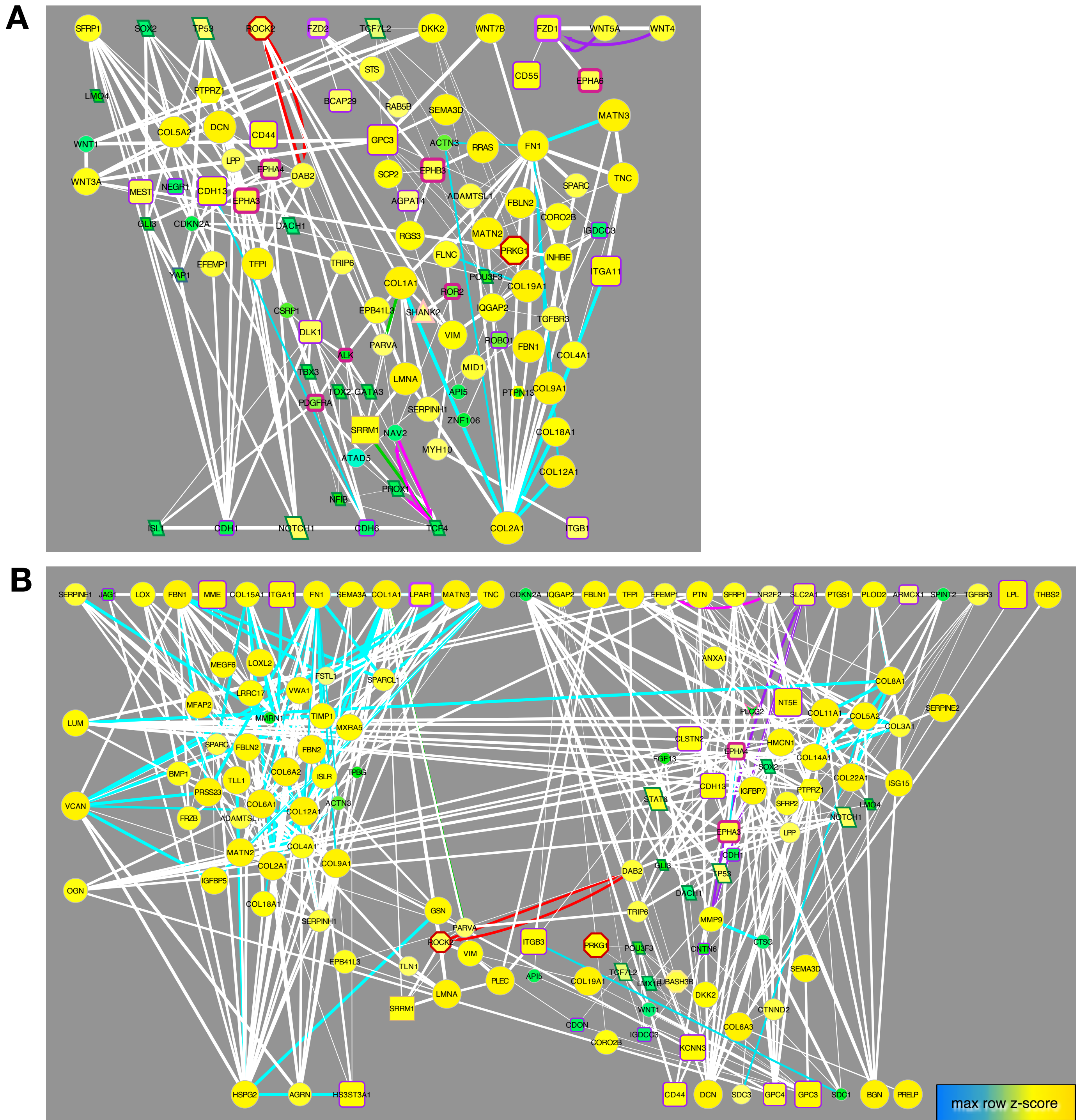


Figure S8. Cluster-filtered networks (CFNs) of selected proteins that interact with two pathways.

(A) Selected interactions between pathways *Wnt signaling pathway* (top row) and *Neural crest differentiation* (bottom row; proteins in both pathways are aligned vertically at the left). Ligand-receptor pairs identified in chondrocytes from proteomic and snRNASeq data and their nearest neighbors were selected from all paths between pathways of the CFN derived from proteomic data (Figure S7A). (B) Friends of mutual friends from pathways *TGF-β regulation of extracellular matrix* (top row) and *Glycosaminoglycan metabolism* (bottom row; common nodes at left). "Mutual friends," defined as proteins that connect to at least one member of each pathway in the CFN (Figure S7B), were selected along with proteins that connect to them. Node size and color represents the max log_2 fold change in organoids over NCSCs (color scale bar shown in B). Node border and shape and edge colors are defined in Figure S7C.

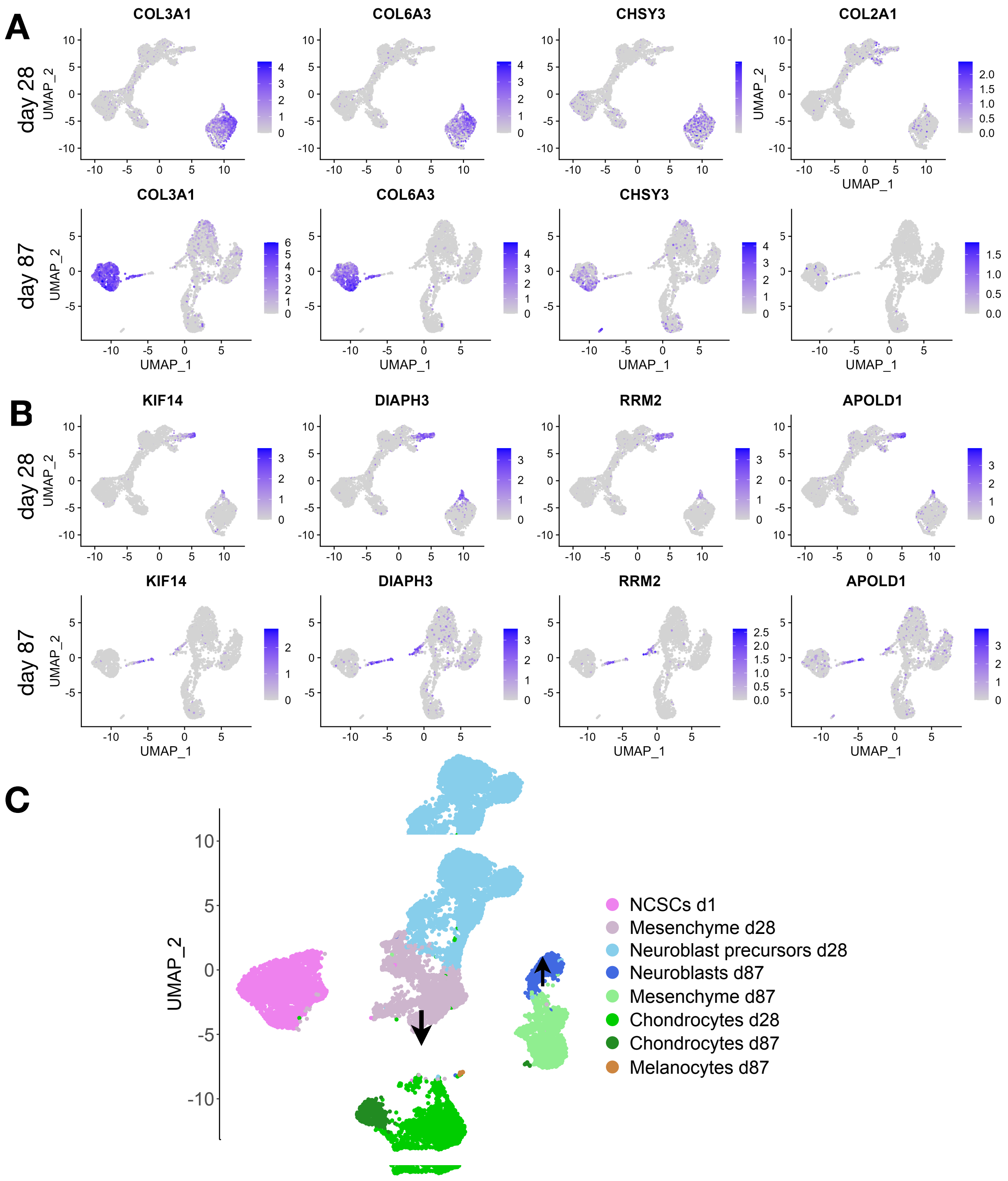


Figure S9. UMAP projections of markers in cells from snRNASeq data 28 and 87 days after NCSC sorting.

(A) Markers expressed in Chondrocytes (see Figure 7A): collagens COL3A1 and COL6A3, CHYS3 (Chondroitin Sulfate Synthase 3). COL2A1 expression is low in mesenchyme cells and chondrocytes at these time points.

(B) Cell cycle markers of proliferating cells (see figure 7B)*:* KIF14 (kinesin family protein expressed in M phase), DIAPH3 (involved cytokinesis, cell movement, adhesion), and RRM2 (ribonucleotide reductase). APOLD1 (endothelial cell early response protein) has a similar distribution in proliferating cells. N=3.

(C) Proposed map of NCSC differentiation. snRNA-seq data from days 1, 28, and 87 were combined, clustered and graphed using UMAP. Arrows indicate hypothesized differentiation trajectories.