

# Further Dimensionality Reduction on Embryonic Liver Stem Cells

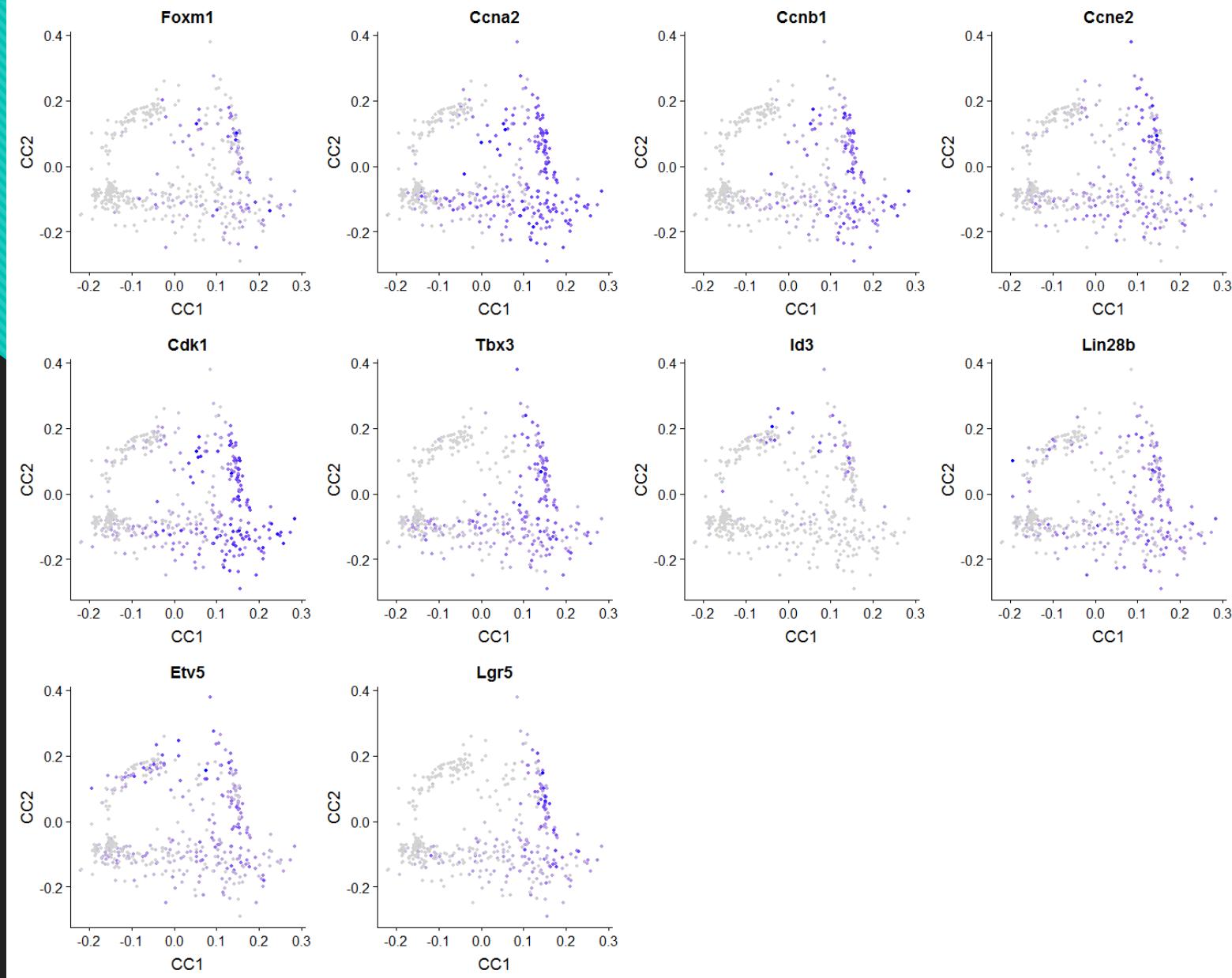
Nick Wawee

# Details

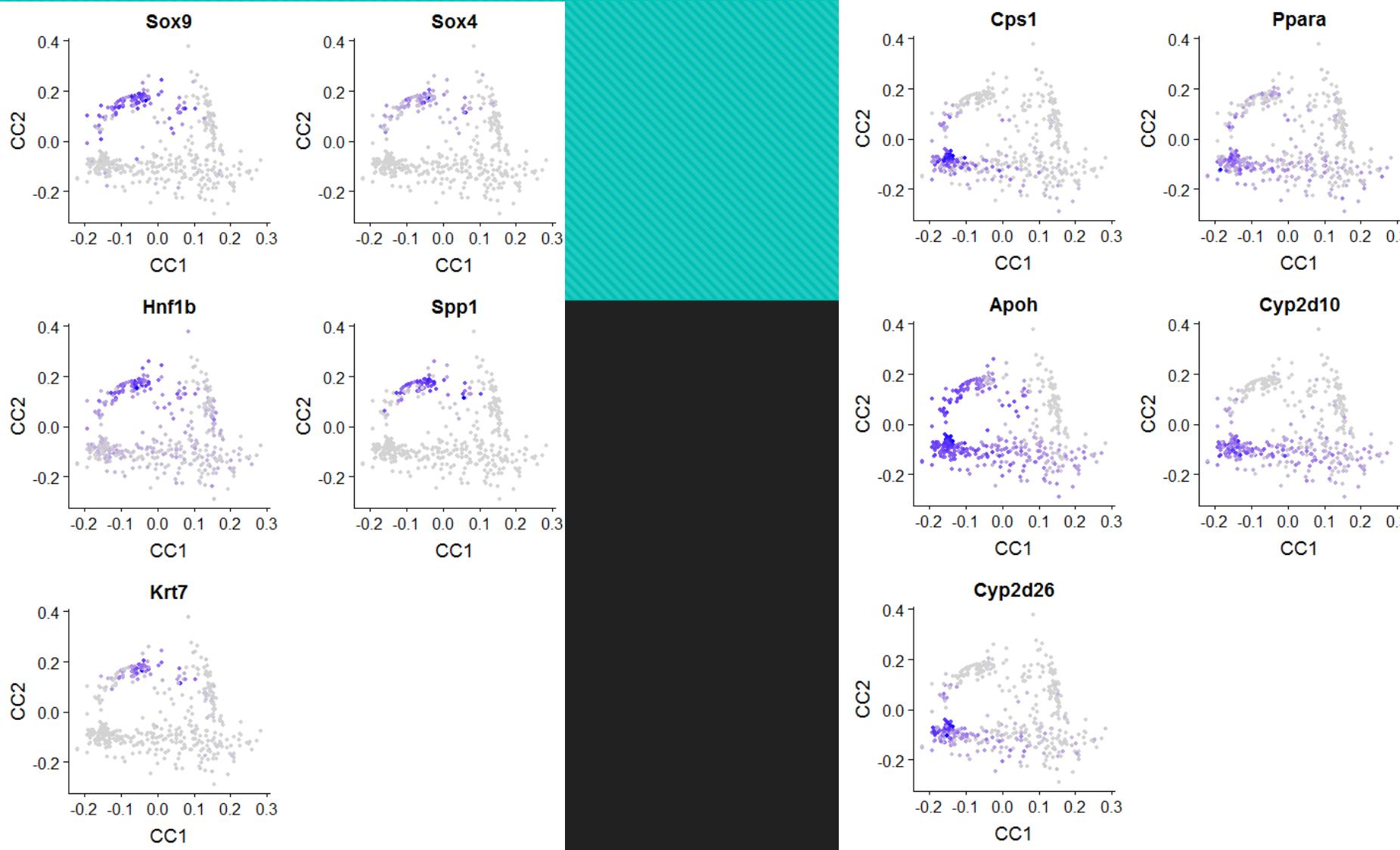
- Absolute read count data from Yang et al. underwent a CCA and a tSNE to identify differentially expressed genes between cell clusters
- Procedure followed Seurat's tutorials available on the Satija Lab's website
  - Used multi.cca on each embryonic day designation
- tSNE and CCA plots were generated depicting gene expression levels and embryonic day designation

# CCA

Interesting structures in gene expression were found using the two dimensional representation of the cells, CC1 and CC2



Hepatoblast/Cell Cycle Genes

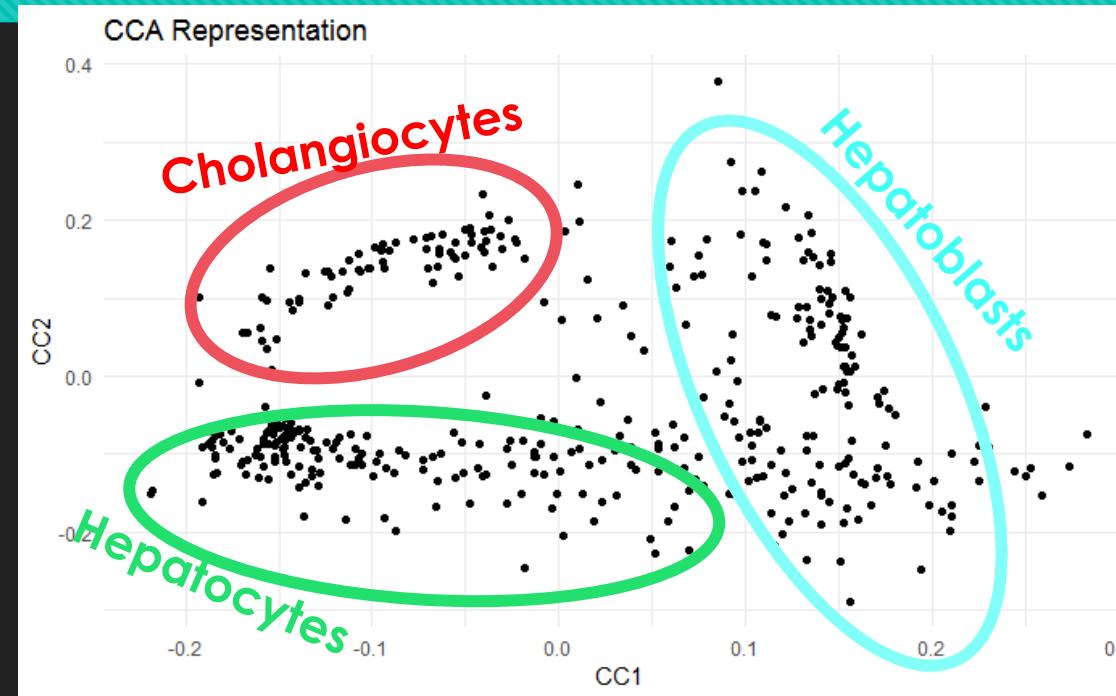


Cholangiocyte Genes

Hepatocyte Genes

# CCA Plot

Cell Specification ↑

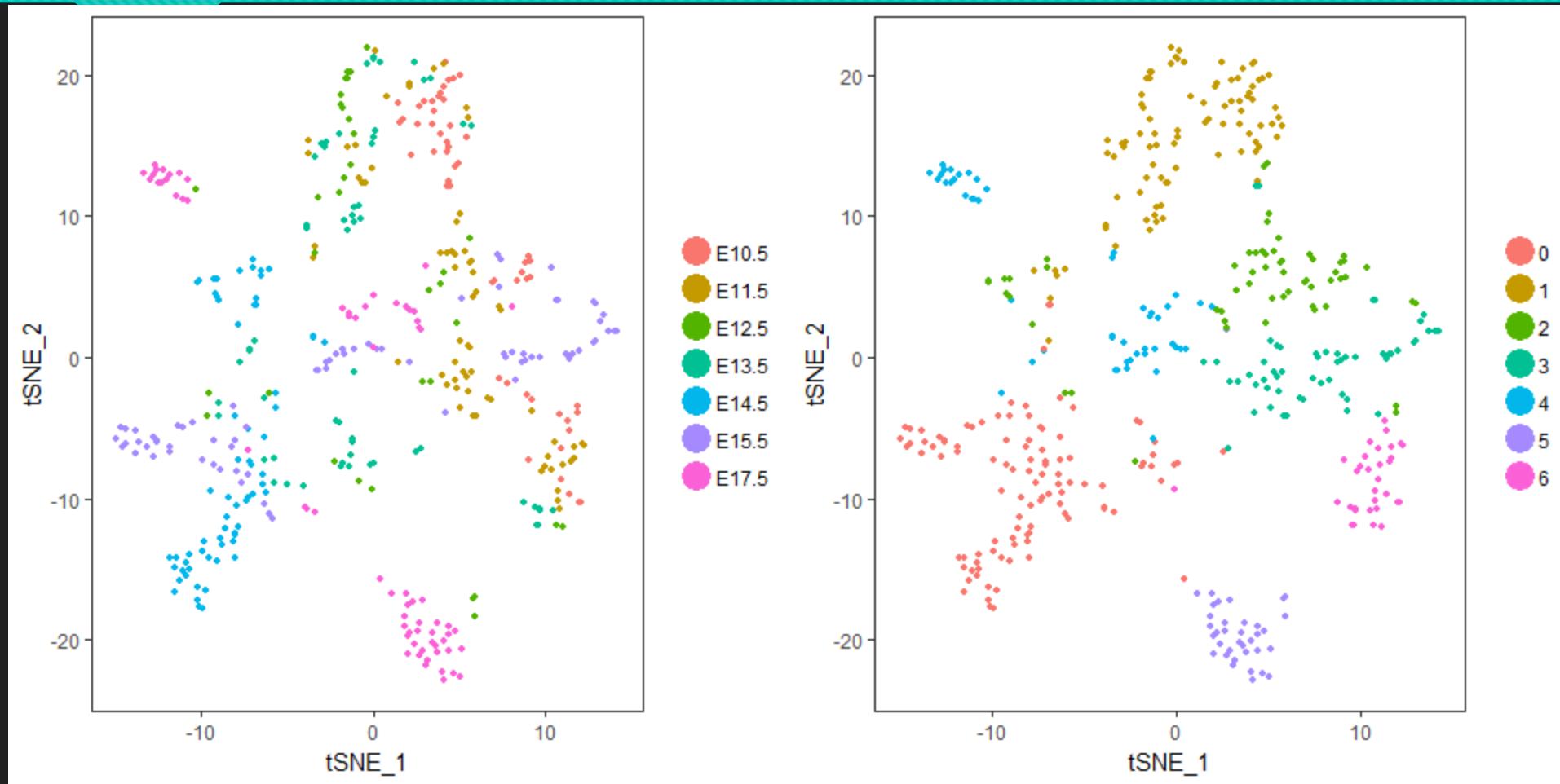


← Differentiation

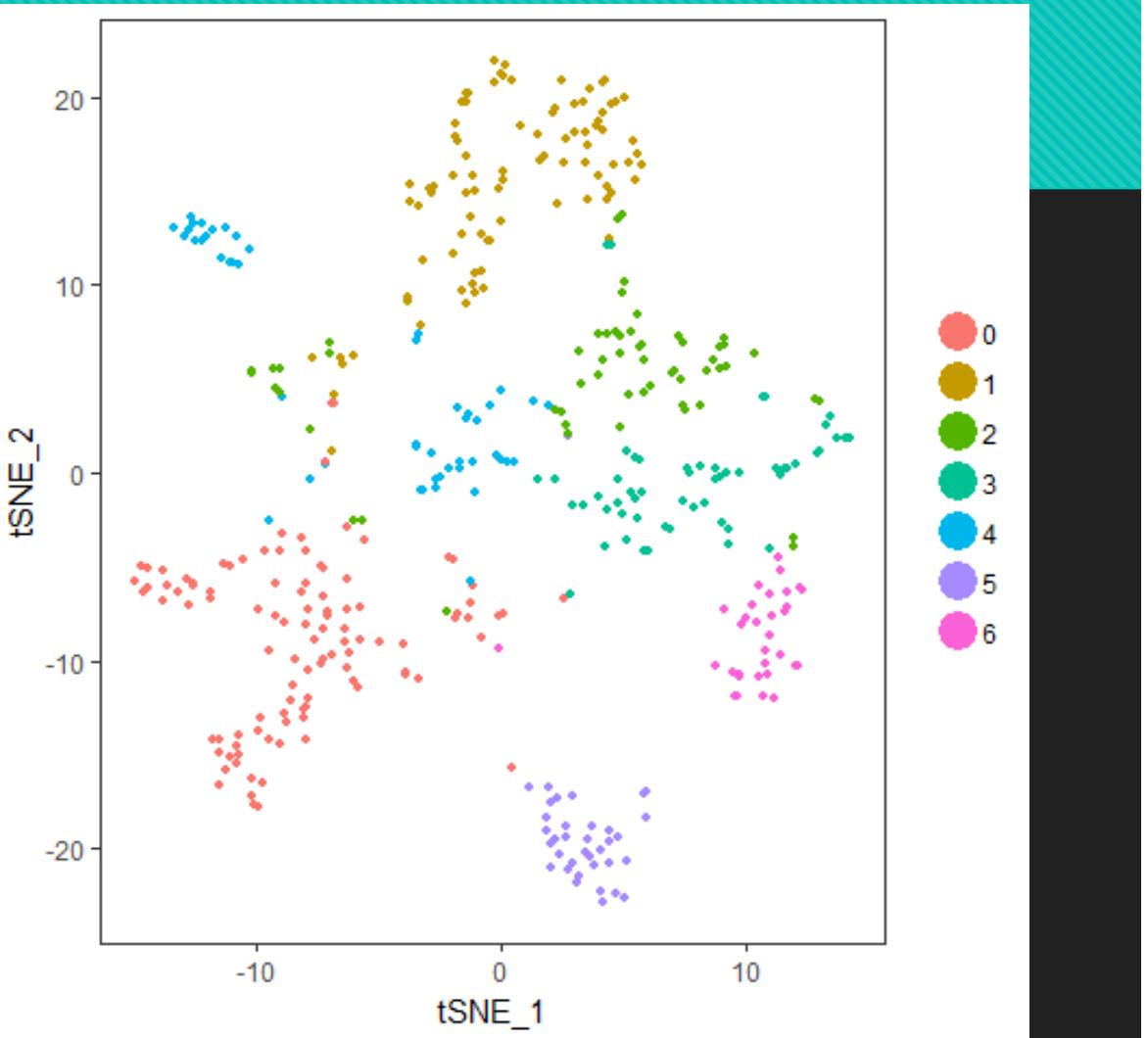
Genes can be classified as either differentiation or specification driven

# tSNE Plots

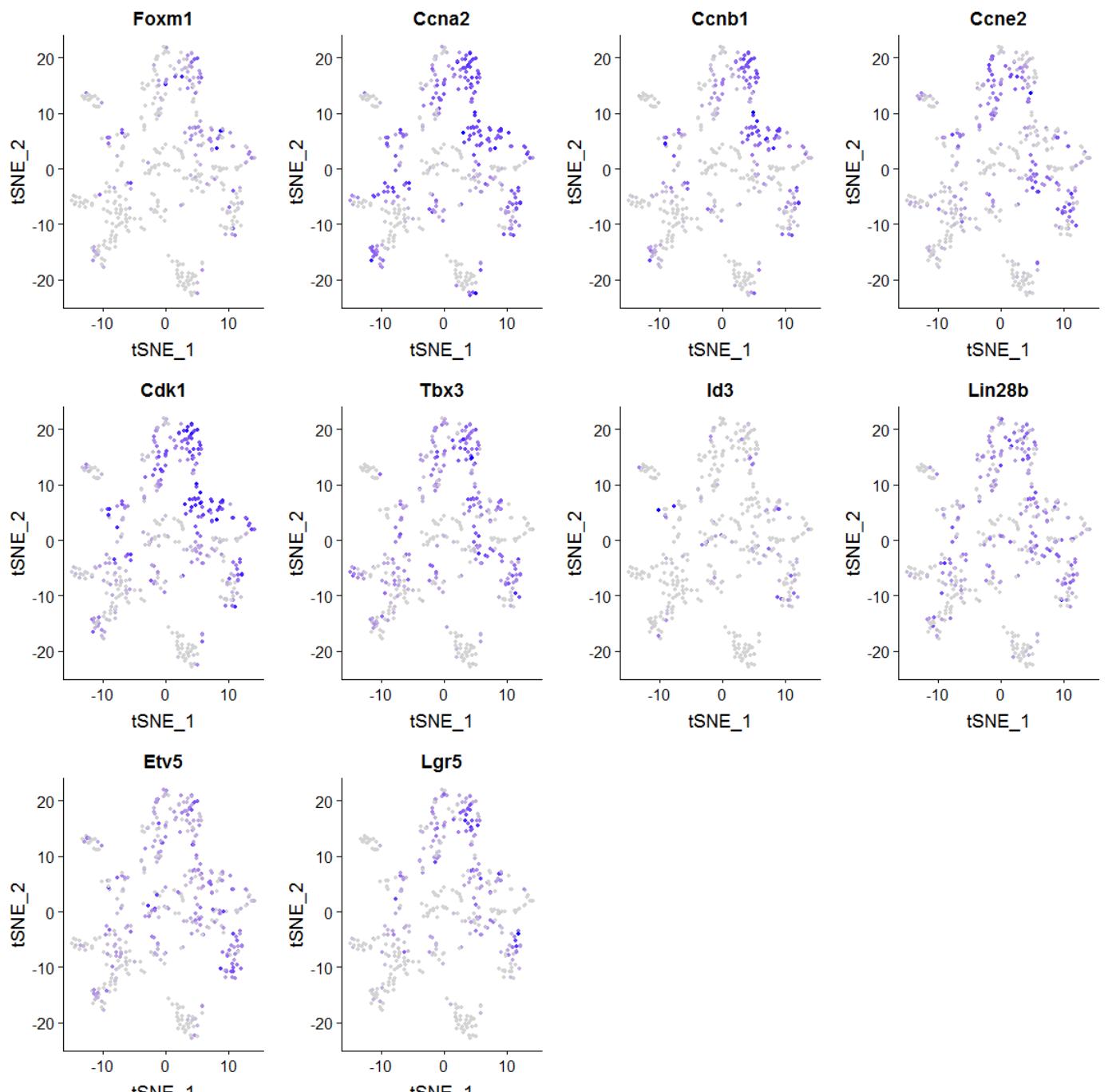
# Embryonic Designation and tSNE clustering



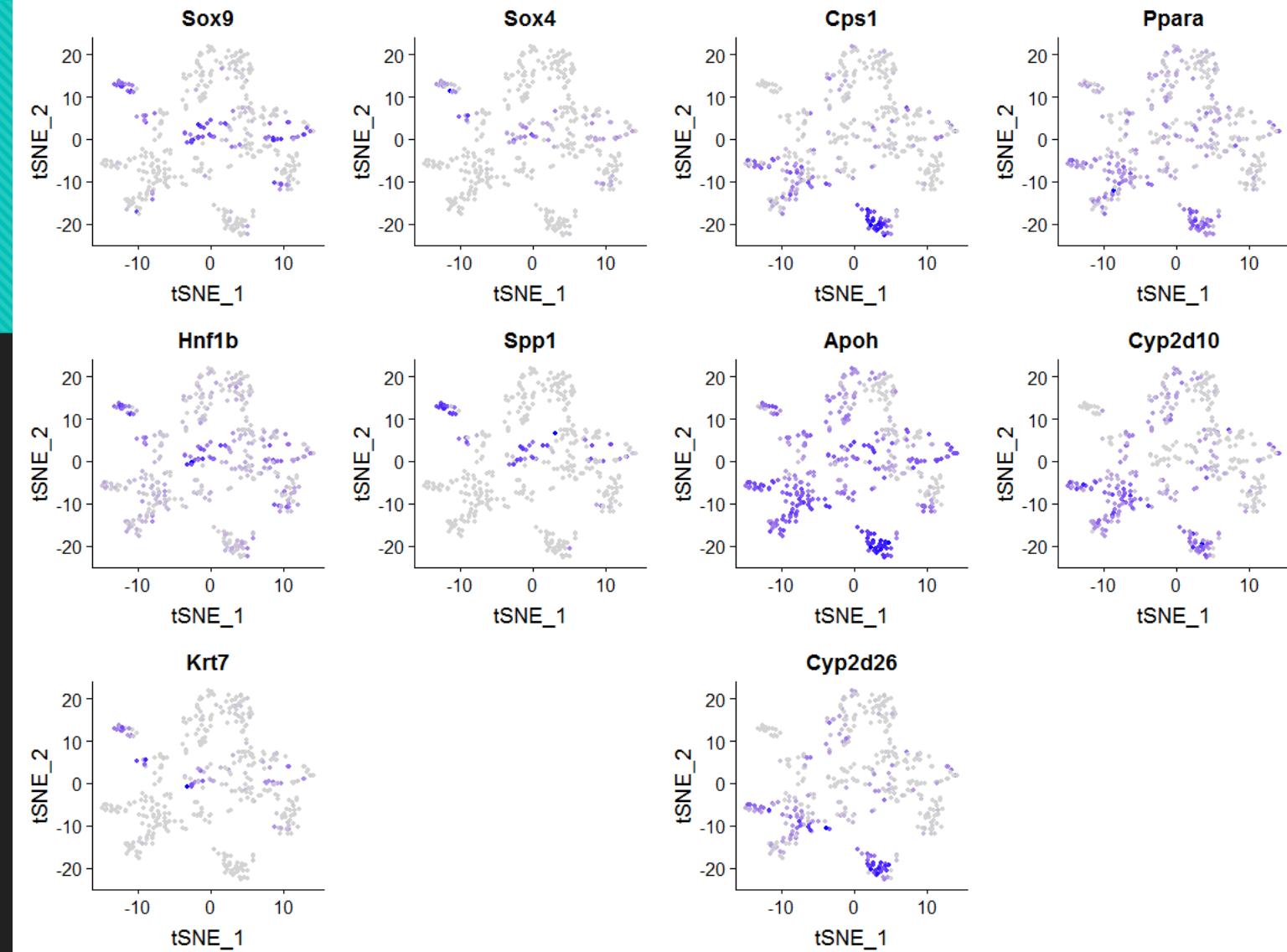
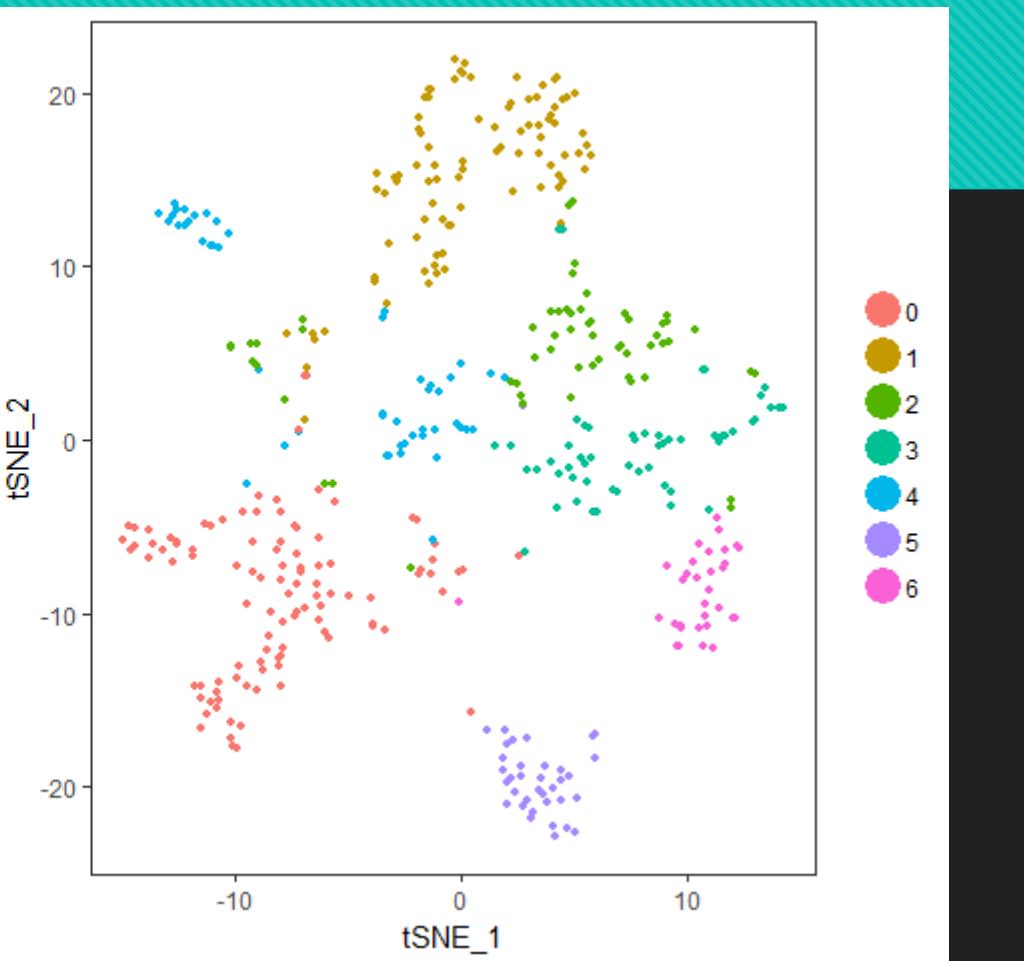
# Hepatoblast Genes



Hepatoblast and cell cycle gene markers are present in clusters 1,2,3, and 6. (some in 0)



# Cholangiocyte and Hepatocyte Genes



Cholangiocyte markers are present in cluster 4, some expression in clusters 3 and 6

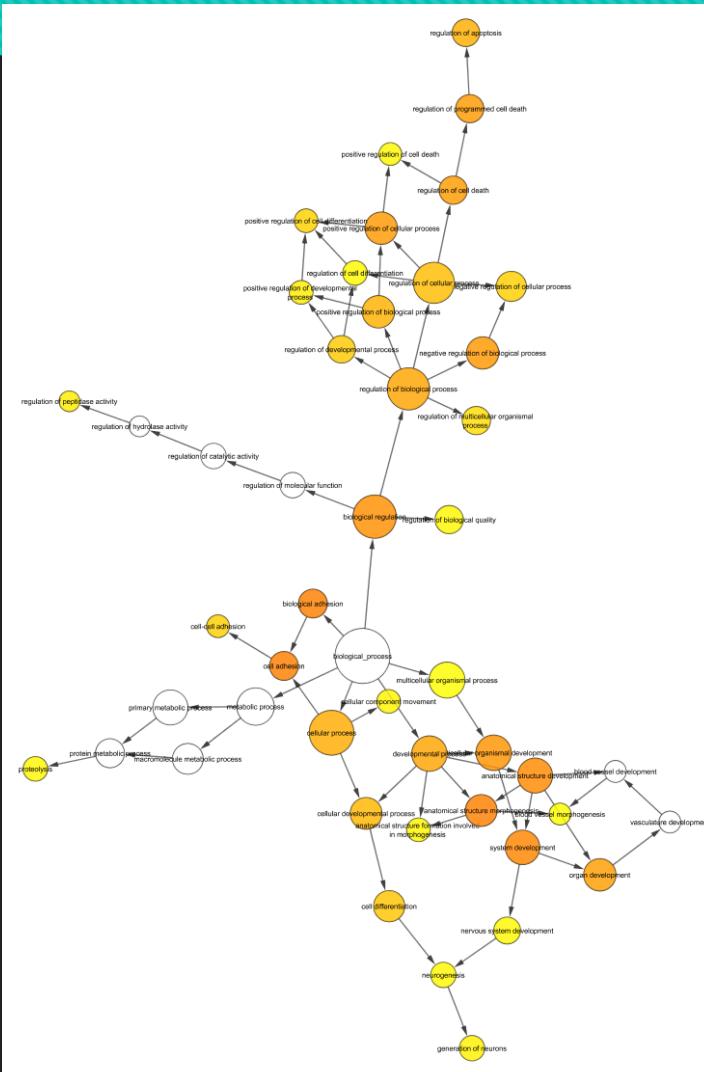
Hepatocyte markers are present in clusters 0 and 5

# Differentially Expressed Genes and GO Analysis

- Wilcoxon Rank Sum Test (non-parametric t-test) was ran on the fold changes between the hepatoblast and differentiated clusters to see which genes were the most differentially expressed
  - Cholangiocyte Differentiation Groups: 1,2,3,6 → 4
  - Hepatocyte Differentiation Groups: 1,2,3,6 → 0,5
- GO analysis in BiNGO was ran on top 100 genes with lowest P-value in each set of markers
  - Hypergeometric test
  - Bonferroni Family-Wise Error Rate (FWER) Correction
  - Level=0.05, so 95% certain that the over-represented categories contain no false positives
- Each set of markers along with fold-change and p-values are saved in .xlsx format

# Cholangiocyte Differentiation GO Plots

# Cholangiocyte Differentiation BiNGO Map

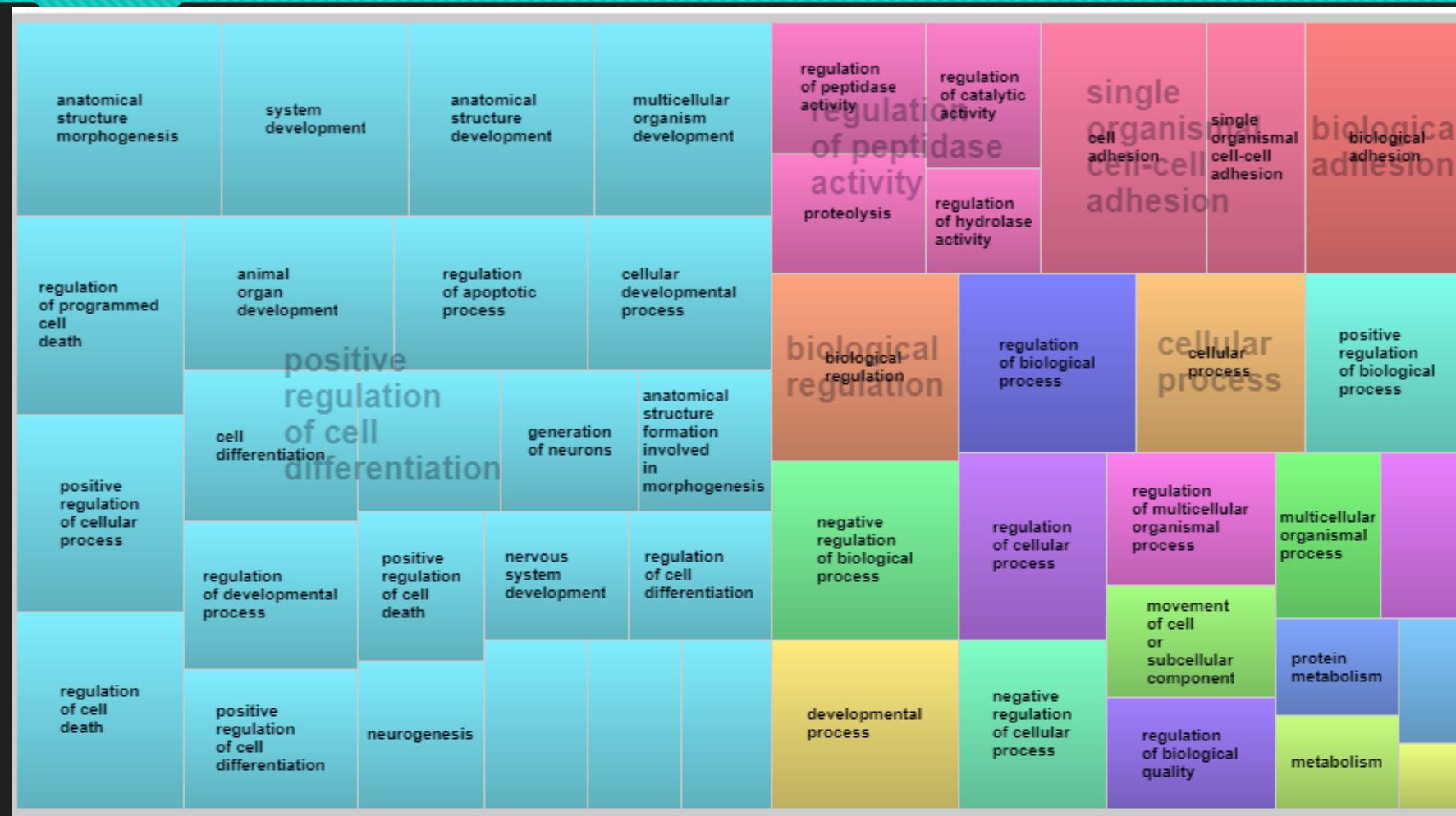


5.00E-2

< 5.00E-7

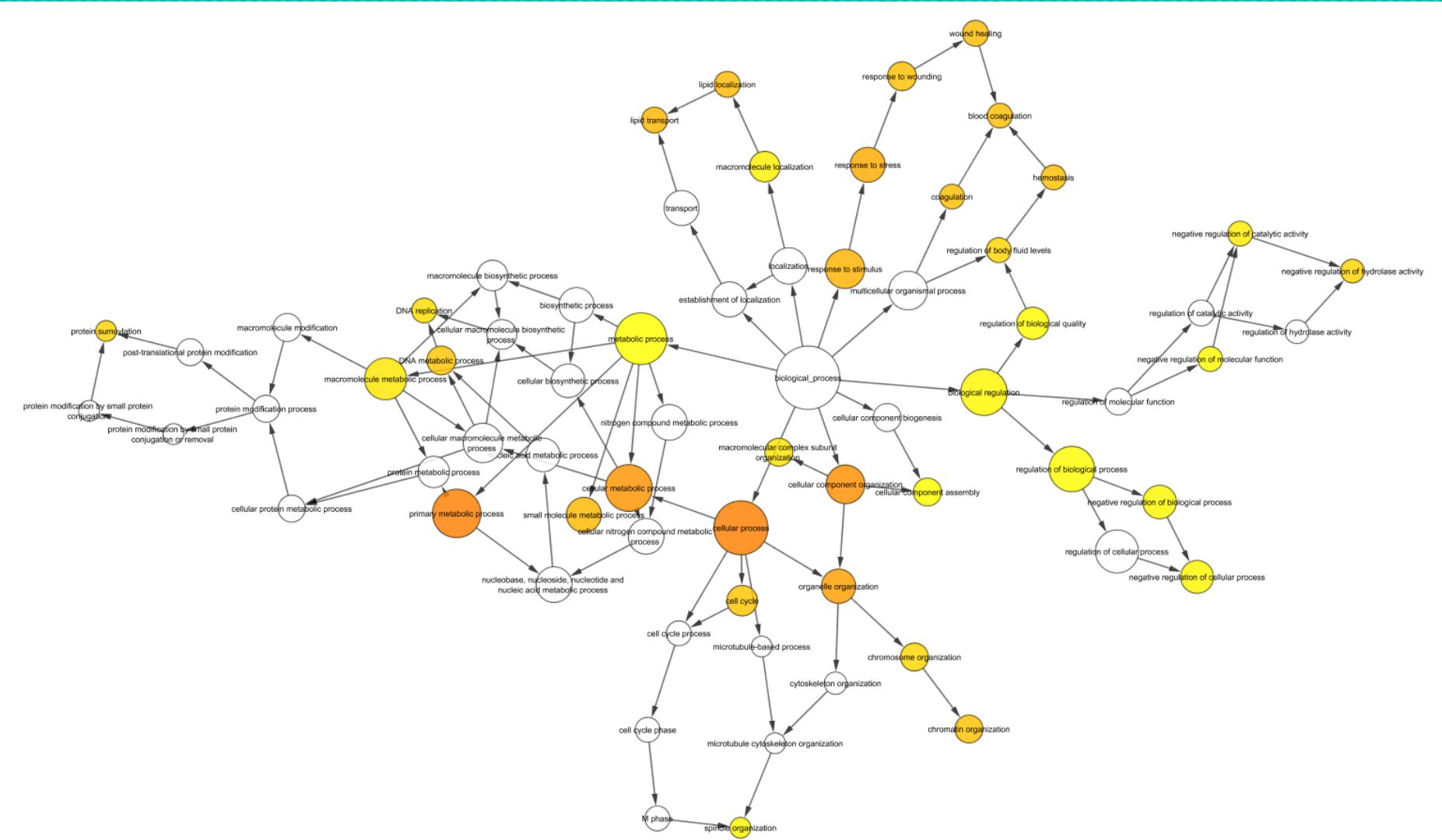


# Cholangiocyte Differentiation Treemap

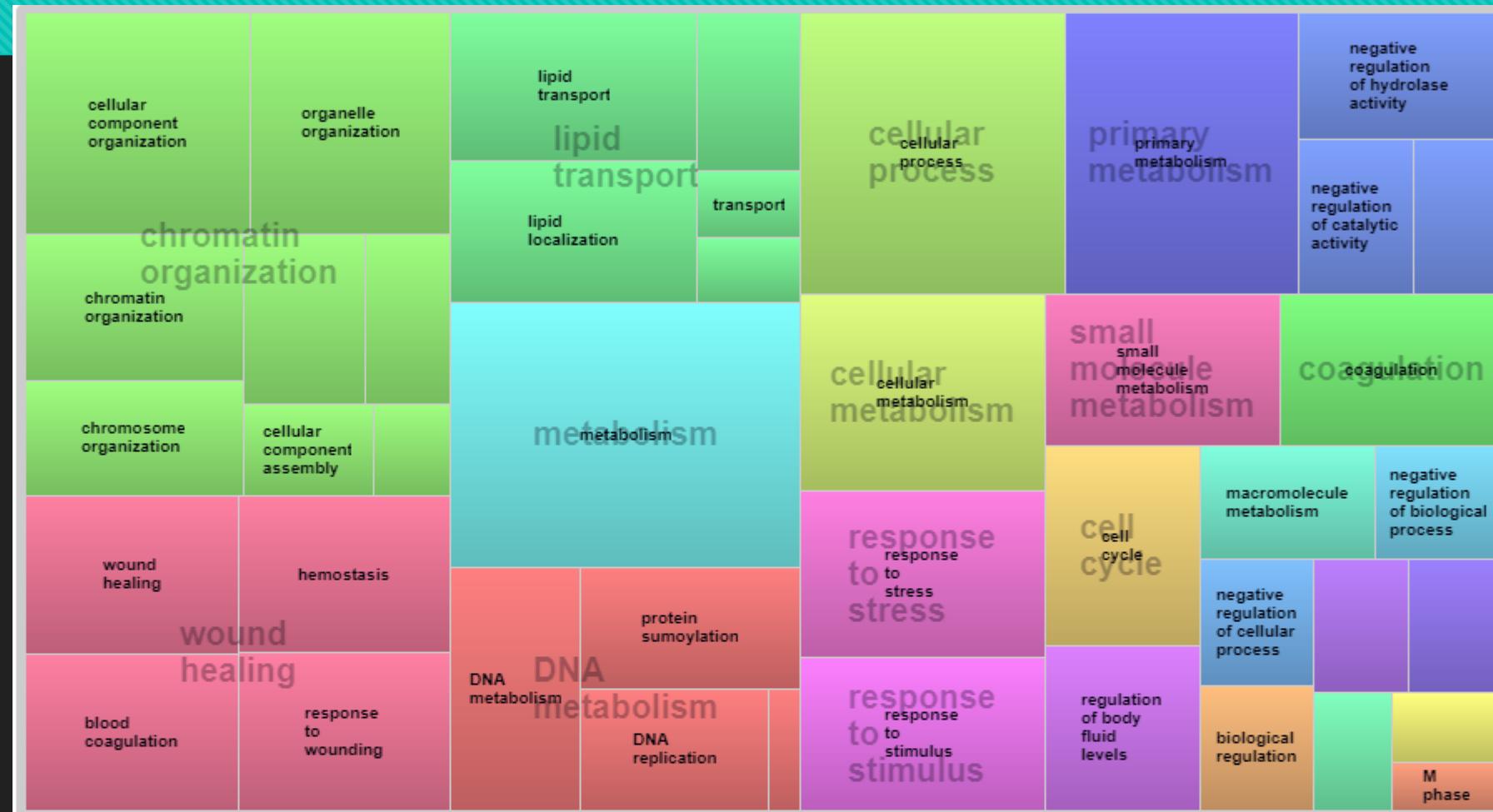


# Hepatocyte Differentiation GO Plots

# Hepatocyte Differentiation BiNGO Map



# Hepatocyte Differentiation Treemap



# Potential Further Work

- Somehow emulate in single cell gene expression using bulk-RNA seq dosage data?
- Cross-validation with monocle findings
  - Manifold plots with pseudotime
- Create aggregate “cholangiocyte/hepatocyte gene expression” cell manifold plots
- Analyze top gene loading values along CC axes and compare to the top differentially expressed the Seurat package outputs
  - Compare to other genes identified as differentially expressed in combined datasets
- UMAP implementation (in the works)
- Incorporation of other single-cell embryonic dataset