Characterizing the tumor microenvironment of metastatic ovarian cancer by single cell transcriptomics

Susan Olalekan1*, Bingqing Xie1, Rebecca Back1, Heather Eckart1, Anindita Basu1,2*

1 Department of Medicine, University of Chicago
2 Center for Nanoscale Materials, Argonne National Laboratory

*Correspondence: Anindita Basu (onibasu@uchicago.edu), Susan Olalekan (solalekan@uchicago.edu)
Abstract

Epithelial ovarian cancer is a highly heterogenous disease with a common metastatic site, the omentum. The presence of CD8⁺T cells within the tumor is associated with overall survival however not every patient responds to checkpoint immune blockade aimed at restoring T cell function. Determining if a patient has T cells that respond to immunotherapies, their characteristics and how they can be manipulated to target cancer cells is an area of intense investigation in cancer therapy. The objective of this study was to determine the cellular composition and their transcriptional state in metastatic ovarian cancer samples using single cell RNA sequencing (scRNA-seq). Hierarchical clustering stratified our patient cohort into 2 main groups: a high T cell and 2) a high monocyte signature group suggesting differences in immune response. To assess the immune response in these patient samples, we performed an unsupervised clustering of the T cell population in the two groups. The T cell population clustered into 4 and 3 subpopulations in the high T cell and high monocyte group respectively. A granulysin expressing T cell population was unique to the High T cell group. Interestingly, although both groups had resident memory CD8⁺T (CD8⁺Trm) cells, only the CD8⁺Trm cells in the high T cell group expressed TOX, a recently described transcription factor. TOX confers longevity to T cells within immunosuppressive environment such as cancer. Along with TOX⁺ T cells we found unique B cells and other novel macrophage clusters in our high T cell response group. In conclusion, our scRNA-seq data reveal novel subsets of immune cells that may contribute to the success of immunotherapy in ovarian cancer patients.
Introduction

Ovarian cancer is the most lethal malignancy of the female reproductive tract\(^1\). Conventional therapy involving cytoreductive surgery and chemotherapy are 90% effective when cancer is diagnosed at the early stage when it is still restricted to one or both ovaries. Unfortunately, most ovarian cancer cases are diagnosed at stage III when the cancer has metastasized and the diagnosis in these patients result in a 30% 5-year survival rate\(^2\). Additionally, the efficacy of novel therapies such as immunotherapy is not consistent among cancer patients. We therefore need biomarkers for early clinical detection and that are predictive of patient response to novel therapies. Furthermore, to develop efficacious therapies for metastatic ovarian cancer, we need to define the cellular heterogeneity and the transcriptional state within the tumor microenvironment. Immunohistochemical staining and flow cytometry have been useful in categorizing the cell types based on specific cell surface markers. However, cell classification solely on the basis of phenotypic markers masks the heterogeneity within individual cell types. Bulk profiling studies average gene expression and fail to identify the respective contribution of cell subsets. Alternatively, single cell RNA sequencing (scRNA-seq) has emerged a powerful tool to interrogate tumor composition revealing previously uncharacterized cellular heterogeneity as well as revealing gene regulatory networks within the tumor microenvironment\(^3-8\).

Although epithelial ovarian is a highly heterogenous disease, all histotypes commonly metastasize to the omentum. The omentum is mainly an adipose tissue in the peritoneal cavity containing aggregates of immune cells in areas called milky spots. These milky spots act similar to lymph nodes, collecting and responding to antigen within the peritoneal cavity. Interestingly, ovarian cancer cells preferentially colonize adipose tissue with milky spots in the peritoneal cavity\(^9\). Additionally, adipocytes provide adipokines and act as a source of energy for ovarian cancer cells\(^10\). It is likely that integrated signals from adipocytes and immune cells make the omentum a focal point of metastasis of ovarian cancer. In high grade serous ovarian carcinoma, the most common histotype, the presence of tumor infiltrating lymphocytes is associated with significantly longer overall survival suggesting that not all immune cells encourage metastasis of ovarian cancer\(^11\). In this current study, we use scRNA-seq technology to examine the cells within omental tumors from 6 patients with different histotypes of ovarian cancer. We identified 12 cell clusters including epithelial and stromal populations. Patients were divided into two groups based on immune signatures; 1) T cell response 2) monocyte response group. This unprecedented concurrent transcriptomic analysis of all cells within metastatic ovarian cancer unravels patient specific response to cancer and highlights patients suitable for cancer immunotherapy.

Materials and methods

Tissue collection
Ovarian cancer tissue was collected from women undergoing debulking surgery at the University of Chicago. Human tissue acquisition after patient deidentification was approved by the university of Chicago institutional review board for human research. Ovarian cancer tissue was histologically carcinoma classified and staged by a pathologist according to tumor-node-metastasis (TMN) and/or international federation of gynecology and obstetrics (FIGO) classifications (Table 1). Three of the ovarian cancer samples originated from the left fallopian tube (with associated serous tubal intraepithelial carcinoma (STIC)).

Immunohistochemistry

Tissue sections were fixed overnight in 4% formaldehyde at 4°C. After serial dehydration, the tissues were embedded in paraffin and cut into 5µm thick sections. Histological evaluation was done with hematoxylin and eosin (H&E). Immunohistochemical staining was performed to confirm the presence of cytokeratin-7 (Thermofisher), pan-vimentin (DAKO), CD45 (Agilent Technologies), CD3 (Human envision) and CD19 cells. Briefly, sections were deparaffinized and rehydrated through xylene and serial dilutions of EtOH to distilled water. Tissue sections were incubated in citrate buffer, pH 6 and heated in a microwave oven. Anti-cytokeratin-7(1:1000), anti-Vimentin (1:100), anti-CD45 (1:100), anti-CD3 (1:100), and anti-CD19 (1:200) antibodies were applied on tissue sections along with one-hour incubation at room temperature in a humidity chamber. The antigen-antibody bindings were detected with labeled polymer-HRP Envision system (DAKO, K4007) and DAB+ chromogen (DAKO, K3468) system. Tissue sections were briefly immersed in hematoxylin for counterstaining and were covered with cover glasses.

Tissue digestion, Red blood cell lysis and Dead cell removal

Ovarian cancer tissue obtained from the anatomic pathology department was transported in DMEM/F12 containing 10% FBS and 1% P/S (10% DMEMF/12) on ice to the laboratory. Ovarian cancer tissue was minced manually with a scalpel and enzymatically digested using 1.5 mg/ml collagenase IV (Sigma-Aldrich), 1 mg/ml hyaluronidase (Sigma-Aldrich) and 500 µg/ml DNase I (GoldBio) in Hank’s balanced salt solution (HBSS) in a 37°C shaker (200rpm) for 0.5 – 2 h. Following digestion, cells were resuspended in 10% DMEMF/12 and filtered serially through 70 µm and 40 µm strainers. Red blood cells were lysed by incubating the cell suspension in RBC lysis buffer (Sigma-Aldrich) for 2-5 minutes. Lysis was quenched by adding excess 10% DMEMF/12. The number of live cells was enriched using the dead cell removal kit (Miltenyi Biotec, 130-090-101) according to manufacturer’s instructions.

Drop-seq Experiments

Drop-seq experiment was performed as previously described12. Briefly, cells and barcoded beads were loaded at a concentration of 100,000 cells/ml in PBS-BSA and 120,000 beads/ml in lysis buffer in separate 3 ml syringes. Drop-seq experiments were run using a 125-micron microfluidic device at 16 ml/hr (oil), 4 ml/hr (cells) and 4 ml/hr (beads) with ~15 minutes per collection. Following collection, drops were broken and barcode beads with mRNA hybridized unto them were collected and washed. Barcoded cDNA attached to the barcode beads or STAMPS were
generated by reverse transcription, treated with exonuclease 1 and the number of STAMPs was counted. 5000 STAMPs were aliquoted per well in a 96-well plate and the cDNA attached to the STAMPs were amplified through 14 PCR cycles. PCR supernatants from each well were pooled and cleaned with Ampure beads. Purified cDNA product was quantified using Qubit 3.0 (Invitrogen) and 450-650 pg of each sample was used as input for Illumina Nextera reactions (12 cycles). Tagmented libraries were quantified using an Agilent BioAnalyzer High sensitivity chip according to manufacturer’s recommendations before submission for sequencing on Illumina’s Nextseq 500, using NextSeq 75 cycle v3 kits. Paired end sequencing was performed on the samples with 5% Illumina PhiX Control v3 added, with 20 bp for Read 1 and 64 bp for Read2 using custom Read1 primer, GCCTGTCCGCGGAAGCAGTGGTATCAACGCAGAGTAC, according to manufacturer’s instructions.

From alignment of the reads collected from the Dropseq experiments, read count matrices for unique molecular identifiers (UMIs) were generated for both exon regions and intronic regions in the human genome (gencode hg38 v.27) using the STAR version 2.5.3 aligner. The ovarian cancer samples from patients’ metastatic site omentum were sequenced with Drop-seq. A total of 13 sequencing runs were performed for six patients metastatic site samples using Drop-seq where each sample was sequenced twice to maximize the sequencing depth and 3401 sample was sequenced three times. Each sequencing run produced paired-end reads, with one pair representing the 12 bp cell barcode and six bp unique molecular identifier (UMI), and the second pair representing a 60 bp mRNA fragment. The paired-end reads from the same samples were merged and six paired-end fastq files were generated. A snakemake protocol based pipeline was applied to process the paired-end reads from each sequencing run and produce a count matrix representing the expression of every gene in every cell (Supplementary Methods). Individual count matrices were produced for each of the six patients after accounting for UMI duplicates.

The individual count matrix was produced for each patient with a total number of six count matrices. The summarized counts for each gene were inferred based on both exon and intronic reads to produce the gene expression matrix per sample. Expression matrices of six ovarian cancer samples from the metastatic sites in the omentum were generated. To select high quality cells, we applied a filter based on the number of genes detected per cell. Prior to filtering, each sample produced approximately 5000 cells. Based on the median number of genes captured, cells with less than 800 genes were removed from the data sets. For lower quality samples, the threshold was lowered to 600 to 700 as shown in Table 2. We followed a standardized pipeline implemented in single cell analysis tool suite Seurat v3.0.2. A global-scaling normalization method that normalizes the feature expression measurements for each cell by the total expression, multiplies this by a scale factor of log (TP10k) was applied to all samples. Each normalization matrix was then scaled by a linear transformation to center the mean gene expression of cells and normalized the gene expression variance. We applied PCA on the normalized expression matrix to extract the variable components in the sample data, followed by a heuristic elbow plot on the variance of each PC. We selected the number of top variant PCs based on the elbow plot and the number varies from 10 to 15 depending on the sample. The top
PCs were used in further exploration of the data, such as UMAP /tSNE dimension reduction, construction of K-nearest neighbor graph and other graphical clustering analysis.

**Cell type classification (correlation with CellAtlas)**

To assign the cell type to individual cells, we used a bulk sequencing gene expression data-set from 95 cell lines collected by Cell Atlas\(^{18}\) which cover 33 majority of cell types in normal human tissue, including common immune cells, endothelial, epithelial, fibroblast, mesodermal cells and so on. The cell lines can be further divided in to fine-grained 60 subtypes. We inferred the similarity between individual cells in the samples and cell lines by calculating the pairwise Pearson correlation matrix \(C=\{c(i,j)\}\) among any cell, \(i\), from the Drop-seq experiments with any cell line, \(j\), in the cell atlas. We performed a shared nearest neighbor (SNN) modularity optimization-based clustering on the CellAtlas cell lines to group the cell lines into larger and more general cell type categories. This was done to eliminate the bias in CellAtlas dataset and remove some rare or under-powered cell types. The resulting clusters of cell lines were annotated by the most frequent major cell type from the cell line groups. For each cell in our cancer sample, the cell type was then assigned with the most common cell type from top 5 highly correlated major cell type clusters. We collapsed few ambiguous cell types with the neighboring cell types based on their expression profiles. The cell sub-type was assigned with the most highly correlated cell-type from the 60 sub-types in the cell lines.

**Classification with cluster markers, canonical genes, and genetic functions**

Sometimes the Gene Ontology (GO) and Pathway analysis gives general/noisy functional categories which make certain cell types hard to be identified. By leveraging prior knowledge from the CellAtlas mapping, we were able to locate the relevant function categories for those cell types and narrow down their marker genes more efficiently\(^{18}\). Moreover, by the consensus of both GO and CellAtlas analyses, we obtained higher confidence in classifying cell types in each patient sample.

We used the cell types obtained from the cell Atlas correlation as baseline and curated the cell types manually using canonical genes information and functional association for cell types detected. The differentially expressed gene markers were extracted from cell clusters. Cross-referencing was done between gene markers and known canonical genes to the cell types. Gene ontology and pathway-based enrichment was performed on the gene markers to provide additional evidence for the cell type assignment. As a result, we identified 9 major cell types which include epithelial cells, fibroblast cells, three immune cell types, viz., B cells/B cell plasma, T cells, and monocytes, mesenchymal stem cells (MSC), embryonic stem cells (ESC), and endothelial cells.

**Trajectory analysis**

We selected cells based on the classified cell types and performed trajectory analysis using monocle for single cell type and monocle 3 for multi-cell types\(^{19,20}\). The trajectory analysis in
monocle was to reduce the high dimensional gene expression matrix onto a tree structure using DDR tree\textsuperscript{20}. This method produces a single tree with multiple branches. The monocle\textsuperscript{3} was based on UMAP dimension reduction and the resulting trajectory can be disconnected so that the trajectory path for various cell types can be shown simultaneously\textsuperscript{21}.

Results

Patient cohort

Ovarian cancer samples were collected from the metastatic site (the omentum) of six patients (Table 1). Four patients were diagnosed with high grade serous cancer, one with serous carcinoma and one with malignant mixed Mullerian tumor (MMMT). The ages of the patients ranged from 46 to 71 years. All tumors were diagnosed at stage III or later. Four of the patients had received neo-adjuvant therapy prior to surgery including the patient diagnosed with MMMT. The reported ethnic background of patients includes two Whites, two Blacks and two Asians.

Characteristics of the Tumor microenvironment

Normal omentum is comprised mainly of adipocytes as well as aggregates of leukocytes called milky spots\textsuperscript{22}. Adipocytes act as a source of energy for ovarian cancer during invasion of the omentum\textsuperscript{10}. This results in an inverse relationship between ovarian tumor growth (cancer and cancer associated cell growth) and reduction of adipocytes within the omentum\textsuperscript{10}. The area of tumor occupied by adipocyte decreases with disease score\textsuperscript{23}. To quantify the transformation of the omentum by ovarian cancer cells, we used Imagescope, a digital histopathology software, to annotate and calculate the area occupied by adipocyte in each patient sample on hematoxylin and eosin (H&E) stained sections (Figure 1A). The area of adipocyte was reported as a percentage of the total surface area of the sample used as a measure of disease score as previously described by Pearce et al.\textsuperscript{23} which we also used (Figure 1B). Two samples, PT-3401 and PT-5150 had the highest percent area of adipocyte which we categorized as low disease score while the other 4 samples are categorized as high disease score. The presence of cancer cells, total stromal cells and immune cells were confirmed by immunohistochemistry (IHC) using antibodies against cytokeratin-7 (CK-7), vimentin and CD45 antibodies (Figure 1C-H). There was a varying proportion of each cell type across patients. Notably, we observed that malignant cells in some patients (PT-5150, PT-3232, PT-4806 and PT-2834) double stained for CK-7 and vimentin suggesting active epithelial mesenchymal transition (Figure 1C-F). In addition, the MMMT (PT-2834) sample had two different types of malignant cells that stained separately for CK7 and vimentin. Incidentally, the proportion of vimentin staining cells made up by CD45\textsuperscript{+} immune cells positively correlated with the area of adipocytes (Figure 1B&H). We also observed that these immune cells were enriched around the remaining adipocytes and sparsely throughout the rest of the tissue. Together, these data reveal patient variability in cancer cells in different histological types of ovarian cancer and positively correlates the presence of adipocytes and immune cells.
Generation of single cell Data, Cell-type clustering and cell type assignment

For Drop-seq experiments, tumors samples were enzymatically digested into a single cell suspension and ran through drop-seq experiment as previously described. We integrated all six omentum samples resulting in a gene expression matrix which contains the expression values of 9885 cells and 40947 genes mapped from the human reference genome (version GRCh38). A minimum cut off of 600-800 genes per cell was set as shown in Table 2. The clustering was performed by hierarchical clustering using resolution of 0.2 where 12 clusters are detected (Figure 2A). The cells from each sample have membership across all clusters (Figure 2B). For each cluster, a one-against-all upregulated markers detection was performed using Wilcoxon Rank Sum test and genes with adjusted P-value < 0.05 were identified as significantly upregulated gene markers. The dimension reduction was done using both t-SNE and UMAP and was found to be consistent. The cell types were assigned and curated by cell line correlation, canonical genes and functional categories for the significantly differentially express genes on the detected clusters. 9 major clusters out of 12 were identified with cell types (Figure 2C). This included 1 cluster for malignant cells, 1 cluster for fibroblast, 1 cluster for mesenchymal stem cells, 1 cluster for embryonic stem cells, 1 cluster for endothelial cells. Unsupervised clustering of the total cell population of the different patients revealed three groups (Figure 2D); the first group (PT-6885 and PT-4806) had high ratios of epithelial and fibroblast clusters. The second group had a high epithelial and T cell ratio (PT-3401 and PT-5150) and the final cluster was driven by a high epithelial cell ratio (PT-3233 and PT-2834). Neo-adjuvant therapy has been previously described to shift the cellular landscape from epithelial to immune cells in high grade serous ovarian cancer samples. Although we had 4 neo-adjuvant treated patient samples, only two (PT-6885 and PT-4806) clustered together in the epithelial/fibroblast group. The neoadjuvant treated sample in the epithelial group was from the MMMT patient with no significant shift to immune cells. Finally, the neo-adjuvant sample in the high T cell group had the highest percentage of immune cells. These scRNA-seq data reconfirms the heterogeneous nature of the tumor environment as well as the heterogenous response to therapy.

Cancer cell clustering analysis

All 6 patient samples were merged and cells belonging to the central clusters connected to epithelial cells shown in Figure 3 were selected. The corresponding cell types are epithelial cells and ESC according to the cell types assignment results from all cellsAtlas. There was a total of 4,733 cells in this subset. We detected 7 clusters using resolution 0.2 (Figure 3A&B). Cluster 0 and 3 highly expressed EpCAM (Figure 3E) that has previously been associated with platinum resistance and metastasis in ovarian cancer. Cluster 0 expressed the highest levels of IL-6 and TGF-β (Figure 3C&D). IL-6 enhances tumor cell survival and platinum resistance and correlates with poor prognosis in ovarian cancer while TGF-β induces an aggressive cancer phenotype. In addition to EpCAM, Cluster 3 highly expressed NDRG1 and other epithelial markers such as KRT18, KRT19 and CD24 (Figure 3E). NDRG1 is activated by p53 and was shown to inhibit proliferation of metastatic lung cancer cells. Another ovarian cancer marker, PAX8, was expressed by cluster 2, along with high expression of MALAT1 and MUC16 (Figure 3E). High expression of MALAT1 is associated with advanced disease, recurrence and reduced survival.
MUC16 is a biomarker of ovarian cancer and necessary for tumor cell growth and metastasis.\textsuperscript{33,34} Cluster 1 expressed vimentin along with high MKI67 and PCLAF indicative of a proliferating population. Cluster 5 also expressed PCLAF, a marker of proliferation, as well as FN1 and ZEB1 that are markers of epithelial to mesenchymal transition (EMT). Cluster 6 highly expressed vimentin, MKI67 (marker of proliferation), S100A4, a protein implicated in tumor and progression and metastasis(Figure 3E).\textsuperscript{35} These highly expressed markers, along with comparatively lower expression of EMT markers (ZEB1 and ZEB2) in cluster 6 are suggestive of cells undergoing mesenchymal to epithelial transition (Figure 3E). We also investigated the presence of clusters expressing classical ovarian cancer stem cells markers including CD33, CD44, CD117 and CD24.\textsuperscript{36} However, we did not observe cells co-expressing these markers, which may be due to the problem of low sensitivity and drop-outs in our single-cell data. These distinctive clusters reveal the heterogeneity within ovarian cancer population at any given time and the multiple processes that cancer cells harness to progress disease.

**Immune cellular profile of patient samples**

The milky spots of the omentum is normally composed of 70% macrophages, 10% T cells and 10% B cells along with mast cells and stromal cells.\textsuperscript{22} During the initial phase of metastasis and colonization of the omentum, additional macrophages are recruited from the peritoneal cavity without any anti-tumor effect on the cancer cells.\textsuperscript{37,38} The presence of T cells positively correlates with overall survival.\textsuperscript{39,40} To further investigate the immune population in the metastatic omentum tissue, we selected the immune cells from the six patient samples and divided the cells into 3 major categories T cell (956), B cell (409)/B cell plasma (358) and monocytes (695). The T cells can be further divided into CD4 and CD8 positive clusters based on the violin plot of the CD4 and CD8 genes. B cells and plasma cells were distinguished by the B cell specific marker gene CD19 where the expression of CD19 was turned down in plasma cells. The monocytes exhibited a combination of macrophages and monocytes. Similar samples were grouped by a dendrogram resulting in 2 main groups (Figure 4A): 1) a T cell signature group (2 samples) and 2) a monocyte signature group (4 samples). These groups are similar to our previously described low and high disease score (Figure 1). In addition, the high T cell group also expressed a high B cell signature which was absent in the samples with the monocyte/macrophage signature group. Interestingly, there was a difference in the samples of the high T cell group between the neo-adjuvant (PT-5150) and the non-neoadjuvant treated sample (PT-3401). The neoadjuvant sample had a higher ratio of plasma B cells (64.4%) than the sample from the non-neoadjuvant sample (28.0%) (Figure 4B). Conversely the neoadjuvant sample had a lower ratio of B cell cluster 1 (naïve/memory B cells) (35.6%) compared to the non-neoadjuvant sample (72.0%). CD20\textsuperscript{+} B cells colocalized with CD8\textsuperscript{+} T cells positively correlated with overall survival\textsuperscript{41}. Contrastingly, plasma cell infiltration in epithelial ovarian cancer significantly reduced overall survival\textsuperscript{42}. Taken together this data suggests there is varied immune response to neoadjuvant therapy amongst sample/patients and a need for personalized novel therapies.

**Differences in T cell clustering and subtype analysis**
We classified the 6 omentum samples into two disease groups based on the adipocyte ratio detected in the samples as well as the dendrogram grouping. Higher disease score (HighDS) corresponds to lower adipocyte and T cell ratio while the low disease score (LowDS) group had higher adipocyte and T cell ratio. The HighDS group (samples PT- 4806, PT-6885, PT-2834, PT- 3233) and LowDS group yielded 5190 and 4695 total cells respectively. To reveal the functional subtypes of T cells and any difference in the HighDS versus LowDS groups, we clustered a total of 956 total T cell, 820 LowDS T cells and 136 HighDS T cells (Figure 5A). A total of 4, 4 and 3 transcriptionally distinct clusters emerged from total, LowDS and HighDS T cells respectively. (Figure 5A-C). The top genes revealed clusters similar to previously described T cell phenotypes in breast cancer including CD4*IL7R, CD4*FOXP3, resident memory CD8*T cells (CD8*Trm cells) and one population described in lung cancer with high granulysin expression CD4*GNLY. Although there were fewer T cells in the HighDS group they clustered into 3 groups. Both groups had LowDS group. CD8*Trm cells had the highest expression of IFN-γ in both groups (Figure 5B). Additionally, although both groups had CD8*Trm cells, only the CD8*Trm cells (P value < 0.004) in the LowDS group expressed a recently described transcription factor, TOX, which is the master regulator of exhausted T cells (Figure 5C). TOX was also lowly expressed by the GNLY*T cells (p value = 0.407). These TOX+ T cells persist during chronic infection and TOX is expressed on the CD8*T cells that are reactivated in response to PD-L1 immunotherapy. These data suggest that scRNA-seq is capable of determining patients that will benefit from novel therapy.

Transcriptionally distinct plasma cell clusters in LowDS group.

B cells studies in ovarian cancer has been limited to their effect on survival. CD20+ B cells colocalized with CD8+ T cells positively correlated with overall survival. Contrastingly, plasma cell infiltration in epithelial ovarian cancer significantly reduced overall survival. These surface markers alone are not informative about the transcriptional state or function of B cells. To investigate the B cell subtypes in our samples, B cells and plasma cells were distinguished by the B cell specific marker gene CD19 where the expression of CD19 was turned down in plasma cells. In the LowDS group there were a total of 396 B cells and 274 plasma cells while there was a total of 124 B cells detected in the HighDS group. Unsupervised clustering on combined B cells revealed a total of 3 clusters (Figure 6A); naive/effector B cells, plasma B cells and a third cluster enriched with cycling genes. When we performed unsupervised clustering on the two groups separately, the LowDS and HighDS group had 4 and 2 clusters respectively (Figure 6B&C). In the LowDS group, cluster 0 had high expression of MHC class II genes, cluster 1 and 2 were both positive for BLIMP1 indicating they are plasma cells. However, cluster 2 highly express CD38 and CD138 indicative of post germinal center B cells. Both cluster 1 and 2, highly expressed class switched IGHG. Finally, cluster 3, highly expressed proliferative genes and low Blimp1 suggesting a plasmablast population. In the HighDS group, cluster 0 highly expressed CD24 a naive B cell marker while cluster 1 highly expressed BLIMP1. Post germinal center B cells were absent in this group.

Identification of novel macrophage subsets
The samples assigned to the HighDS group had a low T cell to monocyte ratio (0.4: 1) compared to the LowDS group (2.5:1). To assess the differences in the macrophage population, we performed unsupervised clustering on both groups combined (Figure 7A) and separately (Figure 7B&C). There were 4 clusters in the LowDS group and 2 clusters in the highDS. Both groups had a CD14+CD16+CD163+ cluster reminiscent of tumor associated macrophages. In the highDS group, there was also an immature CCR7+FLT3+ population. These FLT3+ progenitors can differentiate into osteoclasts, dendritic cells, microglia and macrophages. The other clusters in the high T cell group included CCR7+CD274+ which suggest a regulatory population likely similar to MDSCs. IRF8+ expressing macrophages (Cluster 2) are induced in the presence of IFN-γ and promote the formation of autophagosome. The final cluster highly expressed NR1H2 inhibits inflammatory genes in macrophages.

Discussion

In this study we analyzed approximately 10,000 cells from 6 patients with 3 histotypes of epithelial ovarian cancer at the most common metastatic site, the omentum. The resolution of our dataset revealed 9 major clusters of which we further interrogated the cancer and immune cells clusters. The concurrent transcriptomic analysis of cancer and stromal cells provides insights into the interactions occurring within the tumor microenvironment that contribute to the progression of cancer. Our single cell data was able to stratify our patients into two groups based on their immune response.

The largest cluster of cells was cancer epithelial cells, which composed about 50% of cells analyzed. These cells expressed a variety of genes associated with metastatic disease including MUC16, PAX8 and MALAT1. Although we combined cells from 6 different patients, cancer cells from each patient expressed a variety of these markers indicating that targeting any of these genes will be redundant since there are compensatory proteins that will continue to progress disease. Another approach to targeting tumors will be the identification of cancer stem cell in ovarian cancer. Many studies have attempted to identify these rare cells without success. We did find a cell population closely resembling embryonic stem cells that highly express proliferative markers. Successful identification for isolation and interrogation of these putative ESC cells might provide useful insight about CSC cells in ovarian cancer.

Infiltration of immune cells into the omentum during metastatic cancer does not always elicit anti-tumor responses. In ovarian cancer, high tumor infiltrating T cells is associated with significantly longer overall survival. However not all patients have a high infiltration of T cells. Additionally, the transformation of adipocytes into cancer cell described here inversely correlates with the number of infiltrating immune cells. The mechanisms behind this difference in immune response remains incompletely understood. There is evidence that associates immunogenicity of neo-antigens generated due to high tumor mutational load and clinical response to immunotherapies. The functional plasticity of leukocytes makes them an attractive therapeutic target. Additionally, immunotherapy avoids the obstacle of drug delivery in solid
tumors by using the cells within the tumor environment to target therapy, making cancer immunotherapy an area of high investigation. To this end, we sought to investigate the immune population within these metastatic tumors. From the 6 patients we investigated in this study, 2 of them had a high T cell response while the other 4 had a high monocyte response. Once we grouped our data into two different immune groups we sought to reconcile the transcriptomic signature in the cancer cells with the immune response. Both high T cell and high monocyte groups had 8 cancer cell cluster. Was there any difference in these cells?

Advances in personalized medicine has revealed TOX as a transcription factor expressed by T cells that respond to immune checkpoint block. In our patient cohort, TOX was expressed by two T cell subsets, CD8^+Trm and CD4^+GNLY in the high T cell group. The high monocyte group also had the CD8^+Trm subset but not the CD4^+GNLY subset. Interestingly, there was no TOX expression in the high monocyte group. TOX confers longevity and persistence on T cell in immunosuppressive environment such as cancer; however, the mechanism by which these TOX^+ T cells are generated is not understood. Coincidentally, there is also a difference in the B cell population. Along with the expression of TOX in T cells, the high T cell response group also had a unique subset of BLIMP1^+CD38^+CD138^+ B cell population. It is likely that these antigen cognate plasma B cells are involved in the generation or maintenance of TOX^+ T cells. A mechanistic study of these cells is necessary as they may hold the key into inducing TOX^+ T cells in patients who do not respond to immune checkpoint inhibitors.

Conclusion

To our knowledge, this is the first scRNA-seq study that allows stratification of patients according to immune response in metastatic ovarian cancer. Concurrent transcriptomic analysis of cancer and stromal cells revealed intratumor heterogeneity and patient variability highlighting the need for personalized medicine. Interrogating tumor infiltrating lymphocytes at the single cell level also feature unique differences in other tumor infiltrating immune cells that are previously undescribed. Mechanistic studies are required to determine the role of these cells in tumor progression and response to novel cancer immunotherapies.

Acknowledgement

We are grateful to Mengjie Chen for helpful discussions on CellAtlas data and Mark Lingen for help with tissue access. This work was performed, in part, at the Center for Nanoscale Materials, a U.S. Department of Energy Office of Science User Facility, and supported by the U.S. Department of Energy, Office of Science, under Contract No. DE-AC02-06CH11357. All patient tissues were histologically graded, de-identified and obtained from the University of Chicago Human Tissue Resource Center. Computational resources and data storage were provided by the University of Chicago Research Computing Center.
References


Figure 1: A) The hematoxylin and eosin (H&E) stained sections with B) the percentage area of adipocytes histogram on the individual patient sample. C) Cytokeratin-7 (CK-7) staining on patient sample slices with D) the percentage of CK-7 positive histogram on the individual patient sample. E) Vimentin staining on patient sample slices with F) the percentage of Vimentin positive histogram on the individual patient sample. G) CD45 staining on patient sample slices with H) the percentage of CD45 positive histogram on the individual patient sample. Images taken at A) 40x; C),E),G) 160x.
Figure 2: A) UMAP plot for high-quality cells, based on number of genes detected per cell in all six omentum samples, colored by clustering results. B) The UMAP colored by the patients. C) UMAP plot for clusters colored and annotated by assigned cell types. D) Heatmap of cell types composition ratio in each patient sample with hierarchical dendrograms on both patients and cell types.
Figure 3: A) UMAP plot for annotated cancer cell-types only from all six omentum samples and colored by patient. B) UMAP plot colored by unsupervised clustering results. C) Violin plot for gene IL-6 expression level across all seven clusters. D) Violin plot for gene TGFB1 counts across all seven clusters. E) Feature plots of relevant marker genes.
Figure 4: A) Heatmap of immune cell types (T cell, B cell, plasma B cell, and Monocyte in each patient sample with vertical and horizontal dendrograms on both patients and cell types. B) The overall percentage of two B cell populations (B cell 1: naïve B cell; B cell 2: plasma B cell) in low disease score group, PT-3401 and PT-5150.
Figure 5: A) UMAP plot for T cells in across all six samples, annotated by T cell sub-population for total, low and high disease score. B) The violin plot of gene IFNG expression level across four T cell sub-populations in all six samples. C) The feature plot for expression level of gene TOX in low disease score samples.
Figure 6. UMAP plot for B cells in all six samples, annotated by B cell sub-population A) total B) low disease score group C) high disease score group along with feature plots for marker genes.
Figure 7. UMAP plot for monocytes in all six samples, annotated by monocyte sub-population in A) total B) low disease score group C) high disease score group along with feature plots for marker genes.
<table>
<thead>
<tr>
<th>Omentum</th>
<th>Age</th>
<th>Race</th>
<th>Origin</th>
<th>Histologic type</th>
<th>Histological grade</th>
<th>Neoadjuvant therapy</th>
<th>Stage (PMN/FIGO)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PT:3232</td>
<td>46</td>
<td>Asian</td>
<td>Left fallopian (STIC)</td>
<td>High grade serous carcinoma</td>
<td>High grade</td>
<td>No</td>
<td>pT3c Nx / IIIC</td>
</tr>
<tr>
<td>PT:3401</td>
<td>56</td>
<td>White</td>
<td>Left ovary</td>
<td>High grade serous carcinoma</td>
<td>High grade</td>
<td>No</td>
<td>pT3c Nx Mx / IIIC</td>
</tr>
<tr>
<td>PT:6885</td>
<td>66</td>
<td>Black</td>
<td>Left fallopian (STIC)</td>
<td>High grade serous carcinoma</td>
<td>High grade</td>
<td>Yes</td>
<td>ypT3c N1a / yIIIC</td>
</tr>
<tr>
<td>PT:5150</td>
<td>62</td>
<td>White</td>
<td>undetermined</td>
<td>Serous</td>
<td>Not applicable</td>
<td>Yes</td>
<td>ypT3a Nx M1/ Ib</td>
</tr>
<tr>
<td>PT:2834</td>
<td>66</td>
<td>Asian</td>
<td>Fallopian</td>
<td>Malignant mixed Mullerian tumor</td>
<td>High grade</td>
<td>Yes</td>
<td>ypT3c Nx/ yIIIC</td>
</tr>
<tr>
<td>PT:4806</td>
<td>71</td>
<td>Black</td>
<td>Left fallopian (STIC)</td>
<td>High grade serous carcinoma</td>
<td>High grade</td>
<td>Yes</td>
<td>pT3c, N1, M1/ IIIC</td>
</tr>
</tbody>
</table>

Table 1. De-identified patient meta-data for all metastatic ovarian cancer tissues collected from the omentum.
Table 2. Filter parameters applied to samples collected from six patients and the resulting number of cells and genes per sample.

<table>
<thead>
<tr>
<th>Gene Cutoff</th>
<th>No of Genes</th>
<th>No of Cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>PT-5150</td>
<td>600</td>
<td>24776</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1244</td>
</tr>
<tr>
<td>PT-3401</td>
<td>800</td>
<td>32024</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3451</td>
</tr>
<tr>
<td>PT-6885</td>
<td>800</td>
<td>24063</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1071</td>
</tr>
<tr>
<td>PT-3232</td>
<td>700</td>
<td>19297</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1102</td>
</tr>
<tr>
<td>PT-4806</td>
<td>700</td>
<td>26499</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1108</td>
</tr>
<tr>
<td>PT-2834</td>
<td>800</td>
<td>26221</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1909</td>
</tr>
</tbody>
</table>