Fig. 1

D_ata = A_mplitudes x P_atterns

(gene weights for each biological process)   (biological process in each sample)

$$D = A \times P$$

The algorithm looks for patterns that define the biological activity in a set of samples (e.g., a collection of tumors) and the amplitude at which each gene contributes to that biological process.
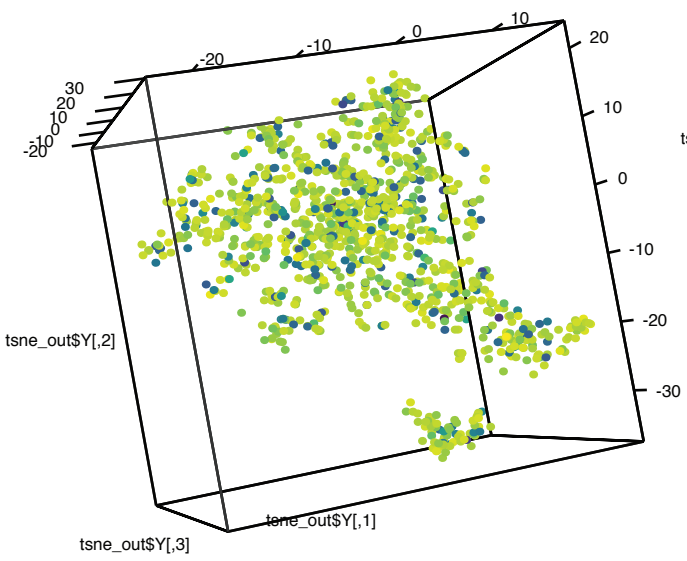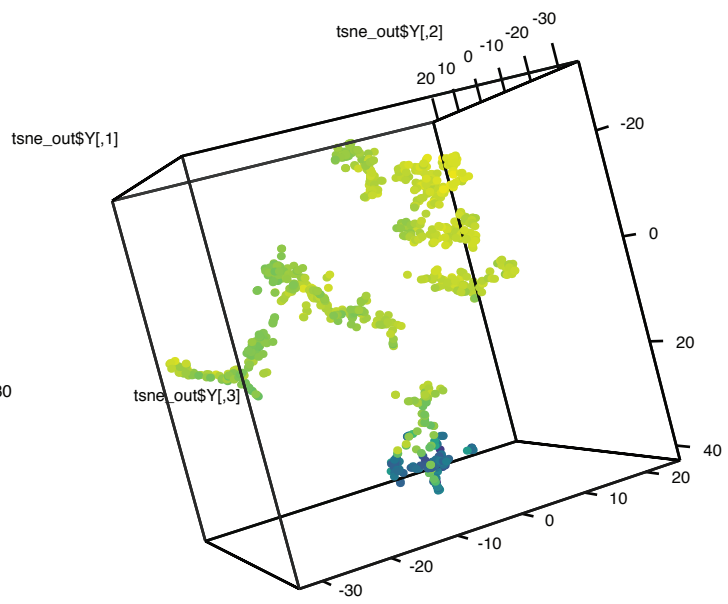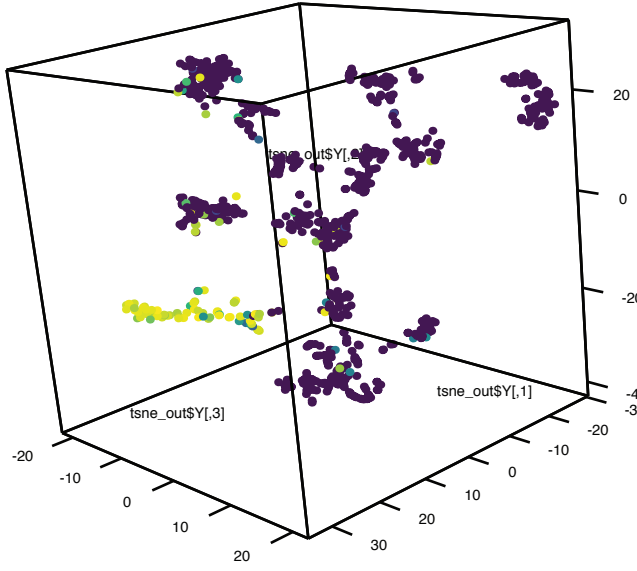
Fig. 3

Fig. 2

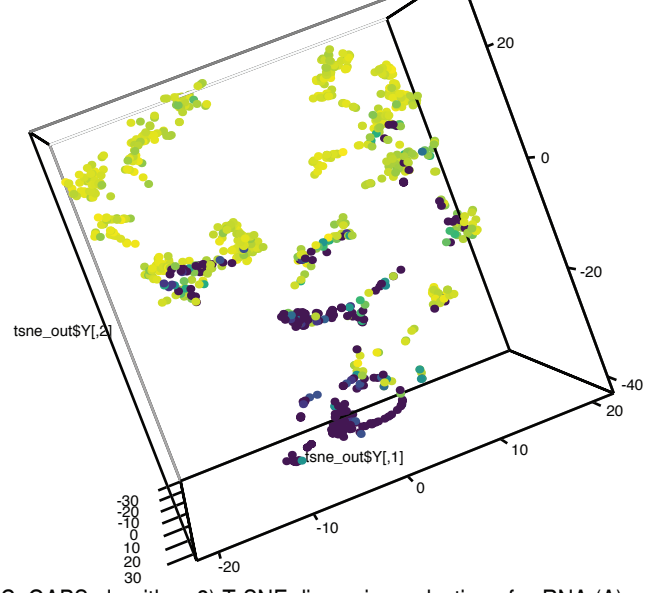A. T-SNE of raw scRNAseq data colored by # genes expressed

B. T-SNE of CoGAPS patterns for scRNAseq colored by # genes expressed

C. T-SNE of CoGAPS patterns colored by Crx expression
(Crx is a marker of photoreceptors)

D. T-SNE of CoGAPS patterns colored by Ccnd1 expression
(Ccnd1 is a marker of cell cycle)

1) Schematic representation of non-negative matrix factorization implemented in the CoGAPS algorithm. 2) T-SNE dimension reduction of scRNA (A) and CoGAPS pattern weights for smart seq data (B-D) of progenitor enriched cells from days E14, E18, and P2 in mouse retina developement. Panels (A and B) are colored by number of genes expressed per cell a surrogate for sequence quality and other technical artifacts. CoGAPS patterns (C and D) are colored by marker gene expression to illustrate CoGAPS ability to find and refine patterns of specific lineages (C) and shared dynamic processes(D). 3) Proposed collaborative network with individual contributions highlighted in blue.