

Figure S1. The positive expression of EpCAM in GSRC. EpCAM is mainly expressed in the cell membranous and partly in the cytoplasmic (n=30). Scale bar: 50μm.

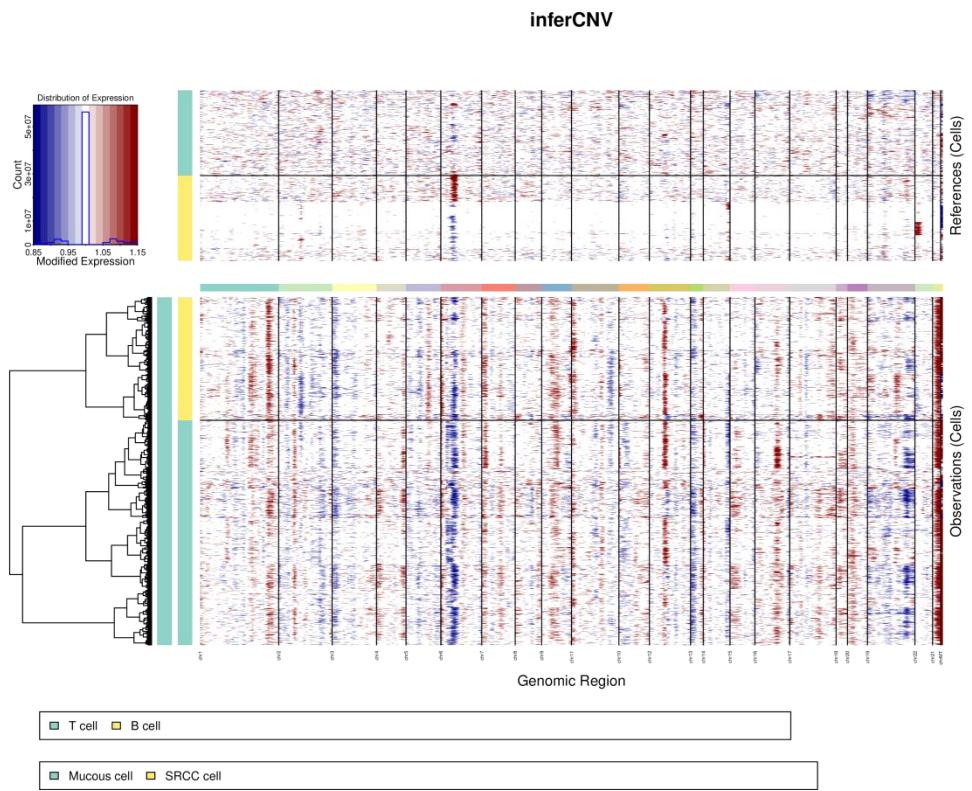
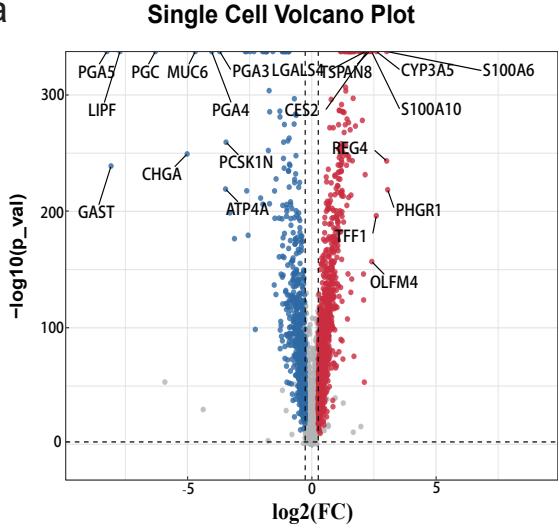


Figure S2 Heatmap showing large-scale copy number variation for each subclusters of mucous cells and SRCC cells. Red: amplifications; blue: deletions. Source data are provided as a Source Data file.

a



b

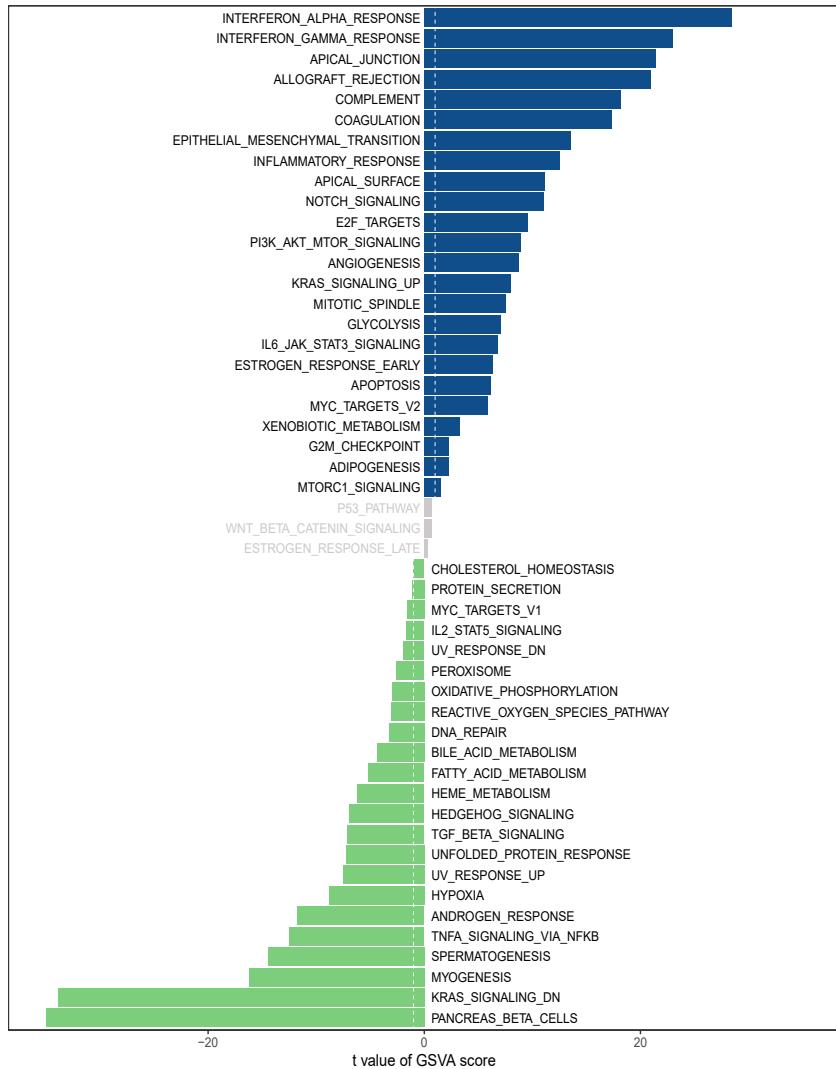


Figure S3 Differentially expressed genes and GSVA between nonmalignant epithelial cells and M/PDA cells. **a** Volcano plot of differentially expressed genes between nonmalignant epithelial cells and M/PDA cells. **b** GSVA of nonmalignant epithelial cells and adenocarcinoma cells. M/DPA cells were mainly enriched in cancer-related signalling pathways, such as epithelial-mesenchymal transition, E2F targets, PI3K/AKT/mTOR signalling, KRAS signalling, etc. GSVA data were plotted according to the t value of limma, and at value > 5 was considered significant. The statistical strategy were two-sided Student's t-test. Source data are provided as a Source Data file.

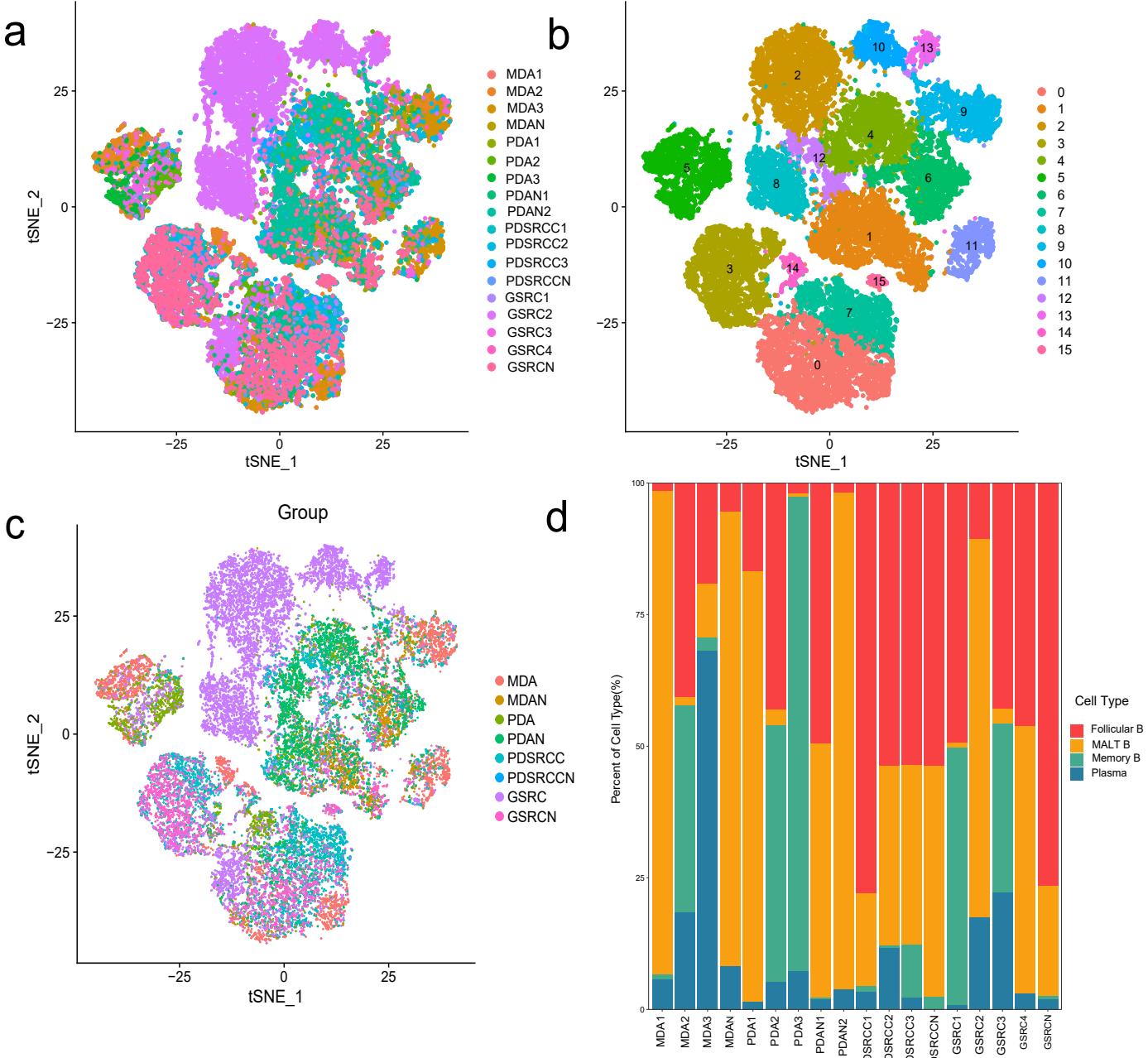


Figure S4 Classification plots of B cells. **a** tSNE plots showing subclusters of B cells (coloured by sample origin). **b** tSNE plots showing subclusters of B cells. **c** tSNE plots showing subclusters of B cells (coloured by group). **d** Scale plot of subclusters of B cells. Source data are provided as a Source Data file.

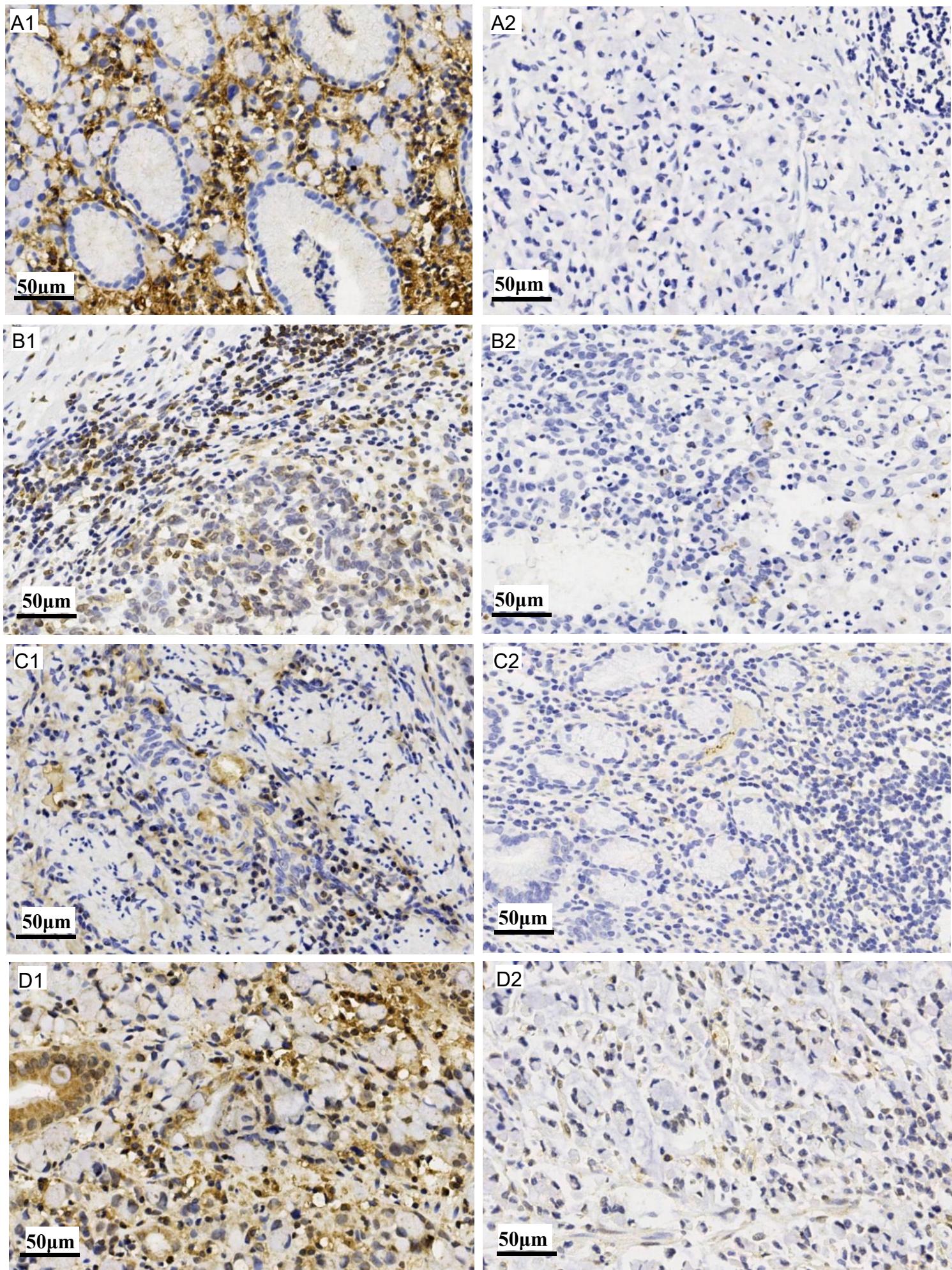


Figure S5 The positive and negative expression of JCHAIN(A1/A2), CD74(B1/B2), FOXP3(C1/C2) and KLDR1(D1/D2) in GA (n=30). Scale bar: 50µm.

Supplementary Table 1 Table of basic information characteristics of each sample