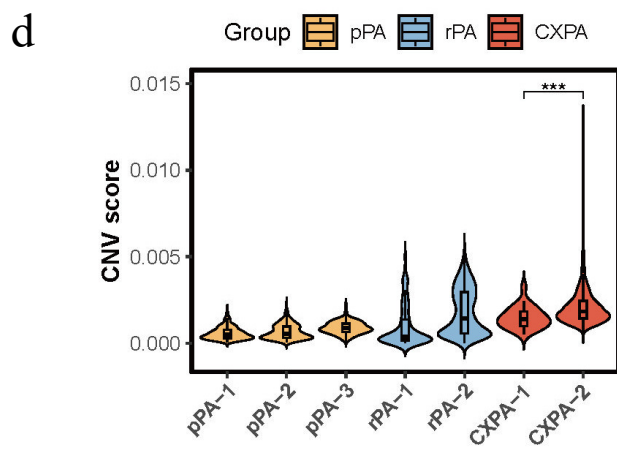
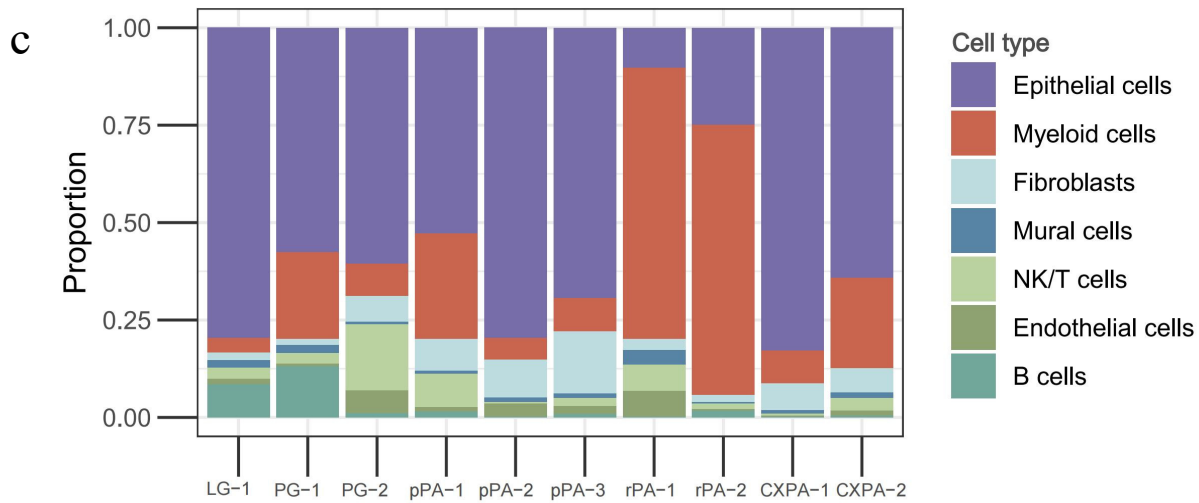
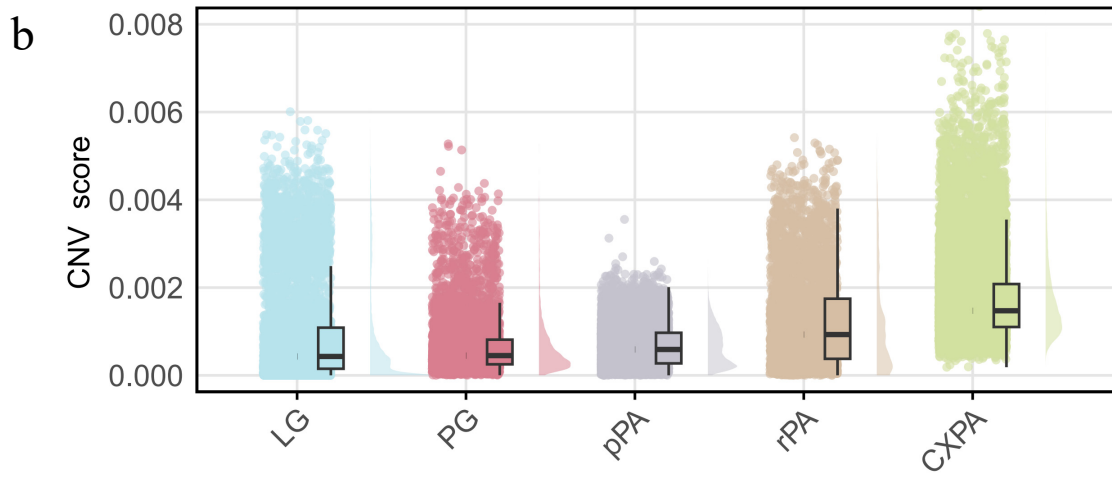
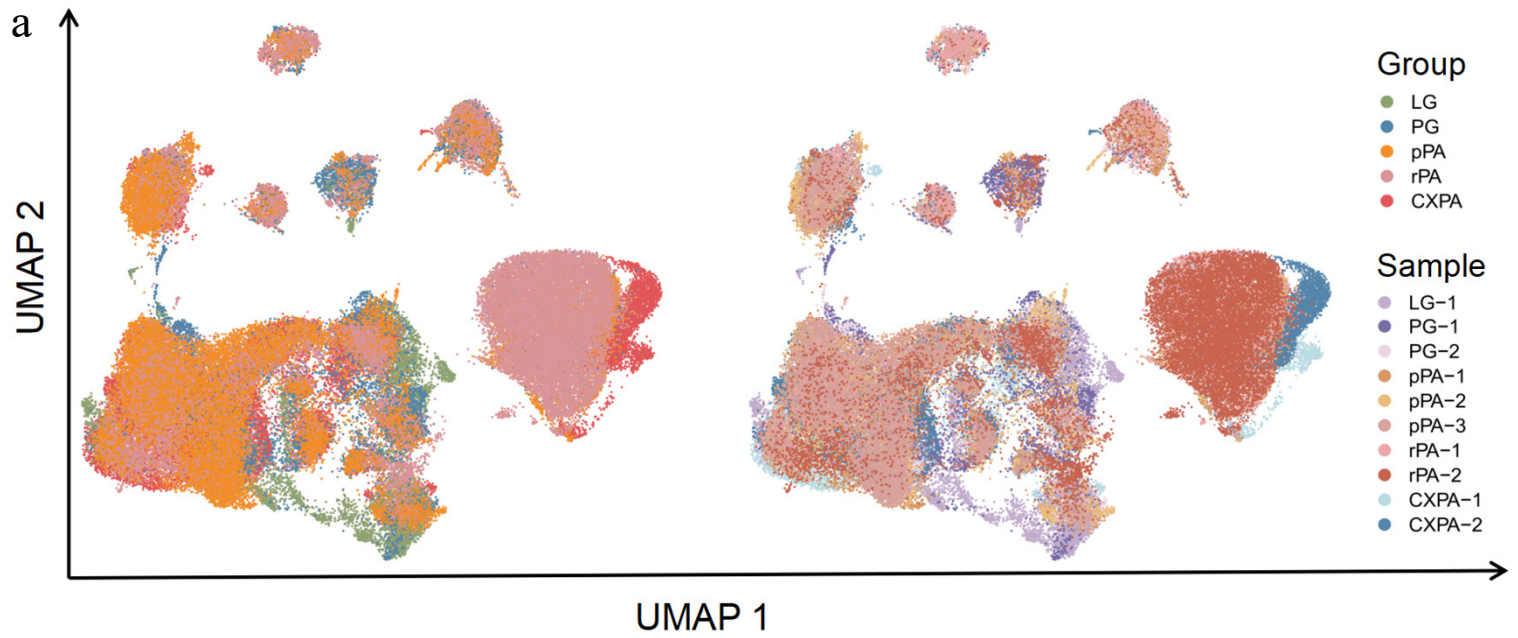
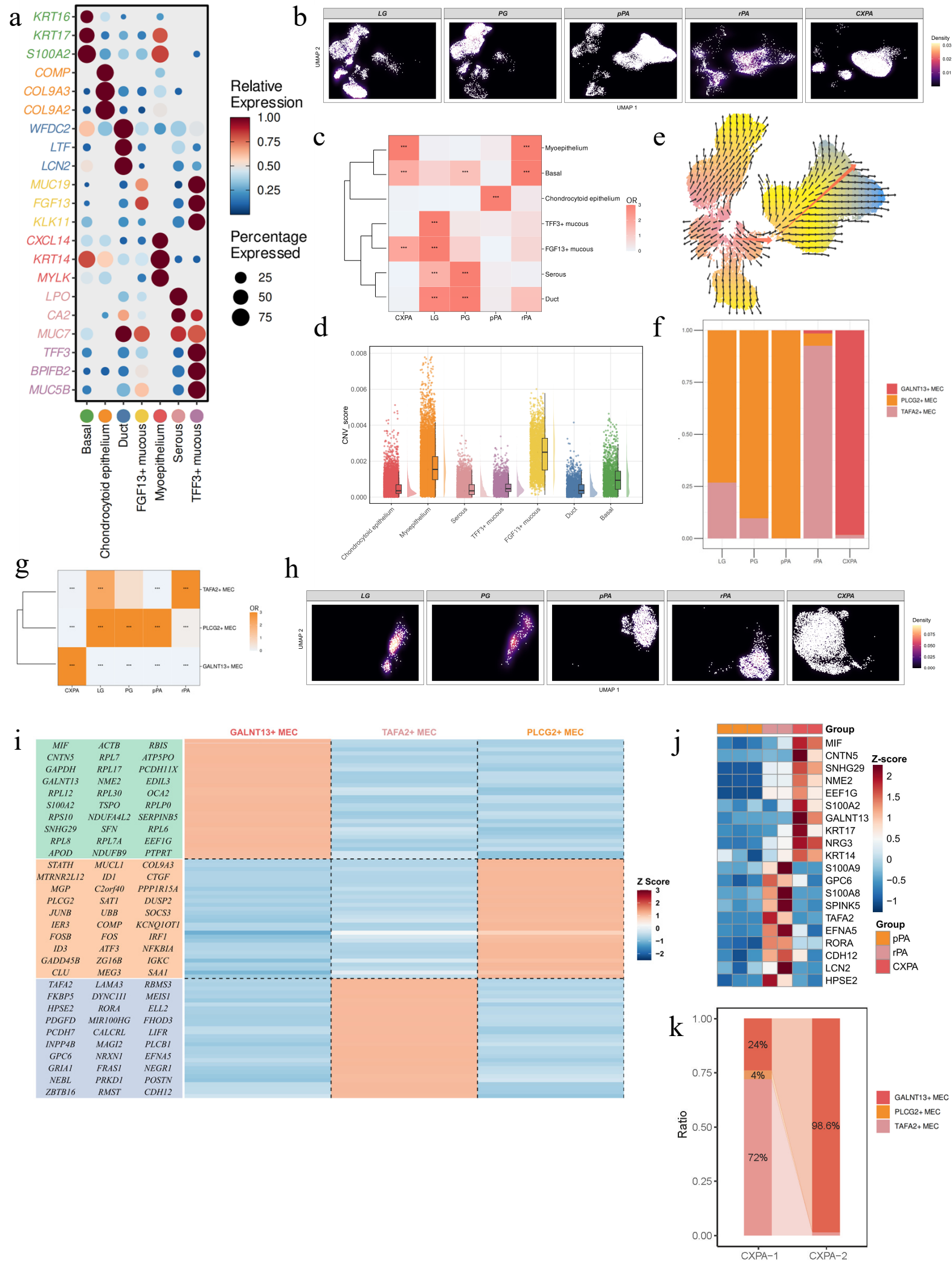




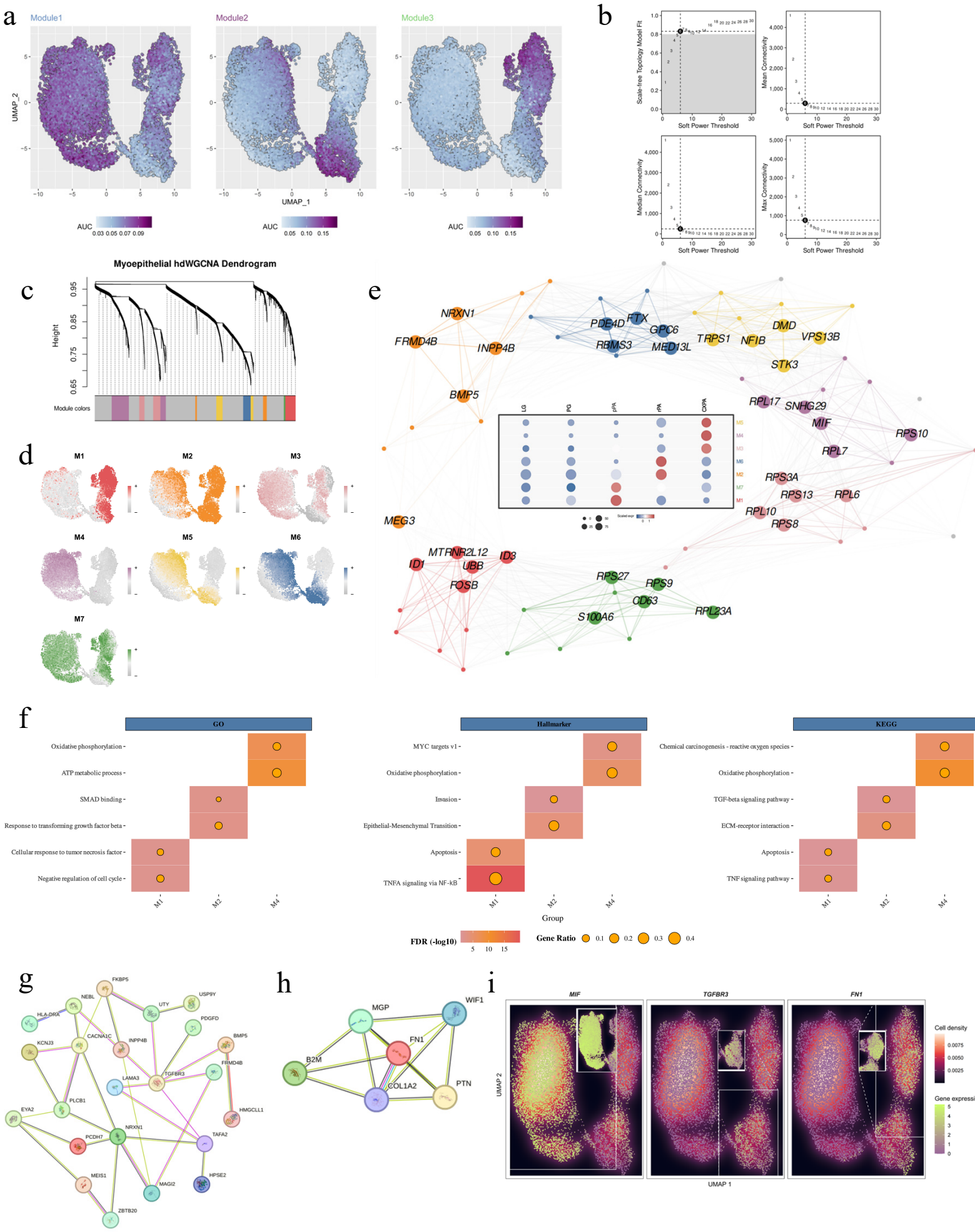
10 Fresh tissue was collected from patients and processed for 10X Genomics 3' scRNA-seq. (a) UMAP plot showing the distribution of different cell populations based on group and sample. (b) Rain cloud plot showing the CNV scores across different groups. (c) The proportion of different cell populations identified in different sample. (d) Violin plot showing CNV scores of myoepithelial cells in different tumor tissue samples.

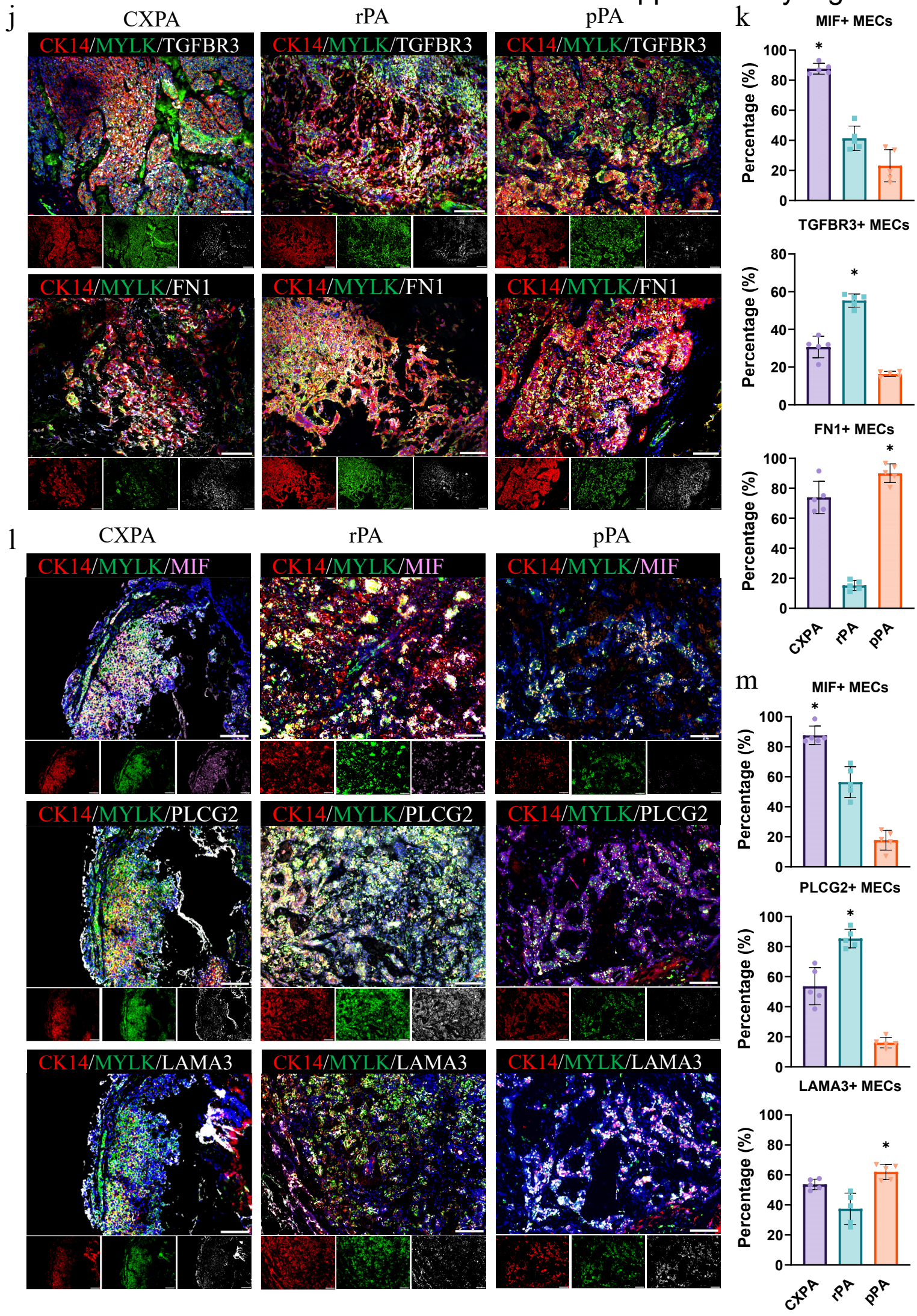
Supplementary Figure 2



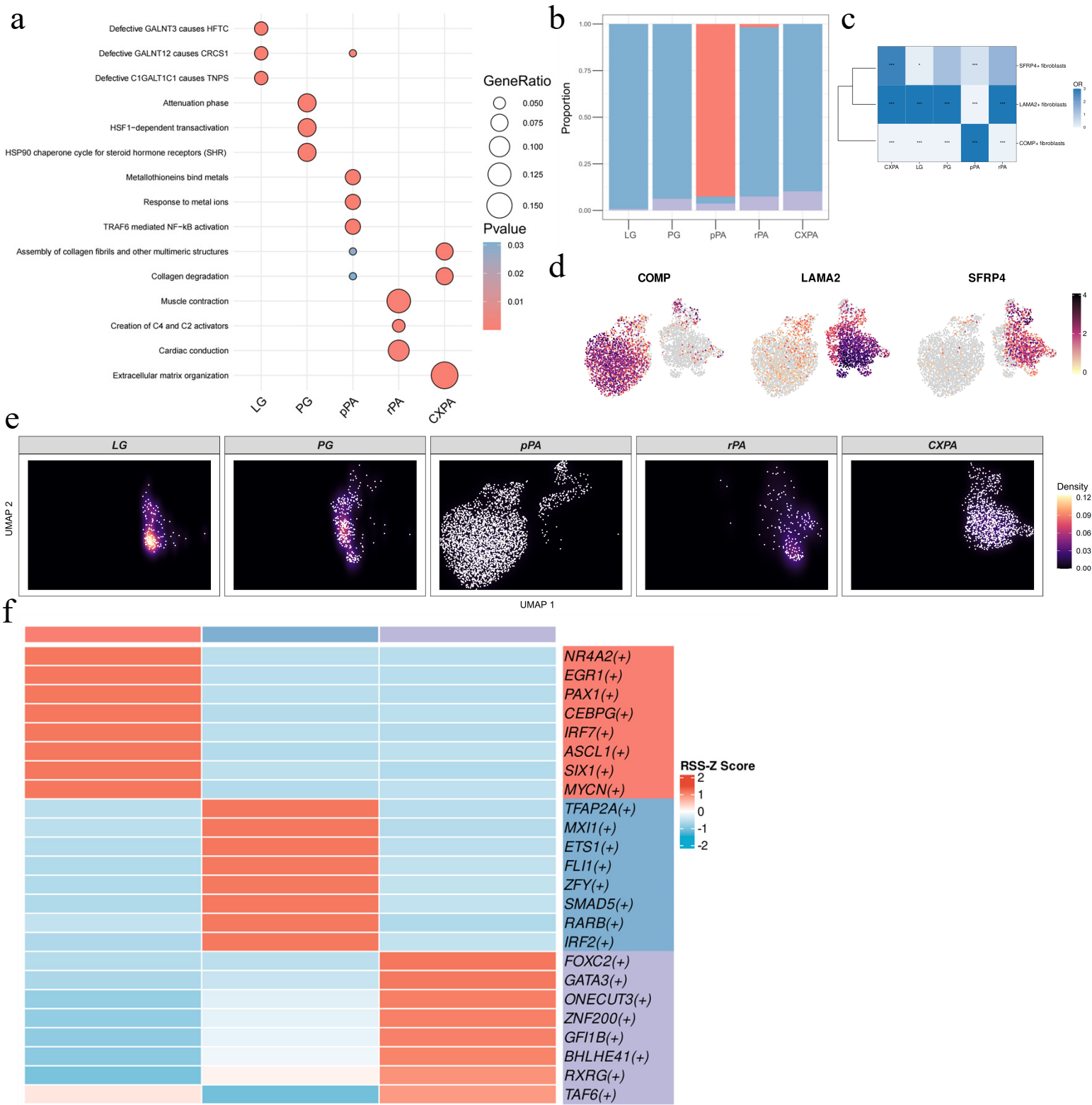
Subclustering of epithelial cells. (a) Dotplot of markers used for epithelial cell type identification. (b) The UMAP distribution of epithelial cells labelled by tissue types. (c) Heatmap showing the odds ratios (OR) for different epithelial subsets across tissue types. (d) Rain cloud plot showing the CNV scores across different epithelial subsets. (e) Trajectory inference showing the epithelial cell subsets development by cell state designation. (f) Stacked bar plot showing the distribution of myoepithelial subset proportions across different groups. (g) Heatmap showing the odds ratios (OR) for different myoepithelial subsets across tissue types. (h) The UMAP distribution of myoepithelial cells labelled by tissue types. (i) A heatmap of marker genes for each myoepithelial cell type. (j) The heatmap displays the most significantly differentially expressed genes in epithelial cells across different samples. (k) Stacked bar plot showing the distribution of myoepithelial subset proportions of CXPA samples.

Supplementary Figure 3



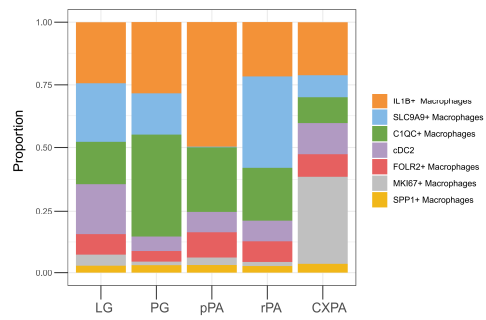


(a) AUC values showing the distribution for each estimated transcription factors module. (b) Evaluation of the soft power threshold in hdWGCNA. (c) Dendrogram of myoepithelial hdWGCNA, showing the hierarchical clustering of genes. (d) UMAP colored by module eigengene for selected modules. (e) Gene co-expression network of myoepithelial cells, where genes are clustered into modules based on their co-expression patterns. Nodes represent genes, and edges indicate co-expression relationships. The inset shows the association of gene modules with different conditions, with color intensity reflecting the module's relationship to each condition. (f) Enrichment analysis of gene sets across different groups using GO, Hallmarker, and KEGG databases. (g) Protein-protein interaction network of intersecting genes in the recurrence group. (h) Protein-protein interaction network of intersecting genes in the primary group. (i) The UMAP gene expression of MIF, TGFBR3 and FN1 in myoepithelial cells. (j) Representative images of mIHC showing the expression of CK14, MYLK, TGFBR3, and FN1 in PA sections; scale bar = 100 μ m. (k) Percentage of MIF+, TGFBR3+, and FN1+ MECs in pPA, rPA and CXPA. (l) Representative images of mIHC showing the expression of CK14, MYLK, MIF, PLCG2, and LAMA3 in PA sections; scale bar = 100 μ m. (m) Percentage of MIF+, PLCG2+, and LAMA3+ MECs in pPA, rPA and CXPA.

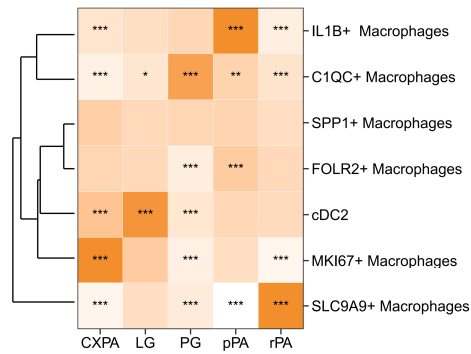


(a) Enrichment analysis of marker genes for fibroblast subsets. (b) The proportions of identified fibroblast populations in each tissue type. (c) Heatmap showing the ORs of fibroblast clusters occurring in each tissue types. (d) UMAP plot showing the expression of selected marker genes in the 3 fibroblast populations. (e) UMAP plots showing the density of cells across different tissue types. (f) Heatmap of differentially expressed transcription factors of fibroblast subsets.

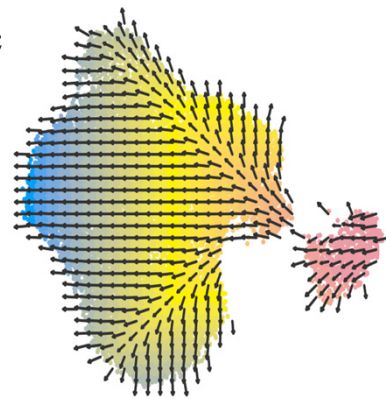
a



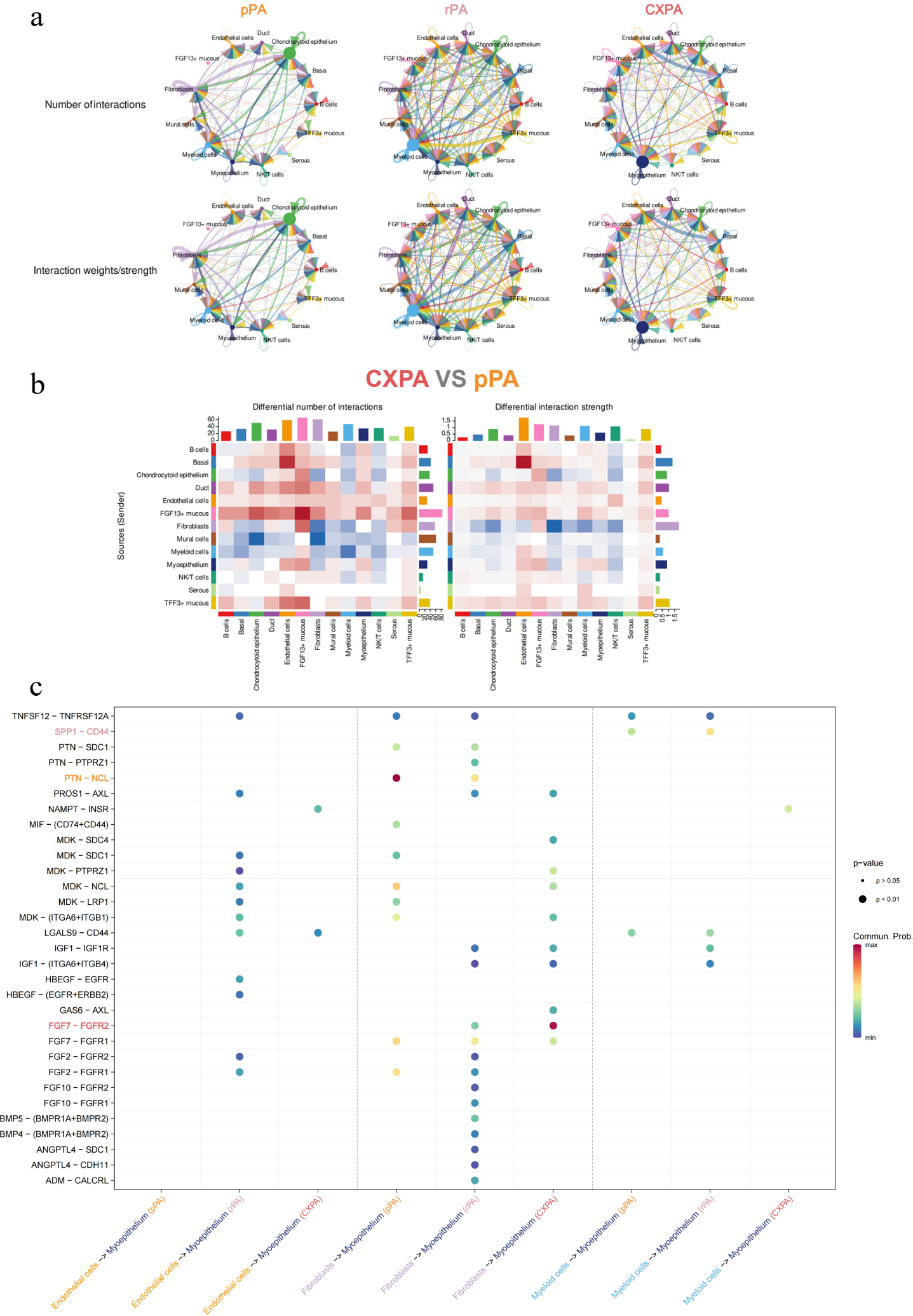
b



c



(a) The proportions of identified myeloid cell populations in each tissue type. (b) Heatmap showing the ORs of myeloid cell clusters occurring in each tissue types. (c) Trajectory inference showing the myeloid cell subsets development by cell state designation.



(a) Number of predicted interactions ($p \leq 0.05$) between tumor cell subsets based on CellChat in pPA, rPA, and CXPA. (b) Interaction networks showing the number of interactions (top) and interaction weights/strength (bottom) between different cell types in pPA, rPA, and CXPA. (c) Communication analysis showing the interactions between different cell types with myoepithelial cells as the output.