Supplementary information

Single-cell transcriptomics reveals tumor landscape in ovarian carcinosarcoma

Junfen XU[1,2], Mengyan TU[1]

[1]*Department of Gynecologic Oncology, Women's Hospital, Zhejiang University School of Medicine, Hangzhou 310006, China*
[2]*Zhejiang Provincial Clinical Research Center for Obstetrics and Gynecology, Hangzhou 310006, China*

Materials and methods

Clinical specimen

Human ovarian carcinosarcoma (OCS) tissue for single-cell RNA sequencing (scRNA-seq) and immunohistochemistry (IHC) validation was obtained from a 61-year-old patient with OCS who had undergone resection in our hospital. She experienced a relapse four months after the last chemotherapy session and passed away 12 months after surgery. This study was approved by the Ethics Committee of the Women's Hospital of Zhejiang University, Hangzhou, China. Informed consent was acquired from the enrolled patient.

Isolation of single cells

After surgical resection, fresh sample was collected and washed with phosphate-buffered saline (PBS) for three times. Tissue was cut into approximately 1 mm^3 pieces and enzymatically digested by collagenase IV (Sigma), collagenase I (Sigma), and DNase I (Worthington) at 37 °C for 30 min. After digestion, the sample was filtered out with a 35-μm cell strainer. Single cells were stained with acridine orange/propidium iodide (AO/PI) for viability assessment and live cells were preferentially sorted for single-cell sequencing.

Single-cell RNA sequencing and statistical analysis

ScRNA-seq was performed as previously reported (Shen et al., 2022; Xu et al., 2022). Briefly, the cellular suspension was loaded onto a 10X Genomics Chromium Controller instrument to generate single-cell Gel Beads-in-emulsions. Then, the scRNA-seq library was constructed using Chromium Single-Cell 3' V3 Reagent Kits (10X Genomics). A 150-bp paired-end sequencing run was conducted on the Illumina platform (Illumina, San Diego, CA) and the data analysis was performed by NovelBio Bio-Pharm Technology Co., Ltd. FASTP was first used with the default parameter to filter the adaptor sequence and remove low-quality reads to achieve the clean data (Chen et al., 2018). Then, Cell Ranger (version 3.1.0, https://github.com/10XGenomics/cellranger) was employed to obtain feature-barcode matrices by aligning to the human genome (GRCh38 Ensemble: version 91). Cells with fewer than 200 expressed genes and those with mitochondria unique molecular identifier (UMI) rate over 30% were removed. Subsequently, Seurat (version: 3.1.4, https://satijalab.org/seurat) was implemented for the procession quality control (QC) to obtain the scaled data. Principle component analysis (PCA) was constructed based on the scaled data with the top 2000 most highly variable genes. The top 10 principals were used for t-distributed stochastic neighbor embedding (t-SNE) construction. We acquired the unsupervised cell cluster results by utilizing the graph-based cluster method (resolution=0.8). We calculated the marker genes by the Seurat FindAllMarkers function with Wilcox rank sum test algorithm using the following criteria: (1) $\log_2(FC)>0.25$; (2) P value<0.05; (3) min.PCT>0.1.

Copy number variation (CNV) estimation

In order to distinguish malignant cells from non-malignant cells in the human OCS sample, we identified somatic CNVs with the R package infercnv (v0.8.2, https://github.com/broadinstitute/inferCNV). Cells with CNV signal above 0.05 and CNV correlation above 0.5 were identified as putative malignant cells.

Single-cell regulatory network inference and clustering (SCENIC) analysis

pySCENIC (v0.9.5) was used to assess the transcription factor regulation strength (Aibar et al., 2017).

Immunohistochemistry (IHC)

The specimen was acquired from a patient with OCS. Formalin-fixed paraffin embedded (FFPE) human tissue sections of 4 μm thickness were processed as previously described (Xu and Lu, 2020). Briefly, slides were deparaffinized in xylene and rehydrated in graded dilutions of aqueous ethanol (100% EtOH; 85% EtOH; 85% EtOH). They were then placed in an antigen target retrieval solution (Dako) and pressure cooked for 15 min for antigen retrieval. The slides were incubated for 1 h in a protein blocking buffer solution and then incubated overnight at 4 ℃ separately with cytokeratin (CK) (#ZM-0308, ZSGB-BIO, 1:100 dilution), E-cadherin (#ZM-0092, ZSGB-BIO, 1:500 dilution), vimentin (#ZM-0260, ZSGB-BIO, 1:200 dilution), Desmin (#ZA-0610, ZSGB-BIO, 1:150 dilution), Ki-67 (#ZM-0116, ZSGB-BIO, 1:200 dilution), p16 (#ZM-0205, ZSGB-BIO, 1:100 dilution), p53 (#ZM0408, ZSGB-BIO, 1:50 dilution), PD1 (#ZM-0381, ZSGB-BIO), EPCAM (#21050-1-AP, Proteintech, 1:200 dilution), KRT7 (#17513-1-AP, Proteintech, 1:2000 dilution), DCN (#14667-1-AP, Proteintech, 1:400 dilution), and COL3A (#22734-1-AP, Proteintech, 1:1000 dilution). The subsequent morning, the slides were washed and incubated with anti-rabbit and anti-mouse labeled polymer-HRP solution (Dako) for 1 h, and then incubated with freshly prepared 3,3-diaminobenzidine chromogen solution and counterstained with hematoxylin. The stained slides were imaged using a digital scanning microscopic imaging system (OCUS) using a 20× objective.

ScRNA-seq data comparison of OCS tumor with open source HGSOC tumors

ScRNA expression data of HGSOC was downloaded from GSE184880. Based on the dimension reduction result and gene expression, cells were identified as epithelial cells (markers: EPCAM, KRT18, and KRT8) and fibroblast (Marker: DCN), respectively. These cells were then integrated with our OCS result separately by Harmony with default parameter to correct the batch effect. Graph-Cluster (resolution=0.8) was applied based on the Harmony dimension reduction result.

References

Aibar S, González-Blas CB, Moerman T, et al., 2017. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods*, 14(11):1083-1086. https://doi.org/10.1038/nmeth.4463

Chen SF, Zhou YQ, Chen YR, et al., 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 34(17):i884-i890. https://doi.org/10.1093/bioinformatics/bty560

Shen YM, Ren Y, Chen KL, et al., 2022. The impact of neoadjuvant chemotherapy on the tumor microenvironment in advanced high-grade serous carcinoma. *Oncogenesis*, 11:43. https://doi.org/10.1038/s41389-022-00419-1

Xu JF, Lu WG, 2020. FAM83A exerts tumor-suppressive roles in cervical cancer by regulating integrins. *Int J Oncol*, 57(2):509-521. https://doi.org/10.3892/ijo.2020.5078

Xu JF, Fang YF, Chen KL, et al., 2022. Single-cell RNA sequencing reveals the tissue architecture in human high-grade serous ovarian cancer. *Clin Cancer Res*, 28(16):3590-3602.

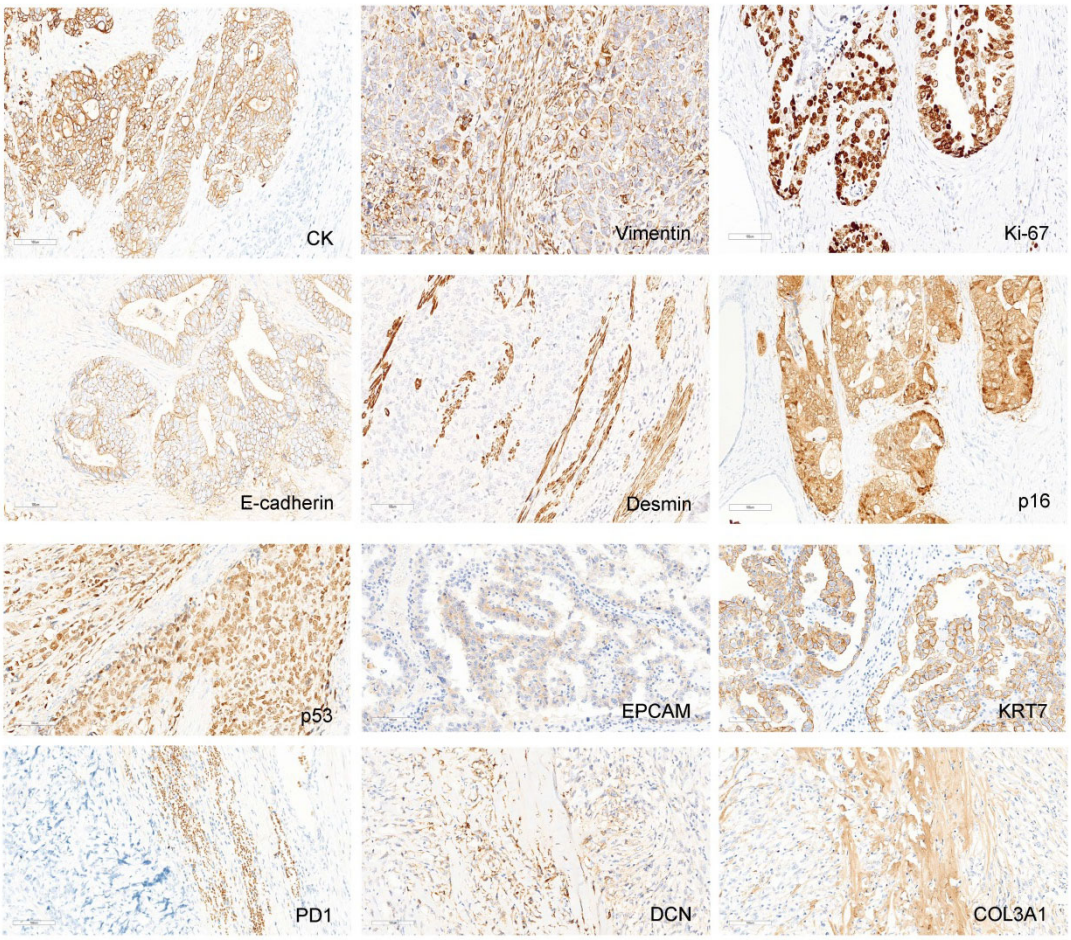

Fig. S1 Immunohistochemistry (IHC) staining in ovarian carcinosarcoma (OCS) tumor. IHC staining for cytokeratin (CK), vimentin, Ki-67, E-cadherin, desmin, p16, p53, epithelial cell adhesion molecule (EPCAM), KRT7, programmed death-1 (PD1), decorin (DCN), and collagen type III α1 chain (COL3A1) in the OCS tumor. Scale bar=100 μm.

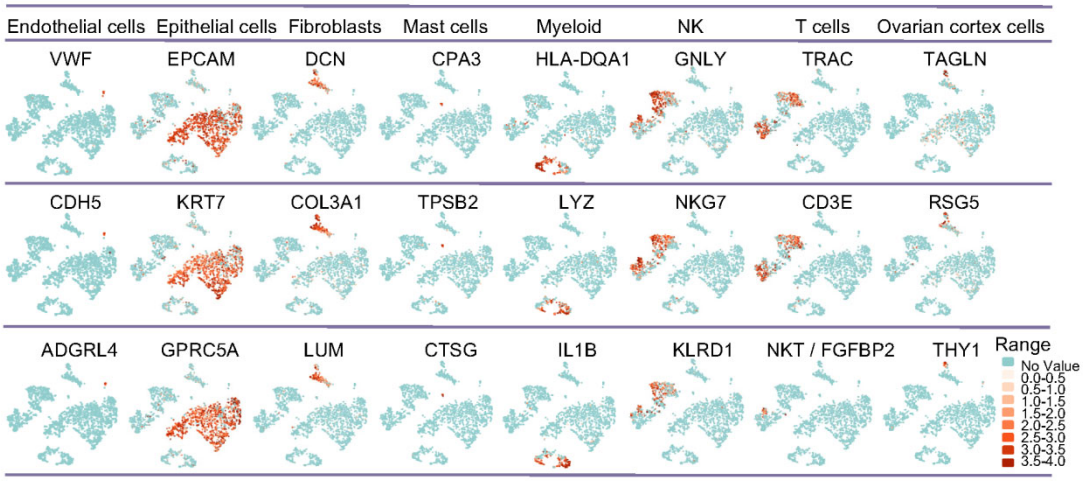

Fig. S2 Expression levels of representative well-known marker genes across the cell types in ovarian carcinosarcoma (OCS). Color key from green to orange indicates relative expression levels from low to high.

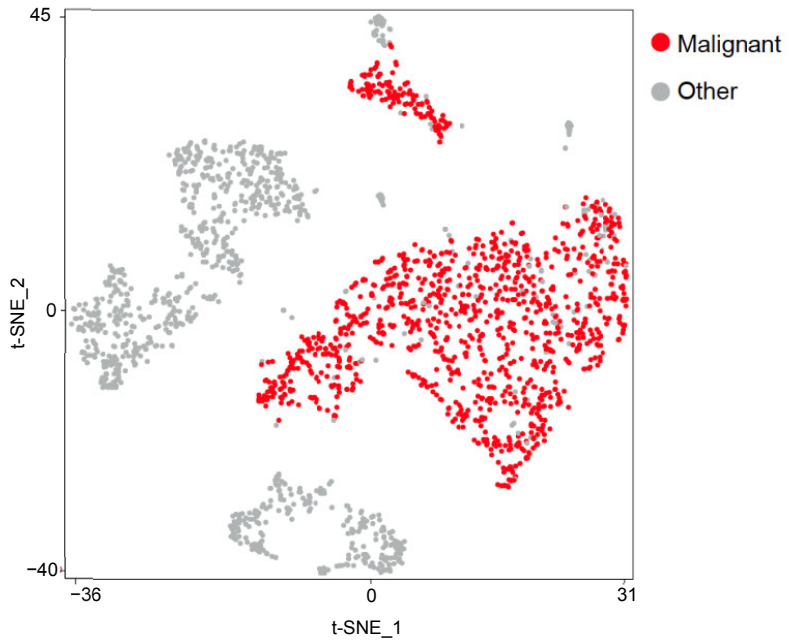

Fig. S3 Copy number variation (CNV) results in t-distributed stochastic neighbor embedding (t-SNE) plots of ovarian carcinosarcoma (OCS) tumor. Malignant cells are colored with red.

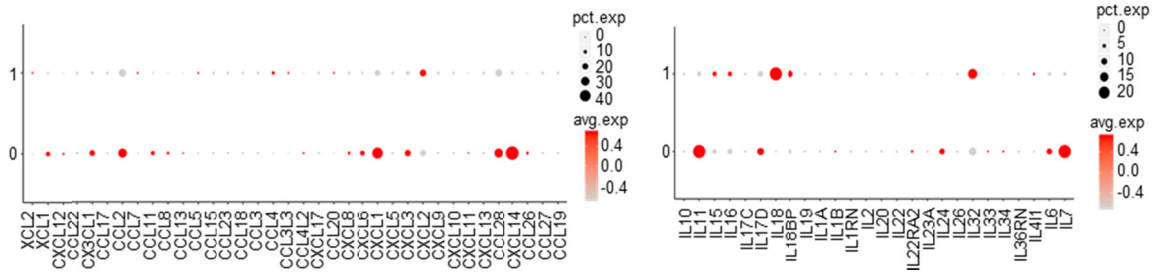

Fig. S4 Characterization of fibroblast subtypes in ovarian carcinosarcoma (OCS) tumor. Dot plots showing expression levels of specific chemokine (left) and interleukin (right) genes in each cell subcluster.