Multi-dimensional characterization of cellular states reveals clinically relevant immunological subtypes and therapeutic vulnerabilities in ovarian cancer

Can Zhang[1†], Si Li[2†], Jiyu Guo[2†], Tao Pan[1], Ya Zhang[1], Yueying Gao[2], Jiwei Pan[1], Meng Liu[1], Qingyi Yang[2], Jinyang Yu[1], Juan Xu[3*], Yongsheng Li[2,4,5*] and Xia Li[1,3*]

Abstract

Background Diverse cell types and cellular states in the tumor microenvironment (TME) are drivers of biological and therapeutic heterogeneity in ovarian cancer (OV). Characterization of the diverse malignant and immunology cellular states that make up the TME and their associations with clinical outcomes are critical for cancer therapy. However, we are still lack of knowledge about the cellular states and their clinical relevance in OV.

Methods We manually collected the comprehensive transcriptomes of OV samples and characterized the cellular states and ecotypes based on a machine-learning framework. The robustness of the cellular states was validated in independent cohorts and single-cell transcriptomes. The functions and regulators of cellular states were investigated. Meanwhile, we thoroughly examined the associations between cellular states and various clinical factors, including clinical prognosis and drug responses.

Results We depicted and characterized an immunophenotypic landscape of 3,099 OV samples and 80,044 cells based on a machine learning framework. We identified and validated 32 distinct transcriptionally defined cellular states from 12 cell types and three cellular communities or ecotypes, extending the current immunological subtypes in OV. Functional enrichment and upstream transcriptional regulator analyses revealed cancer hallmark-related pathways and potential immunological biomarkers. We further investigated the spatial patterns of identified cellular states by integrating the spatially resolved transcriptomes. Moreover, prognostic landscape and drug sensitivity analysis exhibited clinically relevant immunological subtypes and therapeutic vulnerabilities.

†Can Zhang, Si Li and Jiyu Guo contributed equally to this work.

*Correspondence:
Juan Xu
xujuanbiocc@ems.hrbmu.edu.cn
Yongsheng Li
liyongsheng@ems.hrbmu.edu.cn
Xia Li
lixia@hrbmu.edu.cn

Full list of author information is available at the end of the article

Conclusion Our comprehensive analysis of TME helps leveraging various immunological subtypes to highlight new directions and targets for the treatment of cancer.

Keywords Tumor microenvironment, Cellular States, Clinical prognosis, Drug response, Immunotherapy

Background

Ovarian cancer (OV) is the predominant contributor to mortality among gynecologic malignancies, accounting for 5% of cancer deaths in women [1], with an overall survival rate of 45.6% [2]. The majority of patients diagnosed with OV are in the advanced stages of the cancer, and there is currently a lack of effective means of early detection. Standard treatments for patients with OV include cytoreductive surgery and platinum-based chemotherapy [2]. Despite high response rates to chemotherapy and surgery, the majority of patients has a high rate of recurrence and ultimately succumbs to the complications of disease progression [3, 4].

OV is a heterogeneous disease characterized by tumors with varying histological subtypes and microenvironmental characteristics [5]. Thus, it is crucial to identify subtypes that are closely associated with the clinic. Previous studies have revealed immune subtypes associated with OV patient's prognosis [6]. Nevertheless, this classification fails to fully take into account cell types and cellular states in the TME.

Moreover, the development of single-cell mRNA sequencing (scRNA-seq) has been widely used to investigate the TME, characterizing cellular heterogeneity at single-cell resolution [7–9]. For instance, a previous scRNA-seq analysis has discovered that cancer-associated fibroblasts (CAF) in the TME are found to be correlated with lower tumor immune infiltration as well as shorter survival among patients [10]. Moreover, as it has been reported that macrophage subtypes in OV have different clinical and immunological behaviors [11]. Furthermore, there is a correlation between activated CD4$^+$ T cells and tumor infiltrating lymphocytes with favorable survival outcomes in patients with OV [12]. Although scRNA-seq technology can decode the TME, sequencing is expensive and it is still difficult to implement large-scale cohort studies based on certain phenotypes. Moreover, there is currently a lack of large-scale bulk transcriptomes analysis of OV TME. Recent studies have introduced EcoTyper, which deconvolves cell types and cellular states from large amounts of gene expression data to better predict cancer progression, outcome and response in a vast cohort of patients [13].

Here, we performed a pan-cohort characterization of cellular states in OV based on bulk, single-cell and spatially resolved gene expression data. We identified 32 transcriptionally defined cellular states in OV TME, extending the previous immunological classifications. We also validated these findings with multi-dimensional analysis. Based on the marker genes of cellular states, we thoroughly characterized the functional landscape of cellular states in OV. In addition, we discovered critical upstream regulators, suggesting that they may serve as potential biomarkers for OV. We also identified most immunological cellular states associated with clinical outcomes, providing a prognostic atlas of the cellular states in OV. In particular, we constructed spatial tumor microenvironments for cellular states and compared the differences in scores of marker genes for cellular states in the malignant (Mal) regions, tumor boundary (Bdy) regions and non-malignant (nMal) regions. Finally, we discovered cellular states associated with drug therapy responses in OV patients. To summarize, our large-scale study resolved the OV microenvironment and unveiled new avenues and targets for the treatment of cancer.

Methods

Manually collection of OV transcriptomes

Gene expression profiles and clinical information of OV patients were obtained from Gene Expression Omnibus (GEO) [14], The Cancer Genome Atlas (TCGA) [15] and one recent study [16]. A total of 40 public OV datasets were analyzed in this study, including 38 public bulk datasets, one scRNA-seq dataset and one spatial transcriptome. For datasets with the original CEL files, we processed using the RMA algorithm from the 'affay' R package [17]. For other platforms, available transcriptome data were downloaded from the original studies. In addition, we downloaded the annotation files of each platform for probe annotation. If a probe corresponds to multiple genes, the probe was deleted. If a gene corresponds to multiple probes, the probe with the highest expression value was selected as the expression value of the gene [18].

Six datasets (GSE26193, GSE9891, GSE26712, GSE32062, GSE49997 and GSE140082) with survival information were used as the discovery cohorts. It included 1,421 tumor samples and 5 different microarray platforms. The selection of discovery cohorts included the following two criteria: the patients were with overall survival information; and the datasets with a larger number of samples. The remaining 33 datasets were used as the validation cohorts. The datasets with incomplete survival or drug responses information were integrated by platform, naming them as integrate_GPL570 and integrate_GPL96. Detail information about the 40 OV datasets was listed in Additional file 1: Table S1.

To obtain the gene expression profile of the discovery cohort, the Entrez gene IDs common to each platform of the discovery cohort were screened for analysis and the Median Rank Score Normalization method (MNORM) was used for cross-platform standardization [19, 20]. Next, batch correction was performed using the 'combat' function of the 'sva' R package [21].

Single-cell sequencing of OV

We gathered OV scRNA-seq transcriptome data from the literature [22] and obtained the corresponding annotation files. Notably, we screened scRNA-seq data from primary and metastatic tumors for subsequent analysis. Quality control and downstream analysis were performed using the 'Seurat' R package [23]. Cells with less than 500 and more than 6,000 expressed genes were removed, cells with more than 10% mitochondrial gene expression were also removed, and only cells with a total count of more than 500 were retained.

The unique molecular identifier (UMI) counts were normalized based on 'NormalizedData'. The function 'CellCycleScoring' was used to calculate cell cycle scores. Next, the data were scaled based on 'ScaleData', with the parameters 'vars.to.regress' = 'S.Score' and 'G2M.Score' to remove the effect of the cell cycle. The function of 'FindVariableFeatures' was used to identify highly variable genes. Downstream analysis was performed based on the top 2,000 highly variable genes.

The 'RunPCA' function was applied to perform dimensional reduction and the top 30 principal components (PCs) were selected for downstream analysis. Next, cells were divided into diverse clusters utilizing the 'FindNeighbors' function and 'FindCluster' function. Finally, Uniform Manifold Approximation and Projection (UMAP) was used for visualization. All relevant ethical regulations were followed the original study of the datasets and the authors of the source studies had also obtained informed consent from participants.

Cell-cell interaction analysis

CellChat was used to perform cell-cell communication analysis for the cellular states based on the known ligand-receptor interactions [24]. CellChat takes the normalized gene expression data as input to model the probability of cell–cell communication by combining gene expression with the existing database. For the main analyses, the main functions 'computeCommunProb', 'computeCommunProbPathway' and 'aggregateNet' were used with default parameter. To reveal the interactions between cellular states, the CXCL signaling pathway, PARs signaling pathway, BAG signaling pathway and VEGF signaling pathway were selected for further visualization.

Cell-type specific gene profiles

To determine the cell type-specific gene expression profiles of OV patients, we employed CIBERSORTx [25] to infer gene expression profiles and estimate the abundance of each cell types in mixed cells. To investigate the major cellular components in OV, we utilized two signature matrices validated in solid tumors, TR4 [25] and LM22 [26] in EcoTyper [27]. TR4 is a feature matrix that encompasses of Epithelial, Endothelial, immune and Fibroblast populations. LM22 contains 22 human immune subpopulations. The 22 subpopulations of LM22 were aggregated into 9 major lineages which included B cells, CD4 T cells, CD8 T cells, NK cells, Plasma cells (PCs), Monocytes/Macrophages, Polymorphonuclear neutrophils (PMNs), Dendritic cells and Mast cells. Based on the transcriptome of discovery cohort, we obtained the abundances of 12 cell types. In each tumor sample, the aggregate abundance of the 12 cell types was normalized to 1. Finally, cell type-specific gene expression profiles were determined using the default parameters of CIBERSORTx [25].

Discovery of cellular States in OV

Upon getting the cell type-specific gene expression profiles, EcoTyper was operationalized to identify the clusters of each cell type. EcoTyper primarily applied the non-negative matrix factorization (NMF) in conjunction with specific heuristics to identify and quantify the cellular states [27]. By utilizing cell type-specific gene expression profiles, EcoTyper determined the most stable number of cellular states for each cell type by calculating the cophenetic coefficient. We determined the number of cellular states corresponding to a cophenetic coefficient closest to 0.99. To ensure the acquisition of high-quality cellular states, we selectively retained cellular states with more than 10 marker genes. Additionally, an adaptive false-positive index (AFI) was used to eliminate potentially false-positive cellular states [28].

Recovery of OV cellular states from external cohorts

We used the EcoTyper reference-based framework based on the NMF model to recover the cellular states identified from the discovery datasets in other external datasets [27]. Concisely, EcoTyper applied the learnt model from the discovery cohort to the validation cohorts of OV. Drawing from a gene expression matrix, EcoTyper created a coefficient matrix for each cell type, where each cellular state was depicted as a weight. Statistical evaluation of the recovery of each cellular state in the external cohorts was conducted through permutation test, and the z-score was calculated as the statistical confidence level. Cellular states that were significantly recovered were identified based on the criterion that z-score was greater than 1.65 and one-sided *p*-value was less than

0.05. We applied this framework to all external cohorts and scRNA-seq data derived from OV patients.

OV ecotypes detection

Besides, EcoTyper employed algorithms that can recognize co-associated networks of cellular states, referred to as 'ecotype', within different tissue specimens [27]. We employed EcoTyper to identify ecotypes of OV patients. First a binary matrix C is defined, with cellular states represent rows and samples represent columns. If cellular state i is assigned to sample j, $C_{ij} = 1$, otherwise $C_{ij} = 0$.

Next, we used the Jaccard index to quantify the degree of overlap between each pair of cellular states in the tumor samples. We calculated the pairwise combinations of all the cellular states in C, generating a 32*32 Jaccard matrix based on the number of OV cellular states. Once the Jaccard matrix has been generated, we performed a hypergeometric test for each pair of cells under the null hypothesis that there was no overlap between two cellular states. If the p-value > 0.01, the Jaccard index was set to 0 and 1 otherwise. Subsequently, unsupervised clustering of the Jaccard matrix was performed using the function 'hclust' in the R package 'stats'. Finally, to determine the optimal number of clusters, we maximized the silhouette width.

Functional analysis of cellular States

Based on EcoTyper [27], we obtained the marker genes of cellular states. Next, we utilized the 'org.Hs.eg.db' and 'clusterProfiler' [29] R packages to perform functional enrichment analysis. We calculated the ratio of 'GeneRatio' to 'BgRatio' for each pathway, which was defined as the enrichment score. The remarkably enriched pathways were screened with p.adjust < 0.05.

Meanwhile, we obtained the cancer hallmark pathways from MsigDB [30] and the immune-related pathways were derived from ImmPort [31, 32] database. To discover the potential pathways enriched by the marker genes of the cellular states, hypergeometric test was used to calculate the enrichment p-value. It is noteworthy that we used FDR to correct the p-value and considered pathways with p.adjust < 0.01 were significantly enriched. Top three genes were shown based on fold change values of marker genes for cellular states, for repetitive marker genes that appeared in multiple cellular states, the genes with the highest fold change values were shown. To further analyze the genes with high expression in each cellular state, we utilized the 'GSEA' R package [33] for performing Gene Set Enrichment Analysis. 'GseaVis' R package (https://github.com/junjunlab/GseaVis) was used to visualize the enriched pathways.

Identification of transcriptional regulators for cellular States

Available RNA-binding proteins (RBPs) and motifs were collected from the ATtRACT database (http://attract.cnic.es) [34]. Based on the previous study [35], we obtained a ranked matrix of genes corresponding to each motif of RBPs. Then, we selected the top 20 ranked genes to form the RBP-gene relationship pairs. Next, we used the Wilcoxon's rank sum test to obtain the highly expressed genes. First, we categorized patients into two groups based on the cellular state within versus without cellular state i. All genes were sorted as follow:

$$S(i) = -\log(p) * sign(\log(FC(i)))$$

Here, p was the Wilcoxon's rank sum p-value to compare the expression difference of two groups. *FC*, referring to fold change, represented the average expression in the cellular state i group divided by the average expression of patients with or without the cellular state i group. Secondly, the top ten genes were selected for analysis. Finally, according to the regulatory relationship between genes and RBPs, we visualized the RBPs-genes network based on Cytoscape [36].

Simultaneously, we selected p.adjust < 0.05 genes for Rcistarget analysis [37] to identify overexpressed transcription factor binding motifs. We annotated the motifs using the 'motifAnnotations_hgnc' annotation file, preserving only those motifs with high normalized enrichment scores. Furthermore, the top ten motifs were visualized. The gene-motif sequencing database used by RcisTarget was 'hg19-tss-center-10 kb-7species.mc9n. feather'. We acquired the transcription factor (TF)-target interactions from DoRothEA [38], and the interactions confidence score between TFs and targets were categorized into five grades from A (highest confidence) to E (lowest confidence). The confidence scores of the current study were all A.

Survival analysis of cellular States

To determine the relationship between the cellular state and prognostic performance, survival analysis of OV cellular states was performed using the 'survfit' function from the R package 'survival'. The survival curves were plotted using the 'ggsurvplot' function. In addition, we compared the survival time of patients with one cellular state to the other cellular states. Significantly, the 'coef' coefficients for each cellular state were calculated using the 'coxph' function to determine the risky and protective cellular states in OV. It was considered that p-value < 0.05 was statistically significant.

Zhang *et al. Journal of Translational Medicine*   (2025) 23:519

Page 5 of 18

Machine learning LASSO regression analysis

We performed COX univariate regression analysis of marker genes expression associated with adverse cellular states and survival data from the discovery cohort using the R package 'survival', with a threshold of $p < 0.05$. 271 genes were used for machine learning LASSO regression analysis using the R package 'glmnet'. Based on lambda. min, we selected 62 genes to construct risk model to predict the OV patient survival. The risk score for each sample was calculated by the function 'predict' and the samples were divided into high and low risk groups based on the median of riskscore. The R package 'timeROC' was used to evaluate the predictive performance of risk score. TCGA-OV dataset was used for validation.

Cellular States predict drug responses

To determine whether cellular states can predict drug responses, we investigated the drug therapy responses data in OV (GSE63885) using the 'survfit' function from the R package 'survival'. First, cellular states were categorized into sensitive and resistant according to the differences of survival time between responders and non-responders for each cellular state. Moreover, we compared the differences in sensitivity and resistance of cellular states under both platinum/cyclophosphamide and platinum/taxane.

In addition, we used data from GDSC2 (https://www.cancerrxgene.org/) [39] and CTRP (https://ocg.cancer.gov/programs/ctd2/data-portal) [40] as a training set to predict the drug sensitivity of TCGA-OV patients using the 'calcPhenotype' function from the 'oncoPredict' R package [41]. A total of 743 drug responses values were predicted for all OV patients. In light of the predicted drug responses mean values, patients were categorized into sensitive and resistant groups. The *p*-value was calculated for both groups using the 'survdiff' function, and the drugs with *p*-value < 0.05 were considered significantly associated with the cellular states.

Statistical analysis

Continuous data were compared using appropriate statistical tests, including Wilcoxon rank-sum test, or Kruskal–Wallis test. Categorical data were analysed using the Fisher's test. The log-rank test was employed to determine statistical significance in survival analysis. A *P*-value < 0.05 was considered indicative of a significant difference for all statistical tests. R software was utilized for all statistical analyses.

Results

Integrated multi-cohort gene expression profiling identifies cellular States of OV

We manually curated the OV transcriptomes and in total 1,421 OV samples from 6 public datasets were used as the discovery cohort (Additional file 1: Table S1). To comprehensively characterize the OV TME, we leveraged EcoTyper [27] to analyze the complex cellular states in the discovery cohort as well as the interaction patterns between cellular states (Fig. 1A). We first utilized CIBERSORTx to obtain 12 cell type-specific gene expression profiles. Next, we identified 32 distinct cellular states from the 12 cell types, each with 2 to 4 cellular states (Fig. 1B). In addition, the cellular states were defined by specific marker genes (Additional file 2: Fig. S1A). We next compared the proportion of patients in each cell type that were in different cellular states and found significant differences across cell types (Fig. 1C). For example, in endothelial cells and dendritic cells, patients were predominantly in cellular state S03, while in fibroblast cells and NK cells, patients were primarily in cellular state S01 (Fig. 1C). These results revealed the extensive cellular heterogeneity among OV patients.

Extensive studies have reported that tumor-infiltrating T cells are closely associated with clinical outcomes in OV [42–44]. It has been demonstrated that CD4 T cells are associated with improved survival in OV [12]. Here, we performed in-depth analysis of CD4 T cells and CD8 T cells. We obtained three CD4 T cells states (Fig. 1D) and found that immune-related genes, such as CD2, CD96 and KLRB1, were highly expressed in the CD4 T cells S02 (Additional file 2: Fig. S1B-D). CD2 is a glycoprotein that plays a crucial role in T cell activation [45]. Increasing evidences have indicated that CD2 may serve as a novel biomarker for predicting immune efficacy [46]. CD96 is a potential immunomodulatory factor for cancer [47]. Moreover, we also determined two CD8 T cells states (Fig. 1E), and found that the CD8 T cells S01 highly expressed many chemokines such as CXCL13, CCL5 and CXCR6 (Additional file 2: Fig. S1E-G). It has been shown that the high expression of CXCL13 in tumors indicates enhanced immunoreactivity in the TME, which tends to prolong patient's survival [48]. CCL5 is a chemotactic ligand that may play a fundamental role in anti-tumor immunity in OV [49]. Together, these results suggested that the immune subtypes we identified are relevant to clinical outcomes and the subtype-specific marker genes play important roles in OV.

TME are composed of complex cellular communities and recent studies have demonstrated that characterizing the ecotypes is essential for understanding the mechanisms of cancer [27, 28, 50]. We firstly identified the genes co-expressed in different cellular states based on WGCNA (Additional file 2: Fig. S1H). For example, PLAAT4, CXCL11, TAP1 and HLA-G collectively served as marker genes for dendritic cells S03 and NK cells S03. VCAN collectively served as marker genes for epithelial cells S01 and PMNs S02. These results suggested the potential association between different cellular states. To
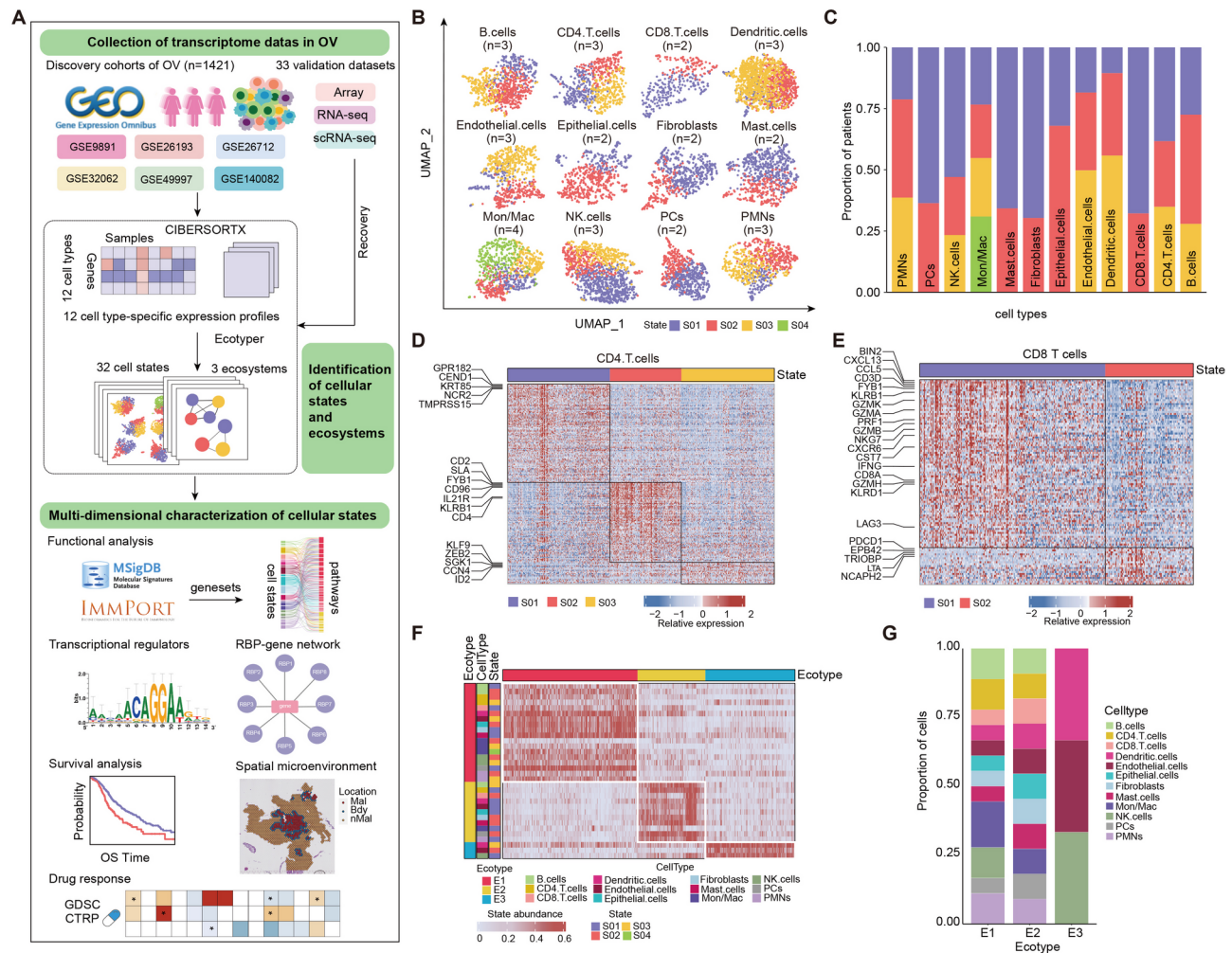
Fig. 1 Framework for pan-cohort determination of cellular states and ecosystems in OV. (**A**) Workflow for dissecting cellular states and ecosystems in OV. (**B**) UMAP plot showing cellular states of OV tumor patients identified by EcoTyper. (**C**) Cellular states distribution in each cell type of OV patients, colored by cellular states. (**D**) Heat map showing three CD4 T cells states identified from OV bulk transcriptomes in the discovery cohort. (**E**) Heat map showing two CD8 T cells states identified from OV bulk transcriptomes in the discovery cohort. (**F**) Cellular state abundance profiles in the OV patients, categorized into three ecotypes. Tumor samples (columns) are ordered based on the most abundant ecotype per specimen. (**G**) Cell types distribution of each ecotype in OV patients, colored by cell types

explore the interactions and co-occurrence patterns of different cellular states, we discovered three distinct OV ecotypes (Fig. 1F), comprising 3 to 18 co-occurring cellular states per community (Additional file 2: Fig. S1I). We found significant differences in cell composition of these ecotypes (Fig. 1G). E1 consisted of 12 cell types; E2 consisted of 11 cell types, while E3 consisted of three cell types. We next compared the distribution of cell types, cellular states, and ecotypes across different age groups and found no significant differences in immune cells or immune cellular states among the age groups. However, ecotype E1 was predominantly distributed within the age range of [63, 69] (Additional file 2: Fig. S1J-L). These findings contributed to our exploration of the potential cellular interactions of the TME in OV.

To validate the stability of the ecotypes, we recovered three ecotypes in multiple independent datasets (Additional file 2: Fig. S2A). Several studies have shown that ecotypes can predict patient clinical outcomes [27, 28]. We next analyzed the relationship between ecotypes and clinical survival in OV patients (Additional file 2: Fig. S2B-D). Higher E2 abundance was significantly associated with better outcomes (Additional file 2: Fig. S2C), whereas higher abundance of E3 was significantly associated with worse prognosis (Additional file 2: Fig. S2D). To further characterize the three OV ecotypes, we assessed the association of ecotypes abundance with therapeutic responses. E2 abundance was significantly higher in responders than nonresponders (Additional file 2: Fig. S2E), whereas E3 abundance was significantly higher in nonresponders than responders (Additional file 2: Fig.

S2F). These findings suggested that E2 and E3 may serve as candidate biomarkers for predicting response to therapeutic drugs.

Validation of cellular States in bulk tissue and single-cell OV transcriptomes

To assess the robustness of identified cellular states, we recovered them using 33 validation cohorts (Additional file 1: Table S2). Next, we calculated the proportion of cellular states assigned to each cell type in the validation cohorts and found that 100% of the cellular states were significantly recovered (Fig. 2A). For instance, when considering fibroblasts as an example, we identified two distinct fibroblast states (Fig. 2B). We successfully recovered these two specific cellular phenotypes in independent validation cohorts including GSE17260,

GSE63885, TCGA-OV and scRNA-seq dataset (Fig. 2C), suggesting the robustness of cellular states we identified in OV. In particular, we also observed that genes associated with OV development and progression were highly expressed in S01 (Fig. 2B), such as THBS2, FBN1, CTSK and COL5A1 (Fig. 2D). Elevated expression of THBS2 is correlated with unfavorable survival outcomes and can serve as a prognostic biomarker for OV [51]. Meanwhile, FBN1 promotes tumorigenesis and metastasis in OV [52] and CSTK [53] and COL5A1 were associated with OV progression and resistance to paclitaxel treatment [54].

Moreover, a growing number of studies have demonstrated that about 85-95% of OV originate from epithelial cells, and epithelial OV is the most common type of OV [55]. We obtained two cellular states in epithelial cells and recovered both cellular states in bulk transcriptomes as
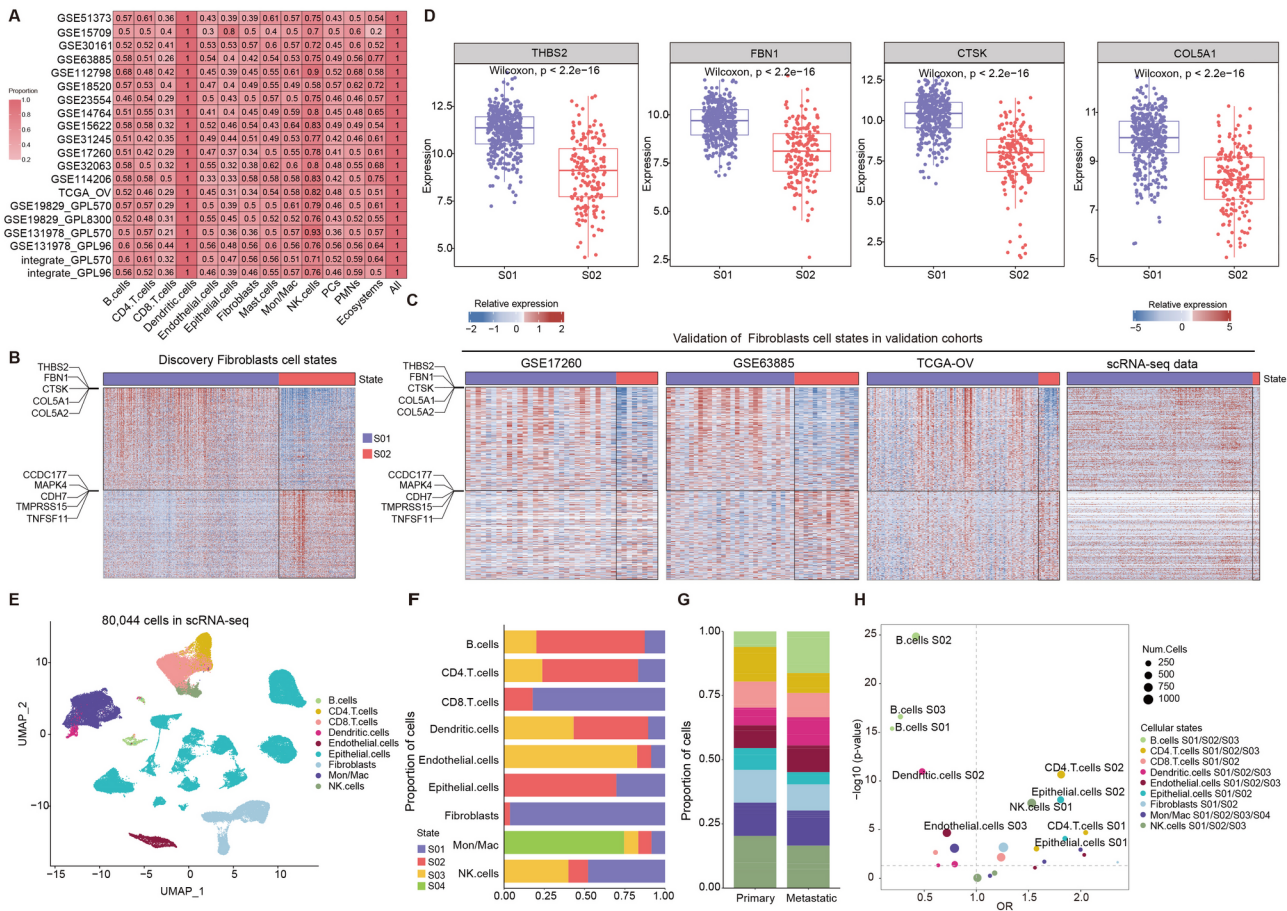
Fig. 2 Validation of cellular states in independent cohorts. (**A**) Heat map showing the percentages of patients recovered in the validation datasets. All represents union of all cellular states. (**B**) Heat map showing two fibroblasts cellular states recognized from OV bulk transcriptomes in the discovery cohort. (**C**) Heat map displaying two fibroblasts cellular states recovered by validation datasets and single-cell data. (**D**) Boxplots showing the expression levels of THBS2, FBN1, CTSK and COL5A1 of patients within two fibroblasts cellular states. *P*-values are calculated using two-sided Wilcoxon's rank sum tests. (**E**) UMAP plot showing the cell types in scRNA-seq transcriptome of OV patients. Cells were designated based on cell types annotated in a previous study. (**F**) Cellular states distribution of each cell type in scRNA-seq transcriptome of OV patients, colored by cellular states. (**G**) Bar plots depicting the distribution of cellular states in scRNA-seq transcriptome of OV patients derived from primary and metastatic tumor groups, colored by cellular states. (**H**) Scatter plot showing the differentially enriched cellular states between primary and metastatic tumor groups. The dotted line represents *p*-value=0.05 and OR=1. The OR is greater than 1 indicates that cellular states are enriched in the primary tumor, while the OR is less than 1 indicates that they are enriched in the metastatic tumor. *P*-values are calculated based on two-sided Fisher's exact tests

well as in scRNA-seq cohort (Additional file 2: Fig. S3A-B). The S02 cellular state was highly expressed genes that were correlated with the survival outcome (Additional file 2: Fig. S3A-B), and including LGALS1, RARRES2, MXRA5 and CD69 (Additional file 2: Fig. S3C). LGALS1 plays a critical role in tumor growth as well as invasion [56], and its high expression leads to poor prognosis in OV patients [57]. CD69 is a marker of lymphocyte activation [58], which may indicate the activation of the immune system.

Next, we validated each cellular state at the single-cell level. We collected single-cell data of OV from a recent study [22] with 80,044 cells from 9 cell types (Fig. 2E). We found that typical marker genes were specifically highly expressed in the corresponding cell types (Additional file 2: Fig. S3D), such as CD8A was highly expressed in CD8 T cells, DCN was highly expressed in fibroblasts, and KRT8 was highly expressed in epithelial cells. In addition, we found significant differences in the proportion of cellular states in each cell type (Fig. 2F). There was a largely consistent distribution of cellular states across cell types both at the bulk transcriptomes and single-cell level (Figs. 1C and 2F). For example, in fibroblasts and CD8 T cells, the majority of patients were in cellular state S01, and in endothelial cells, the cellular state S03 accounted for the highest proportion. Additionally, there existed differences in cellular states distribution between metastatic and primary tumors (Fig. 2G and Additional file 2: Fig. S3E). We used the Fisher's exact test to compare differences in cellular states in primary and metastatic tumors (Fig. 2H), and found that the B cells S01, S02, S03, the dendritic cells S02 and the endothelial cells S03 were predominantly enriched in the metastatic tumors, whereas CD4 T cells S01, S02 as well as the NK cells S01 were mainly enriched in primary tumors. These results indicated that the cellular states we identified are stable and reliable.

We next analyzed the interactions between different cellular states in OV TME, including the number and the strength of interactions, and found that there were interactions between the immune and the stromal cellular states (Additional file 2: Fig. S4A-B), such as an increase in the strength of interactions between CD4 T, CD8 T, and the NK cellular states with the epithelial and the fibroblast cellular states. We further investigated the interactions between cellular states under different signaling pathways and observed an increase in communication between CD8 T cellular states and fibroblast cellular states within the CXCL signaling pathway (Additional file 2: Fig. S4C). Notably, the CXCL12-CXCR4 pair played a critical role in this pathway (Additional file 2: Fig. S4D). It was shown that CXCL12-CXCR4 signaling drove OV cell proliferation, survival and invasion, which could be a promising therapeutic target for OV [59].

Under PARs signaling pathway, CD8 T cellular states and endothelial cellular states communication was increased (Additional file 2: Fig. S4E), with the GZMA-F2R pair playing a significant role (Additional file 2: Fig. S4F). The inefficient communication of GZMA-F2R was correlated with the defective PD-1 mAb therapy [60], suggesting that immunotherapy may benefit the two cellular states. Under the BAG signaling pathway, the communication between NK cellular states and endothelial cellular states was increased (Additional file 2: Fig. S4G), with BAG6-NCR3-PS playing a significant role (Additional file 2: Fig. S4H). Under the VEGF signaling pathway, the communication between Mon/Mac cellular states and endothelial cellular states was enhanced (Additional file 2: Fig. S4I), with VEGFA-VEGFR1R2 playing a crucial role (Additional file 2: Fig. S4J), suggesting that the two cellular states may be associated with the promotion of tumor angiogenesis [61]. These results provided potential therapeutic targets for OV.

Cellular States are enriched in various cancer hallmark pathways

We next aimed to investigate the potential biological functions of the identified cellular states. To achieve this, we employed a hypergeometric test to perform enrichment analysis of hallmark-related and immune-related pathways based on the marker genes of each cellular state. We observed that a variety of cellular states were significantly enriched in hallmark-related pathways (Fig. 3A and Additional file 1: Table S3). For example, epithelial cells S02, fibroblast cells S01 and endothelial cells S03, which were significantly enriched in 'epithelial mesenchymal transition', PMNs S02 and CD4 T cells S03 were enriched in 'tnfa signaling via nfkb', B cells S02 and Monocytes/Macrophages S04 were enriched in 'kras signaling up'. In particular, it has been proven that EMT is closely associated with OV metastasis, chemotherapy resistance and progression, and targeting EMT is expected to prevent OV progression [62]. Currently, several experimental studies have demonstrated that promoting or inhibiting EMT can influence the progression of OV (Additional file 1: Table S4). Our study further revealed several key marker genes including LGALS1, THBS2 and FAP, which were closely associated with EMT, and all of these genes have been validated by rigorous experiments and confirmed to induce EMT in OV and promote the metastasis and invasion of OV cells [63–65]. In addition, several lines of evidence indicated that Kras signaling is related to cancer progression of several cancers [66].

We also analyzed the BP biological pathways enriched by marker genes and found that lymphocyte-associated cellular states including B cells states, CD4 T cells states and CD8 T cells states were associated with the
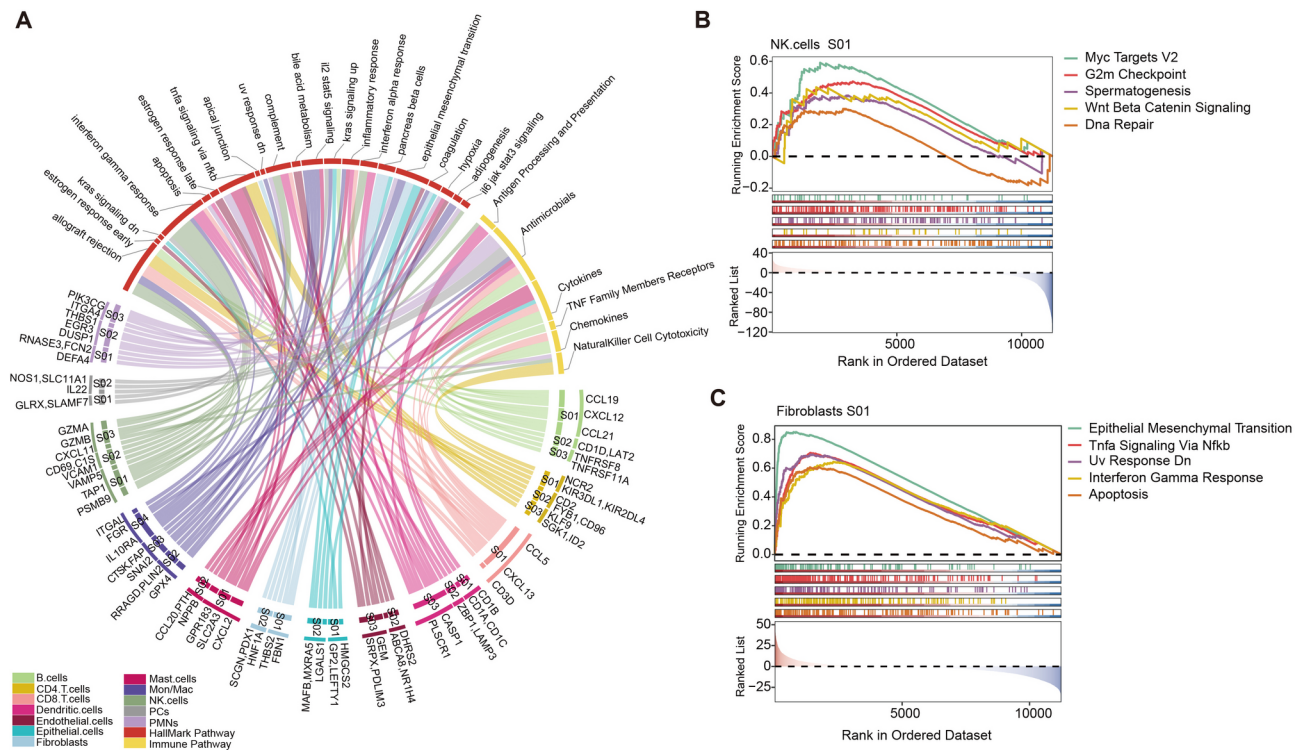
Fig. 3 Functional characterization of cellular states in OV. (**A**) Circos plot demonstrating the genes highly expressed of cellular states are enriched in immune-related pathways and cancer hallmark pathways. The top side displays the enriched pathways, and the bottom side presents the marker genes of the cellular states, the top three marker genes of each cellular state are shown based on the fold change. (**B-C**) GSEA plots showing representative pathways enriched by the differentially expressed genes in (**B**) NK cells S01 and (**C**) fibroblasts S01. Each line is for a pathway

regulation of immune cell activation (Additional file 2: Fig. S5). Furthermore, the marker genes for epithelial cellular states were mainly enriched in pathways such as steroid metabolic process and hormone transport. All these results revealed to some extent the functional heterogeneity of the diverse malignant and immunology cellular states.

NK cells, as part of the innate immune system, are characterized as large granular lymphocytes [67]. It has previously been discovered that NK cells have dual functions: they participate in immune surveillance to eliminate tumor cells, but they also contribute to the development of an immune-tolerant microenvironment that facilitate tumor growth [68, 69]. Recent studies have identified that OV tumor cells can interfere with NK cell function, facilitating the creation of an immune-tolerant microenvironment [70]. Hence, it is important for us to deeply understand the role of NK cells in OV. Next, we analyzed the NK cells S01 in detail and found that cancer-related pathways were significantly enriched (Fig. 3B), including 'Myc Targets V2', 'G2m Checkpoint', 'Spermatogenesis', 'Wnt Beta Catenin Signaling' and 'Dna Repair', suggesting that the markers of NK cells S01 may be related to OV development and progression. In addition, emerging studies have shown that cancerassociated fibroblasts can be used as biomarkers of poor survival outcome [71,

72], which was consistent with the fibroblast S01 marker genes enriched in the cancer-related pathways including 'Epithelial Mesenchymal Transition' and 'Tnfa Signaling Via Nfkb' (Fig. 3C). These findings delineated a functional landscape of diverse cellular states in OV.

Regulation of immunological cellular States in OV

To gain a deeper understanding of the functions of cellular states, we investigated the upstream regulators in OV. In particular, we analyzed NK cells S01 in detail. By utilizing the marker genes of NK cells S01 as a gene set for transcription factor (TF) enrichment analysis, we identified the key transcriptional regulators that regulated the cellular state (Fig. 4A). YY1 is associated with the development of a variety of malignant tumors [73] and is able to drive tumorigenesis, metastasis and drug resistance [74]. Previous studies have demonstrated that YY1 expression is negatively correlated with overall survival in OV patients [75], which aligned with our findings (Fig. 4B). Meanwhile, we also observed that YY1 was highly expressed in S01 (Fig. 4C) and showed the same trend in the validation cohort (Fig. 4C), suggesting that YY1 may be a critical regulator of NK cells S01 marker genes. To construct networks that may regulate gene expression in the NK cells S01, we used the identified the target genes of YY1, retaining only the TF-gene

Zhang *et al. Journal of Translational Medicine*   (2025) 23:519

Page 10 of 18
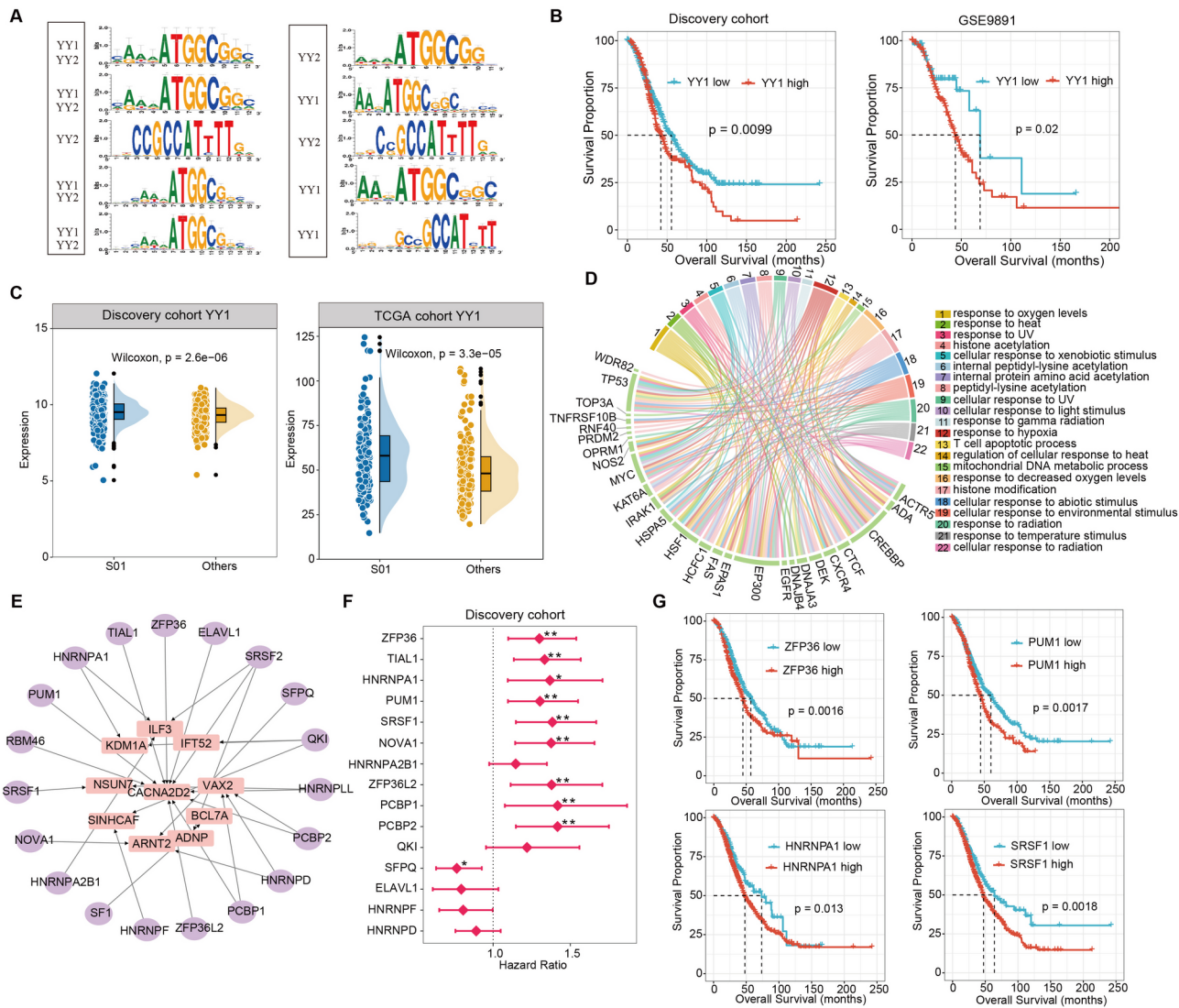
Fig. 4 Potential regulators of NK cells S01 in OV. (**A**) Representative TFs associated with NK cells S01. Top ten TFs are displayed for NK cells S01 of OV. (**B**) Kaplan–Meier curves showing overall survival in the discovery cohort and GSE9891 cohort stratified according to high vs. low expression of YY1. *P*-values are calculated based on log-rank tests. (**C**) Raincloud plots exhibiting the expression of YY1 with NK cells S01 and other cellular states in the discovery cohort and TCGA cohort. *P*-values are calculated based on two-sided Wilcoxon's rank sum tests. (**D**) Circos plot depicting the biological processes pathways enriched by target genes of YY1 in OV. (**E**) Network visualization showing the relationship between genes and RBPs. The purple nodes represent RBPs, and pink nodes represent genes. (**F**) Hazard ratios (HRs) for expression of RBPs correlated with clinical outcome. The center point is the HR and the border is the 95% confidence interval. * indicates *p*-value < 0.05 and ** indicates *p*-value < 0.01. *P*-values are calculated using log-rank tests. (**G**) Kaplan–Meier curves showing overall survival in the discovery cohort stratified according to high vs. low expression of ZFP36, PUM1, HNRNPA1 and SRSF1. *P*-values are calculated using log-rank tests

interactions with a confidence level of 'A'. Ultimately, we identified 73 target genes (Additional file 2: Fig. S6A). In addition, we performed functional enrichment analysis of the target genes of YY1 and found that the genes were mainly enriched in the pathways of response to oxygen levels, response to hypoxia and cellular response (Fig. 4D and Additional file 1: Table S5). Hypoxia is one of the characteristics of cancer, and emerging evidence has suggested that hypoxic TME may be one of the causes of resistance to antitumor therapy and treatment inefficacy

[76], potentially explaining the low survival rate observed in NK cells S01 patients.

In parallel, we also investigated the RNA-binding proteins (RBPs) that regulated the NK cells S01 marker genes. RBPs are class of proteins that specifically bound to RNA, and the aberrant expression of RBPs may lead to a variety of diseases [77]. An increasing number of studies have shown that RBPs-mediated RNA modification can influence cancer progression [78]. We obtained 19 RBPs associated with NK cells S01 based on RBP-gene relationship pairs (Fig. 4E). We further visualized

the RBPs-target relationship (Additional file 2: Fig. S6B), which contained 267 nodes and 519 interaction pairs. The construction of regulatory networks revealed complex interaction patterns between RBPs and their targets, providing new insights into the mechanisms of disease development. In particular, we found most of the RBPs were associated with poor survival outcomes (Fig. 4F and Additional file 2: Fig. S6C-G), while SFPQ exhibited the opposite trend (Additional file 2: Fig. S6H). Notably, RBPs such as ZFP36, PUM1, HNRNPA1 and SRSF1 exhibited a negative correlation between their expression levels and patient survival rates (Fig. 4G). We next compared the expression of RBPs between S01 and other states in NK cells and found 67% of RBPs were significantly highly expressed in S01 (Additional file 2: Fig. S6I). In conclusion, these findings provided valuable clues for us to gain insight into the regulatory mechanisms of NK cells S01.

Immunological cellular States are associated with clinical benefits

Recent studies have explored the association between cellular states and patients prognosis [27, 28, 50]. Thus, to further decipher the relationship between OV cellular states and clinical benefits, we constructed a prognostic atlas for 32 cellular states and categorized them into protective cellular states and adverse cellular states (Fig. 5A and Additional file 1: Table S6). Remarkably, we found that the cellular states corresponding to seven cell types were significantly associated with patient outcomes (Fig. 5A and Additional file 2: Fig. S7A-E). Precisely, six cellular states including NK cells S03, Mon/Mac cells S02, dendritic cells S02, PMNs S03, endothelial cells S02 and fibroblasts S02 were associated with better outcomes (Fig. 5A-B and Additional file 2: Fig. S7F-J). Next, we confirmed the functions of these marker genes of favorable cellular states through the available research (Additional file 1: Table S7) and found that high expression of most of the marker genes inhibited tumor growth. This consistency in the relationship between cellular states and marker genes and patient's survival validated the reliability of our conclusions. By contrast, eight cellular states including Mon/Mac cells S03, PMNs S02, dendritic cells S03, NK cells S01, endothelial cells S03, CD4 T cells S03, B cells S01 and fibroblasts S01 were associated with worse outcomes in OV (Fig. 5A-B and E and Additional file 2: Fig. S7K-P). Similarly, we have elucidated the functions of these marker genes of poor prognostic cellular states from existing studies and found that high expression of most of the marker genes promoted OV cells growth (Additional file 1: Table S7). These findings not only deepened our understanding of the mechanism of action of these specific key genes in tumor development, but also reinforced the strong link between marker genes and cellular states, and ultimately patient's

survival. Further, we explored the two cellular states of fibroblasts and found that the patients assigned to S01 cellular state had inferior outcomes (Fig. 5B), with consistent results observed in the validation cohort (Fig. 5C and Additional file 2: Fig. S7Q-R). Meanwhile, we found that marker genes for the S01 cellular state were significantly enriched in cancer-related signaling pathways such as 'epithelial-mesenchymal transition', 'tgf beta signaling' and 'kras signaling up' (Fig. 5F), which may explain the poor survival of patients. To further explore the association between S01 cellular state and patient outcomes, we compared the proportion of patients at different stages in two fibroblasts cellular states, and found that the majority of patients with S01 were in III and IV stages (Fig. 5G).

Next, we analyzed the association between NK cells states and patients survival. Remarkably, there occurred significant differences in survival among the three cellular states (Fig. 5D). It was worth mentioning that patients with NK cells S01 had shorter overall survival than other states (Fig. 5E), with consistent results observed in the validation cohort (Additional file 2: Fig. S7S). Besides, we found that the marker genes of NK cells S01 were significantly enriched in the MyD88-dependent toll-like receptor signaling pathway (Fig. 5H), which is a junction protein that mediates innate immune signaling [79]. It has been unveiled that MyD88 signaling is associated with tumor progression and immune escape [80]. However, the marker genes for other cellular states were enriched in immune-related pathways, including chronic inflammatory response and response to lectin (Fig. 5H). The signaling pathways enriched by marker genes for S01 cellular state and other cellular states revealed tumor heterogeneity. Together, these results revealed the association between cellular states and patients prognosis, delineating a prognostic landscape of cellular states in OV.

The above results showed that marker genes for cellular states play important roles in OV progression. To obtain key genes for cellular states, we performed feature screening for marker genes associated with adverse cellular states through LASSO regression (Additional file 2: Fig. S8A-B). A total of 62 genes were selected for risk model construction. We stratified the samples into high-risk and low-risk groups based on the median value of the risk score. Patients in the high-risk group showed shorter survival than those in the low-risk group (Additional file 2: Fig. S8C). ROC analyses revealed the diagnostic value of the risk score. The AUCs for 1-, 3- and 5-year survival were 0.75, 0.73 and 0.75, respectively (Additional file 2: Fig. S8D). In addition, we verified these findings in the validation dataset (Additional file 2: Fig. S8E-F), suggested the accuracy of the risk model. The model may serve as promising tool to predict patient prognosis.
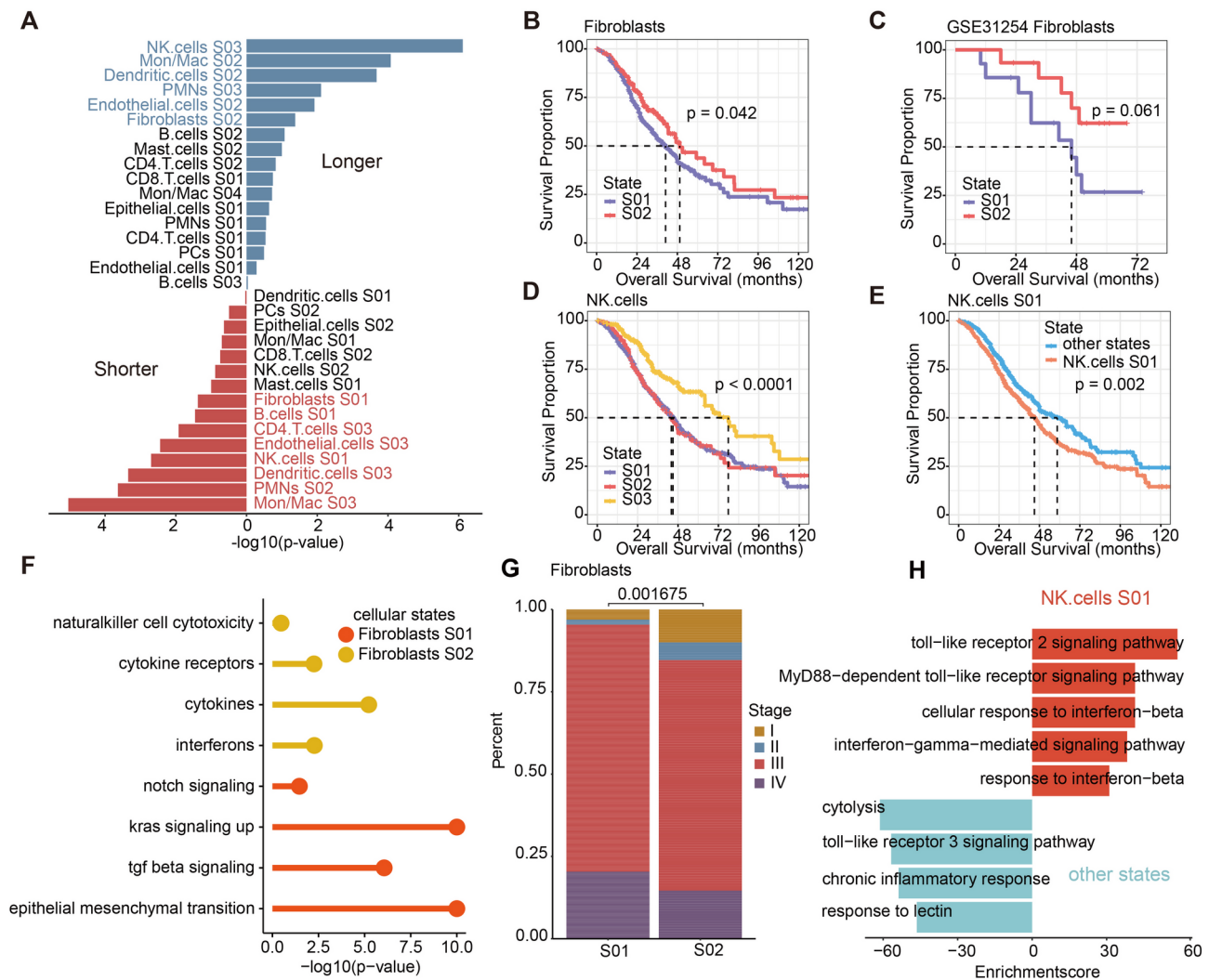
Fig. 5 The prognostic landscape of cellular states in OV. (**A**) Associations between cellular states and survival in the OV discovery cohort. Bar plots depicting the -log10 (*p*-values). Blue represents protective and red represents the risky cellular states. *P*-values are calculated using log-rank tests. (**B**) Kaplan–Meier curves demonstrating overall survival differences for patients with two fibroblasts cellular states. *P*-value is calculated using log-rank test. (**C**) Kaplan–Meier curves showing overall survival differences for patients with two fibroblasts cellular states in validation cohort. *P*-value is calculated using log-rank test. (**D**) Kaplan–Meier curves depicting overall survival differences for patients with three NK cellular states. *P*-value is calculated using log-rank test. (**E**) Kaplan–Meier curves showing overall survival differences for patients with NK cells S01 and other cellular states. *P*-value is calculated using log-rank test. (**F**) Dot plots depicting the functional pathways enriched by genes highly expressed in two fibroblasts cellular states. *P*-values are calculated using permutation tests. (**G**) Bar plots showing the proportion of different stages in two fibroblasts cellular states of patients. *P*-values are calculated using two-sided Fisher's exact tests. (**H**) Bar plots showing the functional pathways enriched by genes highly expressed in NK cells S01 and other cellular states. Red represents the pathways for NK cells S01 and blue represents the pathways for other states
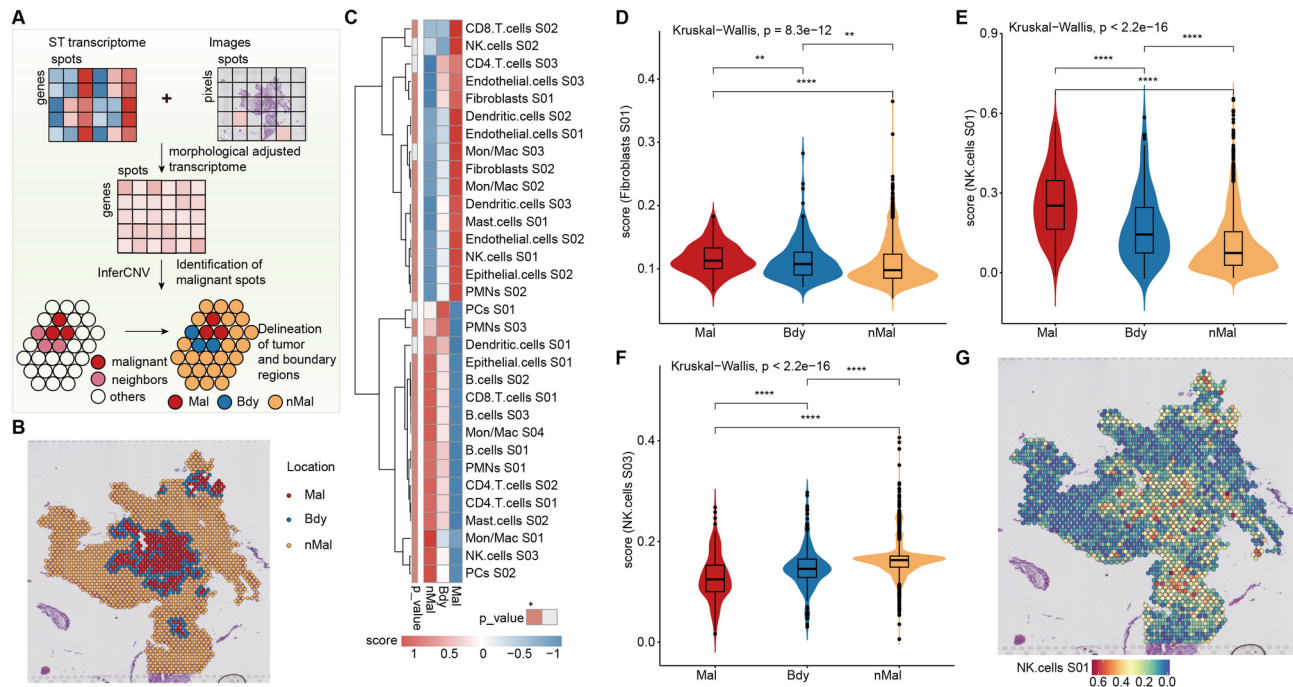
Dissecting the Spatial patterns of cellular States in OV

Next, we dissected the spatial microenvironment of OV tumors. We used Cottrazm [81] for the construction of the microenvironment around tumor boundaries by integrating spatial transcriptomics (ST) with histological imaging, and single-cell transcriptomics. We firstly adjusted gene expression data by utilizing morphology information and subsequently employed inferCNV to identify Mal spots (Fig. 6A). Next, we used Cottrazm to classify the spots as either Mal spots or Bdy spots based on the distance of the UMAP to the center of the tumor centroid (Fig. 6A).

We depicted Mal regions, Bdy regions and nMal regions of OV ST data (Fig. 6B). Notably, distinct tumor boundaries can be clearly observed in the tissue section (Fig. 6B). We next calculated the scores of marker genes for cellular states in Mal, Bdy and nMal spots (Fig. 6C and Additional file 1: Table S8), which showed that 13 cellular states had higher scores in Mal spots and 13 cellular states had higher scores in nMal spots. For example, we found that fibroblasts S01 and NK cells S01 had the

Fig. 6 The spatial patterns of cellular states in OV. (**A**) Overview of delineation of tumor boundary. (**B**) Tissue slide of OV are annotated by malignant spots (Mal, red), boundary spots (Bdy, blue) and non-malignant spots (nMal, orange). (**C**) Heatmap showing the scores of marker genes for cellular states in Mal, Bdy and nMal regions. * indicates *p*-value < 0.05. (**D-F**) Violin plots showing the scores of marker genes for cellular states in Mal, Bdy and nMal regions. *P*-values are calculated using Kruskal-Wallis tests followed by Wilcoxon's rank sum tests for pairwise comparisons. (**D**) for fibroblasts S01. (**E**) for NK cells S01. (**F**) for NK cells S03. ** indicates *p*-value < 0.01 and **** indicates *p*-value < 0.0001 (**G**) Spatial feature plots of marker genes scores of NK cells S01 in OV

highest scores in Mal spots, lower scores in Bdy spots, and the lowest scores in nMal spots (Fig. 6D-E, and Fig. 6G). This was consistent with our previous findings that fibroblasts S01 and NK cells S01 were correlated with poor outcomes of OV patients (Fig. 5A). Meanwhile, we also found that NK cells S03 had the highest scores in nMal spots (Fig. 6F), which was consistent with the better outcomes of NK cells S03 (Fig. 5A). These results suggested the diverse spatial patterns of the cellular states in OV.

Immunological subtypes are associated with drug responses in OV patients

Next, we explored the relationship between cellular states and drug responses. Initially, we categorized cellular states into sensitive and resistant based on the survival times of responders and non-responders (Additional file 1: Table S9). Notably, we found that responders showed longer survival times at 18 cellular states (Fig. 7A). Furthermore, we compared the drug responses of cellular states to both platinum/cyclophosphamide and platinum/taxane (Fig. 7B). Remarkably, a majority of cellular states were significantly sensitive to platinum/taxane (Fig. 7B). Specifically, NK cells S01 showed favorable efficacy to two treatment drugs (Fig. 7B). Additionally, we observed that responders who underwent treatment with these two drugs had significantly better outcomes than

non-responders (Fig. 7C-D). Next, we employed Wilcoxon's rank sum test to identify differentially expressed genes in sensitive and resistant patients of NK cells S01 (Additional file 2: Fig. S9A). Obviously, SCRN1 was highly expressed in the drug-resistant group (Fig. 7E). Our observation was in line with the notion that SCRN1 was a risk factor for OV patients (Additional file 2: Fig. S9B). Moreover, STARD5 is a lipid-binding protein and copy number variations of STARD5 were associated with paclitaxel-platinum chemosensitivity in OV [82]. We also found that STARD5 was highly expressed in the drug-sensitive group and was a protective factor for OV patients (Fig. 7F and Additional file 2: Fig. S9C).

To explore whether cellular states can broadly predict the drug responses, we used the 'oncoPredict' R package [41] to calculate the response values of TCGA-OV data for 743 drugs, including 198 drugs from GDSC2 and 545 drugs from CTRP. We discovered that most cellular states can potentially predict the responses to OV drugs (Fig. 7G). Furthermore, we observed that patients with NK cells S01, dendritic cells S02 and S03 cellular states were sensitive to most drugs (Fig. 7G).

We analyzed in detail the survival response of patients with NK cells S01 cellular state to the drugs and observed that survival time was significantly prolonged in patients who were sensitive to Cisplatin, Olaparib, cediranib and niclosamide drug treatments (Additional file 2: Fig.
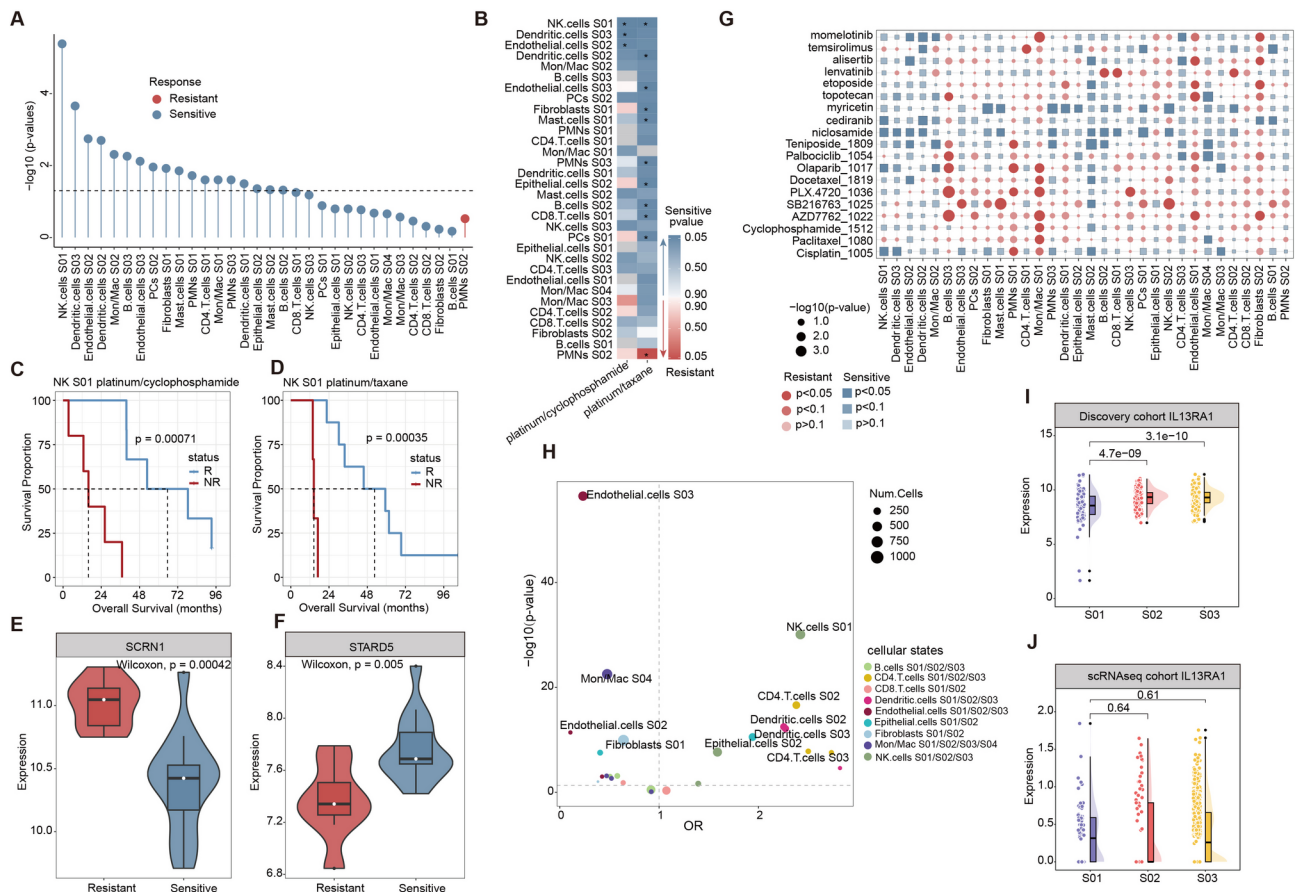
Fig. 7 The association between cellular states and drug therapy. (**A**) Relationship between cellular states and drug responses in OV. Cellular states are sorted by -log10(*p*-values) of the correlations with overall survival. The dotted line represents *p*-value = 0.05. *P*-values are calculated using log-rank tests. (**B**) Heat map showing the association between cellular states with drug responses for platinum/cyclophosphamide and platinum/taxane treatment, separately. * indicates *p*-value < 0.05. *P*-values are calculated using log-rank tests. (**C-D**) Kaplan-Meier analysis of patients in the responders and non-responders groups in NK cells S01. **C** for platinum/cyclophosphamide. **D** for platinum/taxane. *P*-values are calculated using log-rank tests. (**E-F**) Violin plots showing the expression of differential genes for resistant and sensitive groups in NK cells S01. (**E**) for SCRN1. (**F**) for STARD5. *P*-values are calculated using two-sided Wilcoxon's rank sum tests. (**G**) Correlation between cellular states and therapeutic benefits for drugs used for OV therapy in TCGA OV cohort. Circles for resistance and squares for sensitivity. *P*-values are calculated using log-rank tests. (**H**) Scatter plot showing the differentially enriched cellular states between sensitive and resistant tumor groups. The dotted line represents *p*-value = 0.05 and OR = 1. The OR is greater than 1 indicates that the cellular states are enriched in the sensitive tumor, while the OR is less than 1 indicates that they are enriched in the resistant tumor. *P*-values are calculated using two-sided Fisher's exact tests. (**I**) Raincloud plot showing the expression of IL13RA1 of Endothelial cells in the discovery cohort. *P*-values are calculated using two-sided Wilcoxon's rank sum tests. (**J**) Raincloud plot showing the expression of IL13RA1 of Endothelial cells in the scRNA-seq cohort. *P*-values are calculated using two-sided Wilcoxon's rank sum tests

S9D-G). Meanwhile, we compared the distribution of cellular states between chemotherapy-sensitive and chemotherapy-resistant groups using the Fisher's exact test at the single-cell level (Fig. 7H). We observed that NK cells S01, dendritic cells S02, S03 cellular states were also significantly enriched in the chemotherapy-sensitive group, which was consistent with the previous conclusion (Fig. 7G and H). This to some extent reflected the accuracy of cellular states in predicting drug responses.

We observed that endothelial cells S02, S03 were significantly enriched in the chemotherapy-resistant group (Fig. 7H). Recent studies have revealed that the enrichment of IL13RA1+ Endothelial cells suggested that patients were resistant to chemotherapy [22]. We

compared the expression of IL13RA1 in the discovery cohort and the single-cell cohort, and found that IL13RA1 was highly expressed in S02, S03 cellular states (Fig. 7I-J). These consistent findings further validated specific cellular states can represent candidate predictive biomarkers of effective responses to chemotherapy in OV patients.

Discussion

TME is the complex ecosystem of non-malignant cells, endothelial cells, pericytes, immune cells, bone marrow derived cells and extracellular matrix [83]. Together, they play crucial roles for tumor growth, invasion, metastasis and therapy response [84], it is critical to

comprehensively decode the cellular composition and cellular state of the TME. Recent studies have confirmed that parsing cellular states from gene expression data to define cancer subtypes has become an important tool for characterizing TME at high resolution [27, 28]. In this study, we applied EcoTyper, a machine learning framework, to comprehensively characterize the TME of OV patients. Specifically, we obtained 32 diverse transcriptional cellular states and three cellular communities from 12 cell types, many of which were significantly prognostic for patient outcomes. We validated these cellular states using large-scale independent cohorts. These cellular states provided a unique perspective to characterize OV heterogeneity.

In addition, we identified cellular states associated with cancer progression and clinical outcomes. Firstly, we performed functional enrichment analysis of marker genes for cellular states and found that numerous cancer hallmark pathways were significantly enriched. Further, to explore the upstream regulatory mechanisms of cellular states, we identified key TFs and RBPs in OV, focusing on NK cells S01 cellular state. We identified TFs included YY1, which was a poor prognostic marker. High expression of YY1 was correlated with patient poorer survival. It has been reported that that YY1 positively regulated the stemness of OV cells by constructing gain- and loss-of-function experiments, which may explain the negative association between YY1 expression and patient outcomes [75]. Simultaneously, 19 RBPs (e.g. ZFP36, PUM1, and HNRNPA1) and SRSF1 were also implicated. It was discovered that PUM1 can induce OV cell growth, migration and invasion based on wound-healing and transwell experiment. Thus, PUM1 high expression was negatively associated with patient survival [85], which may occur via the above-mentioned molecular mechanisms. Notably, most of the RBPs were significantly associated with worse outcomes. The discovery of these key regulators was expected to be potential candidate targets in future experimental studies. In addition, our research revealed key signaling pathway (epithelial mesenchymal transition, EMT). Several lines of experimental studies suggested that promoting or inhibiting EMT can influence the progression of OV (Additional file 1: Table S4), suggesting EMT played significant roles in OV development. We also found that the drugs we identified have been proven to be used for the treatment of OV validated by experiment (Additional file 1: Table S10). These results provided strong supports for our conclusions. Furthermore, in order to explore the association between cellular states and patients prognosis, we constructed a prognostic maps for 32 cellular states, and 8 cellular states can be used as OV survival risk factors, whereas 6 cellular states can be used as OV survival protective factors. These

findings provided a prognostic landscape of the cellular states that determine patient outcomes in OV.

Current studies have utilized spatial transcriptomics to construct the microenvironment around tumor boundaries [81]. Here, we used Cottrazm to dissect the spatial microenvironment of OV and explored differences in the scores of the NK cells S01 in the Mal, Bdy and nMal spots. It was observed that NK cells S01 exhibited the highest scores in the Mal regions, which was consistent with the poor survival. This revealed that NK cells S01 was closely associated with OV progression. Finally, we explored the relationship between cellular states and OV therapeutic benefits. We observed that treatment effects of OV can be predicted based on cellular states. In particular, for patients with NK cells S01, there was a significant prolongation of survival time after appropriate drug treatment. We also predicted the response values to 743 drug treatments and found that NK cells S01 was significantly sensitive to most drugs such as Cisplatin, Olaparib, cediranib and niclosamide. Moreover, we also found NK cells S01 significantly enriched in the chemotherapy-sensitive group at the single-cell level. These results revealed the predictive potential of cellular states for clinical benefit.

Conclusion
In summary, we performed a comprehensive analysis of TME in OV patients by bulk, single-cell and spatially resolved gene expression data. The results of this study will help provide new directions and targets for the diagnosis and treatment of OV.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12967-025-06521-3.

Additional file 1

Additional file 2

Zhang *et al. Journal of Translational Medicine*   (2025) 23:519

Page 16 of 18

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]College of Biomedical Information and Engineering, Hainan Medical University, Haikou 571199, China
[2]School of Interdisciplinary Medicine and Engineering, Harbin Medical University, Harbin 150081, China
[3]College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, Heilongjiang Province, China
[4]Department of Radiation Oncology, Harbin Medical University Cancer Hospital, 150 Haping Road, Nangang District, Harbin 150040, Heilongjiang, China
[5]Department of Anesthesiology, The First Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, China

References

1. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. CA Cancer J Clin. 2023;73(1):17–48.
2. Matulonis UA, Sood AK, Fallowfield L, Howitt BE, Sehouli J, Karlan BY. Ovarian cancer. Nat Rev Dis Primers. 2016;2:16061.
3. Chobanian N, Dietrich CS. 3rd: ovarian cancer. Surg Clin North Am. 2008;88(2):285–99. vi.
4. Pawlowska A, Rekowska A, Kurylo W, Panczyszyn A, Kotarski J, Wertel I. Current Understanding on why ovarian Cancer is resistant to immune checkpoint inhibitors. Int J Mol Sci 2023, 24(13).
5. Lheureux S, Braunstein M, Oza AM. Epithelial ovarian cancer: evolution of management in the era of precision medicine. CA Cancer J Clin. 2019;69(4):280–304.
6. Sun Y, Yan C, Xu D, Zhang Z, Li K, Li X, Zhou M, Hao D. Immuno-genomic characterisation of high-grade serous ovarian cancer reveals immune evasion mechanisms and identifies an immunological subtype with a favourable prognosis and improved therapeutic efficacy. Br J Cancer. 2022;126(11):1570–80.
7. Baslan T, Hicks J. Unravelling biology and shifting paradigms in cancer with single-cell sequencing. Nat Rev Cancer. 2017;17(9):557–69.
8. Zheng GX, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, Ziraldo SB, Wheeler TD, McDermott GP, Zhu J, et al. Massively parallel digital transcriptional profiling of single cells. Nat Commun. 2017;8:14049.
9. Villani AC, Satija R, Reynolds G, Sarkizova S, Shekhar K, Fletcher J, Griesbeck M, Butler A, Zheng S, Lazo S et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. Science. 2017;356(6335).
10. Kelliher L, Lengyel E. Understanding Long-Term survival of patients with ovarian Cancer-The tumor microenvironment comes to the forefront. Cancer Res. 2023;83(9):1383–5.
11. Yan C, Li K, Meng F, Chen L, Zhao J, Zhang Z, Xu D, Sun J, Zhou M. Integrated Immunogenomic analysis of single-cell and bulk tissue transcriptome profiling unravels a macrophage activation paradigm associated with immunologically and clinically distinct behaviors in ovarian cancer. J Adv Res. 2023;44:149–60.
12. Yang L, Wang S, Zhang Q, Pan Y, Lv Y, Chen X, Zuo Y, Hao D. Clinical significance of the immune microenvironment in ovarian cancer patients. Mol Omics. 2018;14(5):341–51.
13. Rolong A, Chen B, Lau KS. Deciphering the cancer microenvironment from bulk data with ecotyper. Cell. 2021;184(21):5306–8.
14. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, et al. NCBI GEO: archive for functional genomics data sets–update. Nucleic Acids Res. 2013;41(Database issue):D991–995.
15. Tomczak K, Czerwinska P, Wiznerowicz M. The Cancer genome atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn). 2015;19(1A):A68–77.
16. Ganzfried BF, Riester M, Haibe-Kains B, Risch T, Tyekucheva S, Jazic I, Wang XV, Ahmadifar M, Birrer MJ, Parmigiani G et al. curatedOvarianData: clinically annotated data for the ovarian cancer transcriptome. Database (Oxford). 2013;2013:bat013.
17. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy–analysis of affymetrix genechip data at the probe level. Bioinformatics. 2004;20(3):307–15.
18. Sun J, Shi R, Zhang X, Fang D, Rauch J, Lu S, Wang X, Kasmann L, Ma J, Belka C, et al. Characterization of immune landscape in papillary thyroid cancer reveals distinct tumor immunogenicity and implications for immunotherapy. Oncoimmunology. 2021;10(1):e1964189.
19. Glaab E, Garibaldi JM, Krasnogor N. ArrayMining: a modular web-application for microarray analysis combining ensemble and consensus methods with cross-study normalization. BMC Bioinformatics. 2009;10:358.
20. Warnat P, Eils R, Brors B. Cross-platform analysis of cancer microarray data improves gene expression based classification of phenotypes. BMC Bioinformatics. 2005;6:265.
21. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The Sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012;28(6):882–3.
22. Zheng X, Wang X, Cheng X, Liu Z, Yin Y, Li X, Huang Z, Wang Z, Guo W, Ginhoux F, et al. Single-cell analyses implicate Ascites in remodeling the ecosystems of primary and metastatic tumors in ovarian cancer. Nat Cancer. 2023;4(8):1138–56.
23. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. 2018;36(5):411–20.
24. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, Myung P, Plikus MV, Nie Q. Inference and analysis of cell-cell communication using cellchat. Nat Commun. 2021;12(1):1088.
25. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, Khodadoust MS, Esfahani MS, Luca BA, Steiner D, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat Biotechnol. 2019;37(7):773–82.
26. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12(5):453–7.
27. Luca BA, Steen CB, Matusiak M, Azizi A, Varma S, Zhu C, Przybyl J, Espin-Perez A, Diehn M, Alizadeh AA, et al. Atlas of clinically distinct cell States and ecosystems across human solid tumors. Cell. 2021;184(21):5482–e54965428.
28. Steen CB, Luca BA, Esfahani MS, Azizi A, Sworder BJ, Nabet BY, Kurtz DM, Liu CL, Khameneh F, Advani RH, et al. The landscape of tumor cell States and ecosystems in diffuse large B cell lymphoma. Cancer Cell. 2021;39(10):1422–e14371410.
29. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, et al. ClusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innov (Camb). 2021;2(3):100141.
30. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. Cell Syst. 2015;1(6):417–25.
31. Bhattacharya S, Andorf S, Gomes L, Dunn P, Schaefer H, Pontius J, Berger P, Desborough V, Smith T, Campbell J, et al. ImmPort: disseminating data to the public for the future of immunology. Immunol Res. 2014;58(2–3):234–9.
32. Li Y, Jiang T, Zhou W, Li J, Li X, Wang Q, Jin X, Yin J, Chen L, Zhang Y, et al. Pan-cancer characterization of immune-related LncRNAs identifies potential oncogenic biomarkers. Nat Commun. 2020;11(1):1000.
33. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545–50.

34. Giudice G, Sanchez-Cabo F, Torroja C, Lara-Pezzi E. ATtRACT-a database of RNA-binding proteins and associated motifs. Database (Oxford). 2016;2016.

35. Zhou W, Jie Q, Pan T, Shi J, Jiang T, Zhang Y, Ding N, Xu J, Ma Y, Li Y. Single-cell RNA binding protein regulatory network analyses reveal oncogenic HNRNPK-MYC signalling pathway in cancer. Commun Biol. 2023;6(1):82.

36. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498–504.

37. Van de Sande B, Flerin C, Davie K, De Waegeneer M, Hulselmans G, Aibar S, Seurinck R, Saelens W, Cannoodt R, Rouchon Q, et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis. Nat Protoc. 2020;15(7):2247–76.

38. Garcia-Alonso L, Holland CH, Ibrahim MM, Turei D, Saez-Rodriguez J. Benchmark and integration of resources for the Estimation of human transcription factor activities. Genome Res. 2019;29(8):1363–75.

39. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, Bindal N, Beare D, Smith JA, Thompson IR, et al. Genomics of drug sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2013;41(Database issue):D955–961.

40. Seashore-Ludlow B, Rees MG, Cheah JH, Cokol M, Price EV, Coletti ME, Jones V, Bodycombe NE, Soule CK, Gould J, et al. Harnessing connectivity in a Large-Scale Small-Molecule sensitivity dataset. Cancer Discov. 2015;5(11):1210–23.

41. Maeser D, Gruener RF, Huang RS. OncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. Brief Bioinform. 2021;22(6).

42. Zhang L, Conejo-Garcia JR, Katsaros D, Gimotty PA, Massobrio M, Regnani G, Makrigiannakis A, Gray H, Schlienger K, Liebman MN, et al. Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. N Engl J Med. 2003;348(3):203–13.

43. Sato E, Olson SH, Ahn J, Bundy B, Nishikawa H, Qian F, Jungbluth AA, Frosina D, Gnjatic S, Ambrosone C, et al. Intraepithelial CD8 + tumor-infiltrating lymphocytes and a high CD8+/regulatory T cell ratio are associated with favorable prognosis in ovarian cancer. Proc Natl Acad Sci U S A. 2005;102(51):18538–43.

44. Tomsova M, Melichar B, Sedlakova I, Steiner I. Prognostic significance of CD3 + tumor-infiltrating lymphocytes in ovarian carcinoma. Gynecol Oncol. 2008;108(2):415–20.

45. Ohno H, Ushiyama C, Taniguchi M, Germain RN, Saito T. CD2 can mediate TCR/CD3-independent T cell activation. J Immunol. 1991;146(11):3742–6.

46. Pan J, Chen X, Yang L, Song Y, Liu J, Li L, Lin Y, Tang L, Qiu S, Xu Q. A comprehensive analysis of prognostic indicators in serous ovarian Cancer based on leukocyte migration and immune microenvironment. Front Biosci (Landmark Ed). 2023;28(6):130.

47. Feng S, Isayev O, Werner J, Bazhin AV. CD96 as a potential immune regulator in cancers. Int J Mol Sci. 2023;24(2).

48. Yang M, Lu J, Zhang G, Wang Y, He M, Xu Q, Xu C, Liu H. CXCL13 shapes immunoactive tumor microenvironment and enhances the efficacy of PD-1 checkpoint Blockade in high-grade serous ovarian cancer. J Immunother Cancer. 2021;9(1).

49. Han Y, Guo Z, Jiang L, Li X, Chen J, Ouyang L, Li Y, Wang X. CXCL10 and CCL5 as feasible biomarkers for immunotherapy of homologous recombination deficient ovarian cancer. Am J Cancer Res. 2023;13(5):1904–22.

50. Subramanian A, Nemat-Gorgani N, Ellis-Caleo TJ, van Sears IDGP, Somani TJ, Luca A, Zhou BA, Bradic MY, Torres M. IA: Sarcoma microenvironment cell States and ecosystems are associated with prognosis and predict response to immunotherapy. Nat Cancer. 2024.

51. Wang M, Wang J, Liu J, Zhu L, Ma H, Zou J, Wu W, Wang K. Systematic prediction of key genes for ovarian cancer by co-expression network analysis. J Cell Mol Med. 2020;24(11):6298–307.

52. Wang Z, Liu Y, Lu L, Yang L, Yin S, Wang Y, Qi Z, Meng J, Zang R, Yang G. Fibrillin-1, induced by Aurora-A but inhibited by BRCA2, promotes ovarian cancer metastasis. Oncotarget. 2015;6(9):6670–83.

53. Fan X, Wang C, Song X, Liu H, Li X, Zhang Y. Elevated cathepsin K potentiates metastasis of epithelial ovarian cancer. Histol Histopathol. 2018;33(7):673–80.

54. Zhang J, Zhang J, Wang F, Xu X, Li X, Guan W, Men T, Xu G. Overexpressed COL5A1 is correlated with tumor progression, Paclitaxel resistance, and tumor-infiltrating immune cells in ovarian cancer. J Cell Physiol. 2021;236(10):6907–19.

55. Roett MA, Evans P. Ovarian cancer: an overview. Am Fam Physician. 2009;80(6):609–16.

56. Astorgues-Xerri L, Riveiro ME, Tijeras-Raballand A, Serova M, Neuzillet C, Albert S, Raymond E, Faivre S. Unraveling galectin-1 as a novel therapeutic target for cancer. Cancer Treat Rev. 2014;40(2):307–19.

57. Li X, Wang H, Jia A, Cao Y, Yang L, Jia Z. LGALS1 regulates cell adhesion to promote the progression of ovarian cancer. Oncol Lett. 2023;26(2):326.

58. Cibrian D, Sanchez-Madrid F. CD69: from activation marker to metabolic gatekeeper. Eur J Immunol. 2017;47(6):946–53.

59. Salomonnson E, Stacer AC, Ehrlich A, Luker KE, Luker GD. Imaging CXCL12-CXCR4 signaling in ovarian cancer therapy. PLoS ONE. 2013;8(1):e51500.

60. Gao Y, Xu Q, Li X, Guo Y, Zhang B, Jin Y, Zhu C, Shen Y, Yang P, Shi Y, et al. Heterogeneity induced GZMA-F2R communication inefficient impairs antitumor immunotherapy of PD-1 mAb through JAK2/STAT1 signal suppression in hepatocellular carcinoma. Cell Death Dis. 2022;13(3):213.

61. Li S, Lin Y, Gao X, Zeng D, Cen W, Su Y, Su J, Zeng C, Huang Z, Zeng H, et al. Integrative multi-omics analysis reveals a novel subtype of hepatocellular carcinoma with biological and clinical relevance. Front Immunol. 2024;15:1517312.

62. Deng J, Wang L, Chen H, Hao J, Ni J, Chang L, Duan W, Graham P, Li Y. Targeting epithelial-mesenchymal transition and cancer stem cells for chemoresistant ovarian cancer. Oncotarget. 2016;7(34):55771–88.

63. Zhu J, Zheng Y, Zhang H, Liu Y, Sun H, Zhang P. Galectin-1 induces metastasis and epithelial-mesenchymal transition (EMT) in human ovarian cancer cells via activation of the MAPK JNK/p38 signalling pathway. Am J Transl Res. 2019;11(6):3862–78.

64. Gardi NL, Deshpande TU, Kamble SC, Budhe SR, Bapat SA. Discrete molecular classes of ovarian cancer suggestive of unique mechanisms of transformation and metastases. Clin Cancer Res. 2014;20(1):87–99.

65. Fan L, Lei H, Zhang S, Peng Y, Fu C, Shu G, Yin G. Non-canonical signaling pathway of SNAI2 induces EMT in ovarian cancer cells by suppressing miR-222-3p transcription and upregulating PDCD10. Theranostics. 2020;10(13):5895–5913.

66. Tokumaru Y, Oshi M, Katsuta E, Yan L, Satyananda V, Matsuhashi N, Futamura M, Akao Y, Yoshida K, Takabe K. KRAS signaling enriched triple negative breast cancer is associated with favorable tumor immune microenvironment and better survival. Am J Cancer Res. 2020;10(3):897–907.

67. Caligiuri MA. Human natural killer cells. Blood. 2008;112(3):461–9.

68. Chiossone L, Dumas PY, Vienne M, Vivier E. Natural killer cells and other innate lymphoid cells in cancer. Nat Rev Immunol. 2018;18(11):671–88.

69. Morvan MG, Lanier LL. NK cells and cancer: you can teach innate cells new tricks. Nat Rev Cancer. 2016;16(1):7–19.

70. Gonzalez VD, Huang YW, Delgado-Gonzalez A, Chen SY, Donoso K, Sachs K, Gentles AJ, Allard GM, Kolahi KS, Howitt BE, et al. High-grade serous ovarian tumor cells modulate NK cell function to create an immune-tolerant microenvironment. Cell Rep. 2021;36(9):109632.

71. Yang W, Han W, Ye S, Liu D, Wu J, Liu H, Li C, Chen H. Fibroblast activation protein-alpha promotes ovarian cancer cell proliferation and invasion via extracellular and intracellular signaling mechanisms. Exp Mol Pathol. 2013;95(1):105–10.

72. Mhawech-Fauceglia P, Yan L, Sharifian M, Ren X, Liu S, Kim G, Gayther SA, Pejovic T, Lawrenson K. Stromal expression of fibroblast activation protein alpha (FAP) predicts platinum resistance and shorter recurrence in patients with epithelial ovarian Cancer. Cancer Microenviron. 2015;8(1):23–31.

73. Fu X, Ji F, He Q, Qiu X. A Systematic Pan-Cancer Analysis of YY1 Aberrations and their Relationship with Clinical Outcome, Tumor Microenvironment, and Therapeutic Targets. J Immunol Res. 2022;2022:5826741.

74. Kaufhold S, Garban H, Bonavida B. Yin Yang 1 is associated with cancer stem cell transcription factors (SOX2, OCT4, BMI1) and clinical implication. J Exp Clin Cancer Res. 2016;35:84.

75. Qian S, Wang W, Li M. Transcriptional factor Yin Yang 1 facilitates the stemness of ovarian cancer via suppressing miR-99a activity through enhancing its deacetylation level. Biomed Pharmacother. 2020;126:110085.

76. Liao C, Liu X, Zhang C, Zhang Q. Tumor hypoxia: from basic knowledge to therapeutic implications. Semin Cancer Biol. 2023;88:172–86.

77. Mehta M, Raguraman R, Ramesh R, Munshi A. RNA binding proteins (RBPs) and their role in DNA damage and radiation response in cancer. Adv Drug Deliv Rev. 2022;191:114569.

78. Iino K, Mitobe Y, Ikeda K, Takayama KI, Suzuki T, Kawabata H, Suzuki Y, Horie-Inoue K, Inoue S. RNA-binding protein NONO promotes breast cancer proliferation by post-transcriptional regulation of SKP2 and E2F8. Cancer Sci. 2020;111(1):148–59.

79. Deguine J, Barton GM. MyD88: a central player in innate immune signaling. F1000Prime Rep. 2014;6:97.

Zhang *et al. Journal of Translational Medicine*   (2025) 23:519

Page 18 of 18

80. Huang B, Zhao J, Li H, He KL, Chen Y, Mayer L, Unkeless JC, Xiong H: 's Note, editors. Toll-Like Receptors on Tumor Cells Facilitate Evasion of Immune Surveillance. Cancer Res. 2019;79(16):4305.

81. Xun Z, Ding X, Zhang Y, Zhang B, Lai S, Zou D, Zheng J, Chen G, Su B, Han L, et al. Reconstruction of the tumor Spatial microenvironment along the malignant-boundary-nonmalignant axis. Nat Commun. 2023;14(1):933.

82. Li N, Hou JL, Shi ZZ, Li XG, Li N, Sun YC, Xu X, Cai Y, Zhang X, Zhang KT, et al. Copy number changes of 4-gene set May predict early relapse in advanced epithelial ovarian cancer after initial platinum-paclitaxel chemotherapy. Am J Cancer Res. 2014;4(3):285–92.

83. Belli C, Trapani D, Viale G, D'Amico P, Duso BA, Della Vigna P, Orsi F, Curigliano G. Targeting the microenvironment in solid tumors. Cancer Treat Rev. 2018;65:22–32.

84. Elhanani O, Ben-Uri R, Keren L. Spatial profiling technologies illuminate the tumor microenvironment. Cancer Cell. 2023;41(3):404–20.

85. Guan X, Chen S, Liu Y, Wang LL, Zhao Y, Zong ZH. PUM1 promotes ovarian cancer proliferation, migration and invasion. Biochem Biophys Res Commun. 2018;497(1):313–8.

Publisher's note