SUPPLEMENTAL MATERIALS

Systems-Level Mapping of the Tumor Microenvironment Reveals Immune-Mediated Mechanisms and Potential Targets in Platinum-Resistant Ovarian Cancer

Authors:

Adriana Del Pino Herrera[1] (adelpinoherrera@ufl.edu; ORCID: 0000-0001-5740-5779)
Miguel A. Martinez[1]
Monica Kim[2] (1007mkmk@gmail.com)
Kadin El-Bakkouri[1] (kelbakkouri@ufl.edu)
David A. Iglesias[3,4] (iglesiad@ufl.edu; ORCID: 0000-0002-8506-3451)
Meghan C. Ferrall-Fairbanks[1,4*] (mferrall.fairbanks@bme.ufl.edu; ORCID: 0000-0002-3380-6989)

AFFILIATIONS

[1] J. Crayton Pruitt Family Department of Biomedical Engineering, Herbert Wertheim College of Engineering, University of Florida, Gainesville, FL 32611
[2] Department of Computer & Information Science & Engineering, University of Florida, Gainesville, FL 32611
[3] Division of Gynecologic Oncology, Department of Obstetrics & Gynecology, College of Medicine, University of Florida, Gainesville, FL 32611
[4] University of Florida Health Cancer Institute, University of Florida, Gainesville, FL 32611

CORRESPONDENCE

* **To whom correspondence should be addressed.**

Meghan C. Ferrall-Fairbanks, PhD

Assistant Professor, University of Florida, J. Crayton Pruitt Family Department of Biomedical Engineering

P.O. Box 116131, Gainesville, Florida 32611; O: 352-846-2762, Email: mferrall.fairbanks@bme.ufl.edu

SUPPLEMENTAL METHODS

ImageJ Cellular Quantification

Briefly, images were input into ImageJ and the regions with the brightest pixels were set as true. The images were then converted to binary and segmented using different size threshold to identify different particles/cells. For the DAPI channel the size threshold was set to 5-2000 to identify all the different sized cells. For the FITC channel (used to label CD163+ and α-SMA+ cells) the size threshold was set to 5-1000 while for smaller cells in the TxRed and Cy5 channels (used to label CD20+ and CD8+ cells, respectively) size was set to 5-200 and 5-80 respectively. The total number of cells with positive marker expression per tile was collected and the proportion of the different cellular components was calculated based on the total number of cells present on that tile from the DAPI channel.

gentleMACS Dissociation

Briefly, fresh tumors were cut into 5-mm pieces and transferred to a gentleMACS C-tubes where RPMI1640 media and tumor dissociating enzymes were added. The tubes were then placed on the gentleMACS dissociator and the soft tumor dissociation program was run. The dissociated tumor was strained through a 70μM filter, centrifuged at 300 rcf for 10 minutes and cells were counted using a hemocytometer. Single cell suspensions were then frozen in 2 million cell vials in a solution of 10% DMSO-FBS. Vials were placed in a -80°C freezer for 24-48 hours and then moved to a liquid nitrogen dewar for long term storage.

RNA extraction

Briefly, frozen samples were cut to 30-60mg pieces and placed in a gentleMACS M-tube with 600 μL of buffer RLT and RNA program for frozen tissues was run. The five fresh samples dissociated into single cell suspensions were also prepared for RNA extraction. One vial of 2 million dissociated single cells was thawed in the water bath and resuspended in 5mL of warm RPMI. Samples were then centrifuged at 10,000 rpm for 5 minutes and resuspended in 350 μL of buffer RLT. The samples were vortexed for 30 second, sonicated for twice for 15 seconds and centrifuged at 14,000 rcf for 5 minutes to ensure complete cell lysis. The RNA was then extracted for both frozen and cell pellet samples following the centrifuging and washing steps in the Qiagen Kit.

Ovarian Cancer Cell Type Reference Processing

Briefly, the scRNAseq samples were merged into a reference object (n = 15,120 cells) was then filtered to only contain the cells that were present in the **Supplemental Table S4** (n = 13,646). SCTransform was used to normalize the reference and principal component analysis (PCA) was ran to make the reference usable for further applications. The cell type classification was added as metadata using the “cell.type” column from **Supplemental Table S4** which included: B cells, endothelial cells, epithelial cells, fibroblasts, macrophages, NK cells, T cells and an empty slot for other epithelial cells. After normalization the differentially expressed markers for each cell type were identified and the top 10 were plotted on **Supplemental Figure S3B**.

Once the reference was filtered, normalized and cell types were added to each cell as previously stated, the count matrix from the reference Seurat object was extracted. Counts were converted to counts per million and column names were changed from cell barcodes to cell types. Due to space constraints on CIBERSORTx, the count matrix was reduced to the top 3500 most variable genes and the same number of cells per cell type (192 cells per cell type since there were only 192 B cells in the reference). The finalized reference was then uploaded to CIBERSORTx and signature matrix file was created to be used as a reference.

Dissociated Cell Thawing

Initially, vials with 2 million cells per patient were retrieved from liquid nitrogen storage and thawed in a 37°C water bath until a small ice crystal remained. The cells were then transferred to 50mL conical tubes and the vial was rinsed with warm RPMI1640 media supplemented with 10% FBS. An additional 20mL of warm RPMI1640+10%FBS media was added to the conical tube slowly and the solution was then mixed by inverting the tube twice. The suspension was then centrifuged at 150 rcf for 10 minutes and resuspended 2 mL of RPMI1640+10%FBS.

Publicly Available Single-cell RNA-seq Handling

For the GSE154600 dataset, all the samples were downloaded by first downloading the bam files and converting them to compressed FASTQ files. This dataset was of particular interest since the samples were classified into sensitive or resistant. For both GSE181955 and GSE184880 datasets, compressed FASTQ files were extracted from SRA files. All samples were downloaded from the GSE184880 dataset but only five samples were downloaded from the GSE181955 dataset, since samples with depleted CD45 cells were excluded from the analysis.

Platinum-Sensitivity Classification Validation

While patients can be classified as sensitive and resistant, an entire tumor is not likely to be entirely composed of only sensitive or only resistant cells. To provide a more realistic classification, equal quantities of sensitive and resistant cells were down-sampled to build the reference and then could be used to validate sensitivity assignment. There were 42,747 resistant cells and 9,725 sensitive cells across the 5 patients with clinically annotated platinum-sensitivity and a random subset of 4,862 resistant cells and 4,862 sensitive cells were used to create a reference.

The remaining cells from those samples not included in the reference (37,885 resistant and 4,863 sensitive) were used as a query for singleR. Once a new classification of these cells was done, the query and reference were swapped resulting in a query that contained 4,862

resistant cells and 4,862 sensitive cells and a reference where resistant cells were down-sampled to contain 4,863 resistant cells and 4,863 sensitive. This new query and reference were once again run through singleR. This re-classification of cells allowed for samples that were initially formed by all sensitive or all resistant cells to now contain a percentage of both.

QuPath Cellular In-vitro Quantification

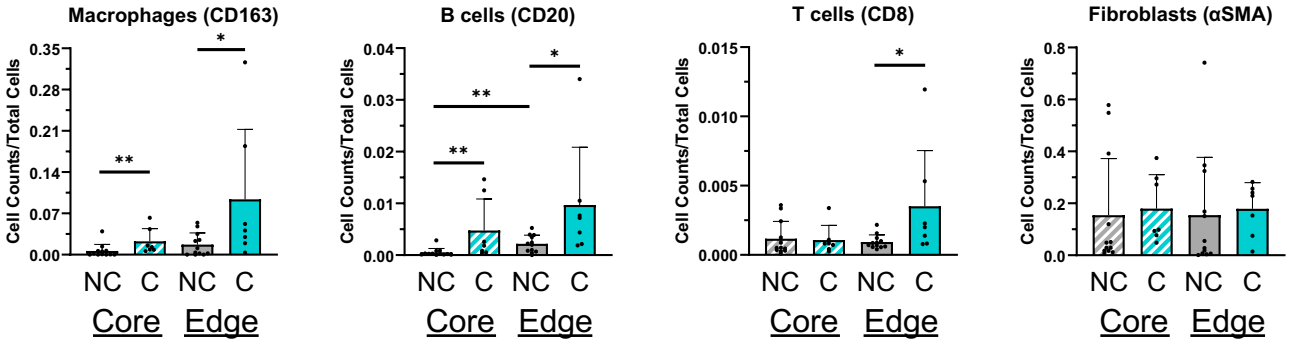
Six images for each well were taken at random using the ECHO Revolution with both TxRed and FITC channels. For both standard and M2 TAM conditioned media conditions, images were taken every day for 14 days. To collect cell counts from the images QuPath was used. Initially the size of 1 cell was measured using images from days 1-4. Then all the images from the TxRed and FITC channels separately were compiled and input to QuPath as a project. The total area covered by the cells was recovered using thresholding separating the background from the fluorescent cell signal. Once the area was collected, the total area was divided over the size of a single cell and cell counts were recorded. The counts for the 6 images taken per well were averaged and the number of cells in a well were estimated based on well and image size.

Bland-Altman Agreement Analysis

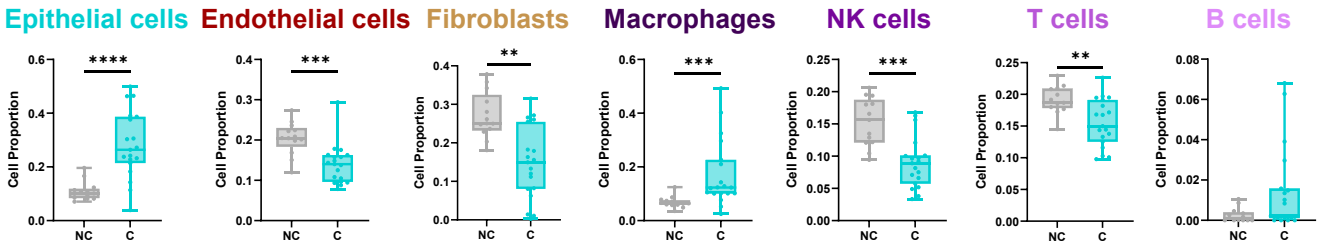
Seven non-cancerous samples and 6 cancerous samples were analyzed using both mIF and bulk RNA-seq technologies. Additionally, the four fresh cancer samples that we processed were sequenced using bulk and single-cell RNA-seq technologies. The agreement between these modalities in paired samples was studied using Bland-Altman plots. Non-cancerous and cancerous samples were combined when comparing the agreement between mIF and bulk RNA-seq technologies and plots were computed using the proportions of the markers of interest (CD163+ cells, CD20+ cells, CD8+ cells and α-SMA+ cells). This analysis calculates a bias between the modalities (solid line), which corresponds to the average of the differences between the cell type proportions. The differences between the cell type proportions across modalities was then plotted against the average between the cell type proportions across modalities. A standard

deviation of the bias was also computed and plotted along with 95% confidence intervals (dashed lines).

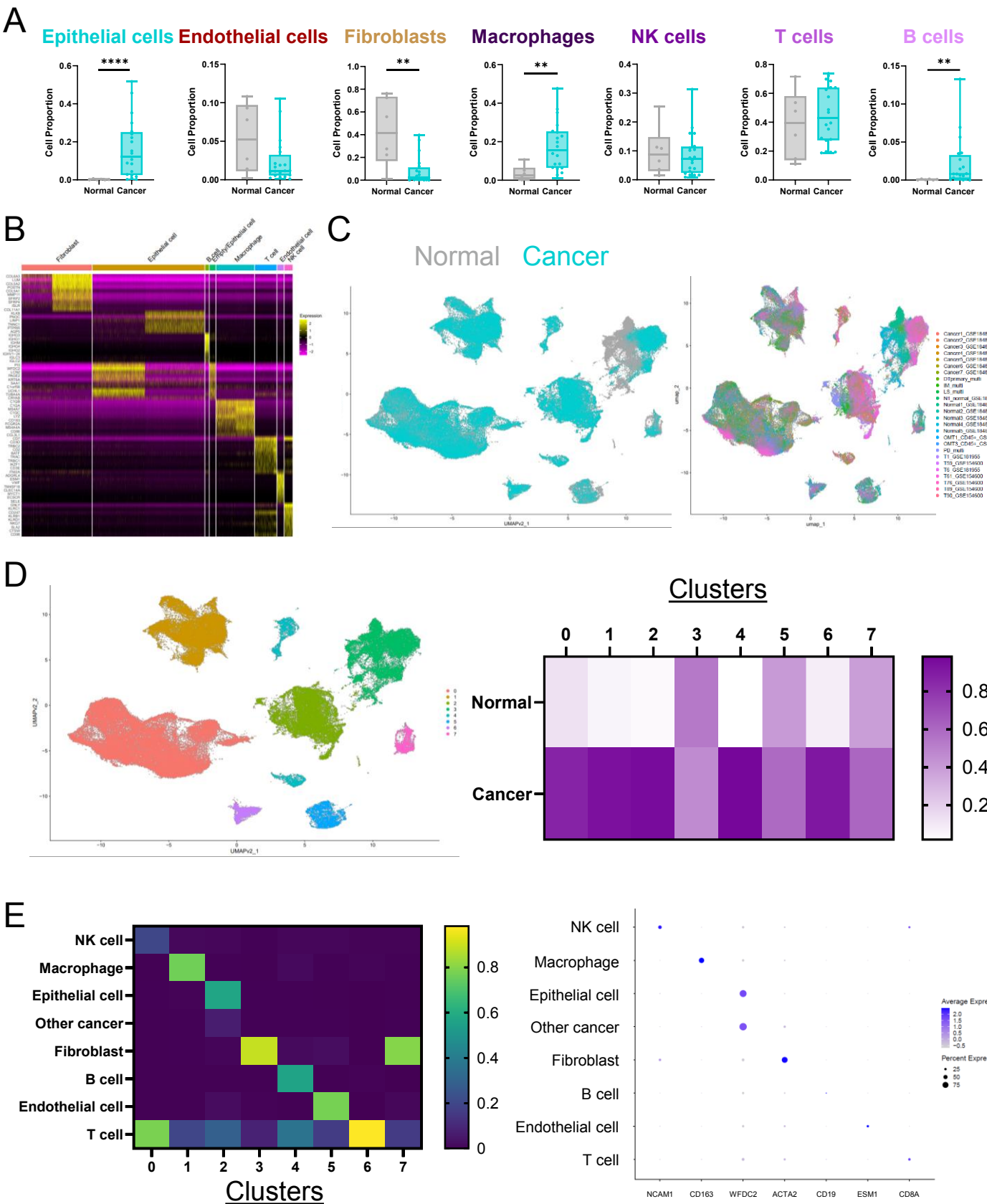
Both modalities/technologies agree when the bias is close to 0 and the points are randomly scattered around the bias and within the limits of agreement. If the bias is higher or lower than 0 then one modality is being favored over the other. If points are following a pattern, such as points are on an increasing straight line, one modality might be favored for small cell proportions while the other is favored for large cell proportions. Bulk RNA-seq was compared to mIF using the CIBERSORTx cell type proportions for bulk RNA-seq and the average core and edge cell type proportions for mIF. Single-cell RNA-seq was then compared to bulk RNA-seq using the singleR cell type proportions for single-cell RNA-seq and CIBERSORTx cell type proportions for bulk RNA-seq.



Supplementary Figure S1. Statistical analysis of mIF samples comparing core vs edge areas in non-cancerous vs cancerous samples across four different cell types. Generally, edge areas showed a higher cell infiltration compared to core areas and cancerous samples showed a higher proportion of macrophages, B-cells, and T-cells compared to non-cancerous.* p-value < 0.05; ** p-value < 0.01

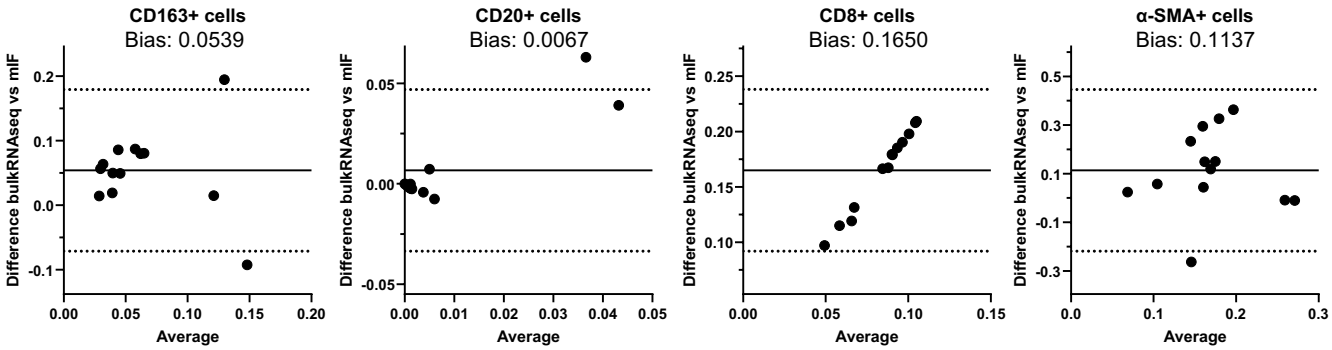


Supplementary Figure S2. Box and whiskers plots for each cell type proportions groups by non-cancerous and cancerous patients in the bulk RNA-sequencing cohort. Generally, non-cancerous patients showed higher proportions of endothelial cells, fibroblasts, NK and T cells while cancerous patient showed higher proportions of epithelial cells and macrophages. ** p-value < 0.01; *** p-value < 0.001; **** p-value < 0.0001

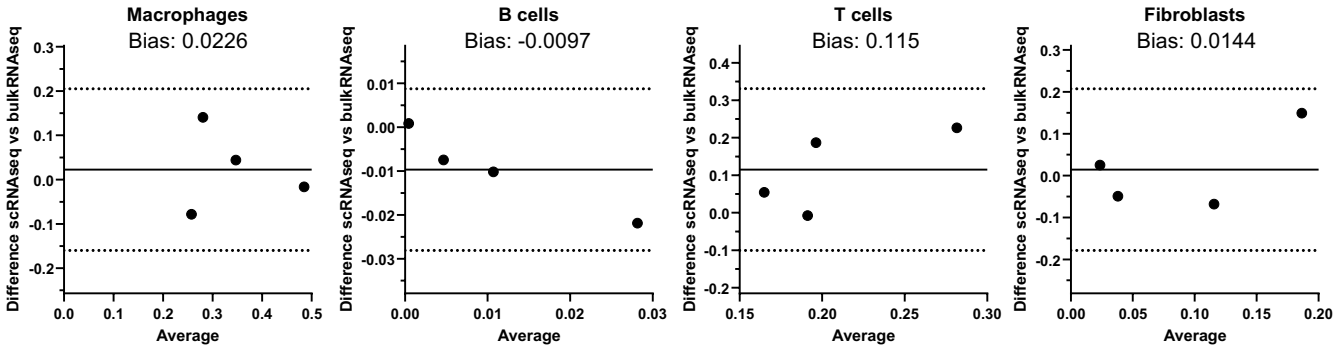


Supplementary Figure S3. A. Box and whiskers plots of cell type proportions grouped by normal and cancer samples in the scRNA-seq cohort. Generally, normal samples shower higher proportions of fibroblasts while cancer samples showed higher proportions of epithelial cells, macrophages and B cells. **B.** Heatmap of characteristic genes for each cell type in the ovarian cancer reference used to deconvolute cell type proportions in CIBERSORT for bulk RNA-seq samples and singleR for single-cell RNA-seq samples. **C.** UMAPs of the single-cell RNA-seq dataset grouped by normal and cancer cells (left) and original identity (right). **D.** UMAP of the single-cell RNA-seq dataset grouped by clusters at a 0.025 resolution. A total of 8 clusters were identified that were composed of different proportions of normal and cancer cells with clusters 1, 2, 4 and 6 being mostly cancer (90%) and cluster 3 being mostly normal cells (53%). **E.** Heatmap (left) of proportions of canonical ovarian cancer cells per clusters showing that a high proportion of each cell type fell into a single cluster. Dot plot (right) of canonical markers per cell type identified in single-cell RNA-seq confirmed the correct labeling of cell types based on marker expression.
** p-value < 0.01; *** p-value < 0.001

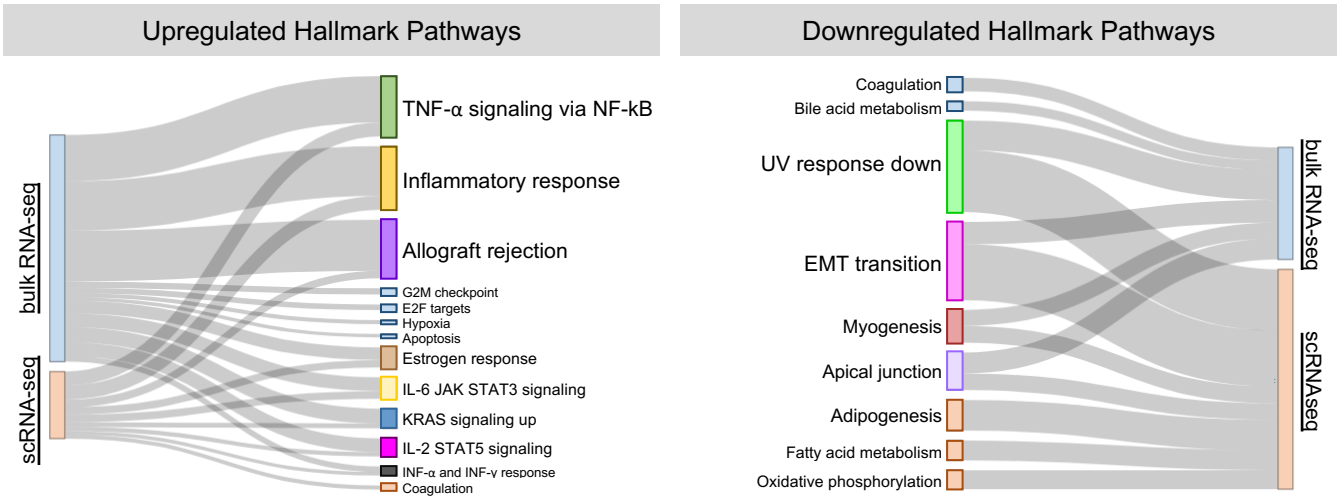
A



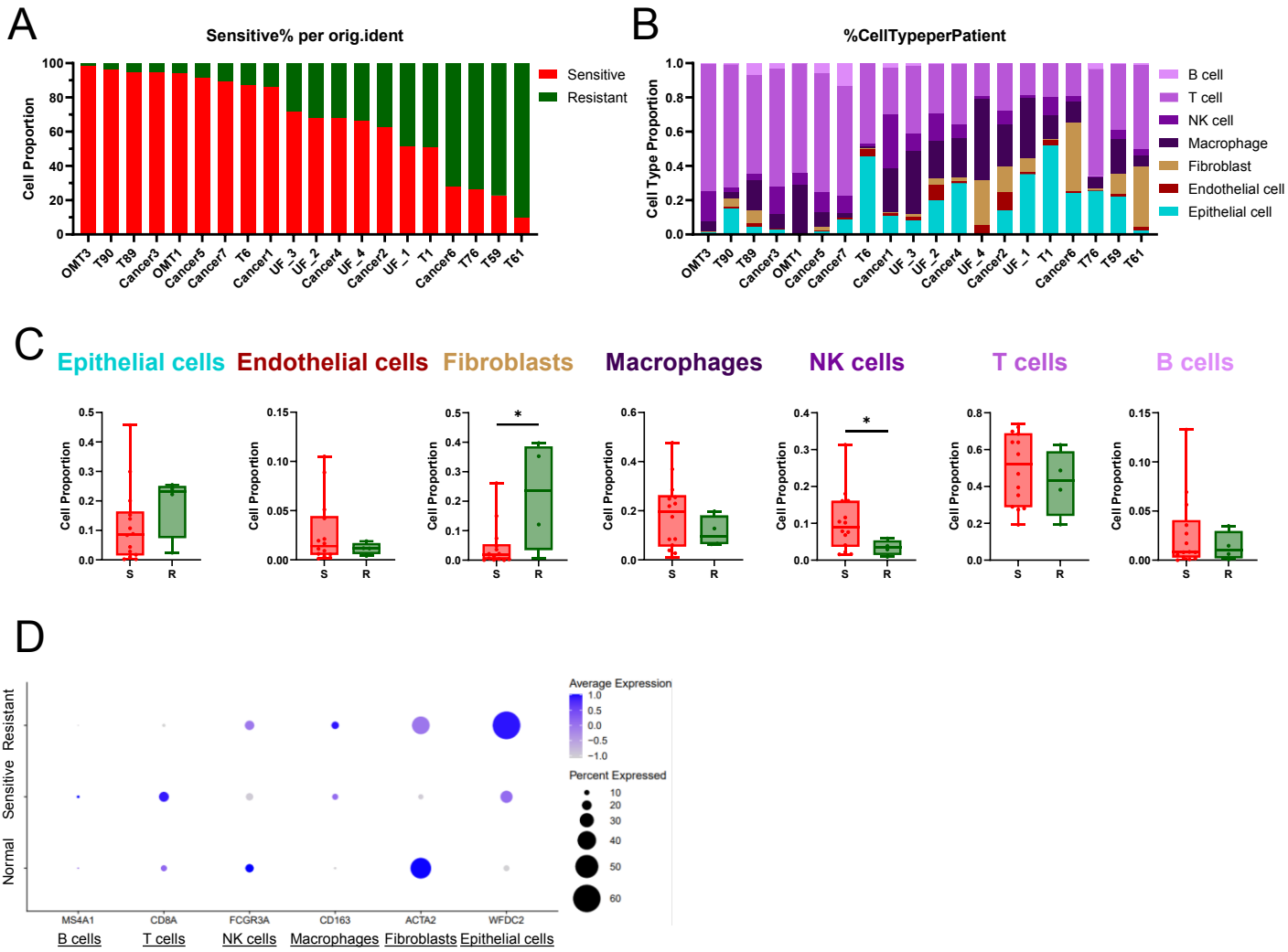
B



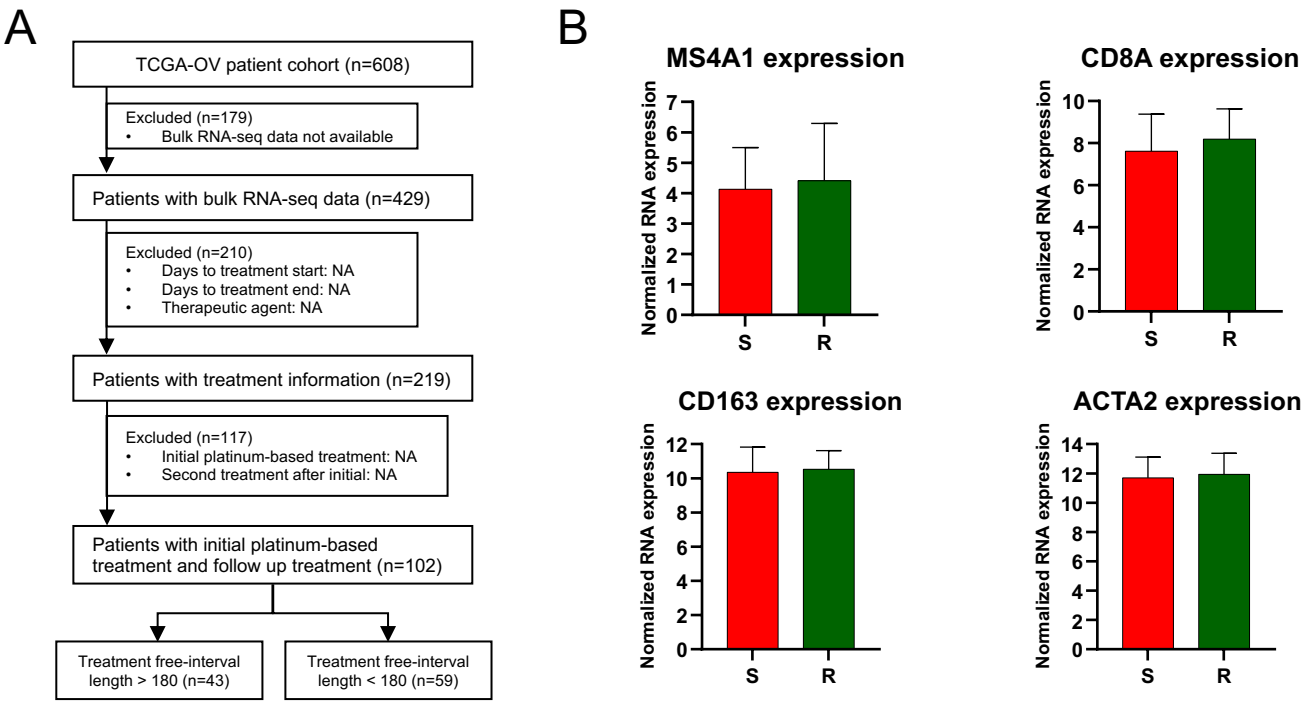
Supplementary Figure S4. Macrophages, B-cells, and fibroblasts were equally captured with mIF, bulk- and single-cell RNA-seq technologies. A. Bland-Altman plots showed agreement between mIF and bulk RNA-seq at capturing proportions of macrophages, B-cells, T-cells and fibroblast in non-cancerous (n=7) and cancerous (n=6) samples. Both technologies had high agreement at capturing macrophages, B-cells, and fibroblast proportions, while T-cell proportions were biased towards bulk RNA-seq at high percentages and biased towards mIF at low percentages. **B.** Bland-Altman plots of four fresh cancer samples with matched bulk- and single-cell RNA-seq information. Macrophage, B-cell, and fibroblast proportions showed high agreement between both technologies, while T-cell proportions were biased towards single-cell RNA-seq.



Supplementary Figure S5. Bulk- and single-cell RNA-seq capture similar Hallmark pathway changes upregulated and downregulated in cancer compared to normal samples. Sankey plot of upregulated and downregulated Hallmarks pathways in cancer compared to normal samples based on adjusted p-values for both bulk- and single-cell RNA-seq datasets.



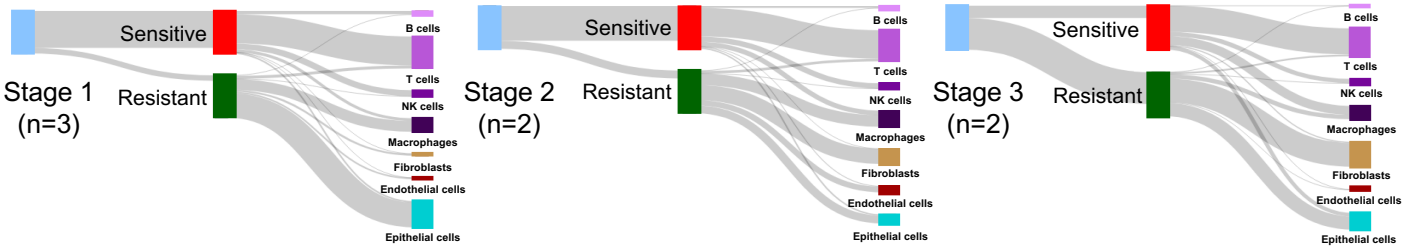
Supplementary Figure S6. A. Sensitivity proportions across all single-cell RNA-seq samples. Initially, T59, T61, and T76 were classified as 100% resistant, while T89 and T90 were classified as 100% sensitive. A subset of cells were used as a reference in singleR for sensitivity re-classification in the GSE154600 cohort and those cells were used as a reference to classify the rest of the cancer cells based on sensitivity. B. Cell proportions per patient samples. C. Box and whiskers plots of cell type proportions grouped by sensitive patients (those with > 60% sensitive cells, n=14) and resistant patients (those with > 60% resistant cells, n=4). Generally, sensitive patients showed higher proportions of NK cells while resistant patients showed higher proportions of fibroblasts. **D.** Dot plot of canonical marker expression of cell types of interest across normal, sensitive, and resistant samples. Macrophage and epithelial cells markers (CD163 and WFDC2) showed highest expression in resistant cells, while CD8A for T-cells showed highest expression in sensitive cells. * p-value < 0.05



Supplementary Figure S7. A. Consort diagram of TCGA bulk RNA-seq sample selection criteria. Ultimately, patients with an initial treatment-free interval (TFI) > 180 days were considered sensitive to platinum-based chemotherapies, while patients with an initial TFI < 180 days were considered resistant. **B.** Normalized RNA expression comparison between sensitive and resistant bulk RNA-seq samples for four markers of interest. There were no significant differences in the expression across the two groups for the markers.

A

Treatment-naïve



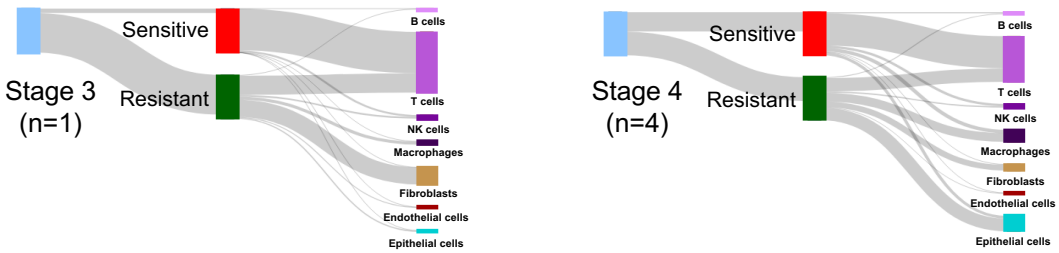
	% per sensitivity	% per cell type
Stage 1	86.99	7.72
		68.08
		14.05
		6.59
		0.01
		0.09
		3.46
	13.01	1.10
		6.84
		0.07
		25.46
		4.52
		3.63
		58.38

	% per sensitivity	% per cell type
Stage 2	81.61	4.39
		66.73
		12.54
		11.16
		0.08
		1.93
		3.17
	18.39	0.69
		6.07
		0.00
		26.97
		35.31
		12.97
		18.00

	% per sensitivity	% per cell type
Stage 3	27.83	0.09
		59.26
		9.86
		20.09
		2.49
		0.18
		8.02
	72.17	0.04
		3.95
		0.18
		9.95
		54.07
		1.56
		30.25

B

Treated



	% per sensitivity	% per cell type
Stage 3	9.75	0.00
		94.62
		4.31
		1.06
		0.00
		0.00
		0.00
	90.25	1.63
		43.71
		3.84
		6.93
		39.08
		1.56
		2.72

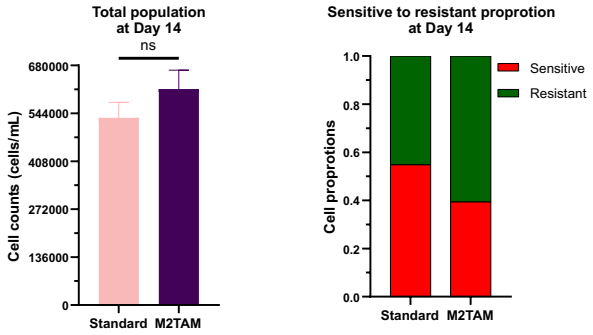
	% per sensitivity	% per cell type
Stage 4	44.65	2.59
		74.84
		5.59
		8.10
		1.50
		0.99
		6.37
	55.35	1.90
		31.00
		2.97
		20.08
		13.48
		1.24
		29.33

Supplementary Figure S8. Sankey and tables plots of single-cell RNA-seq cells at different stages and their proportions of sensitive and resistant cells and cell type proportions out of those 2 compartments. **A.** Stages 1 and 2 were treatment-naïve and composed of mostly sensitive cells. **B.** Stages 3 and 4 were treated and mostly composed of resistant cells.

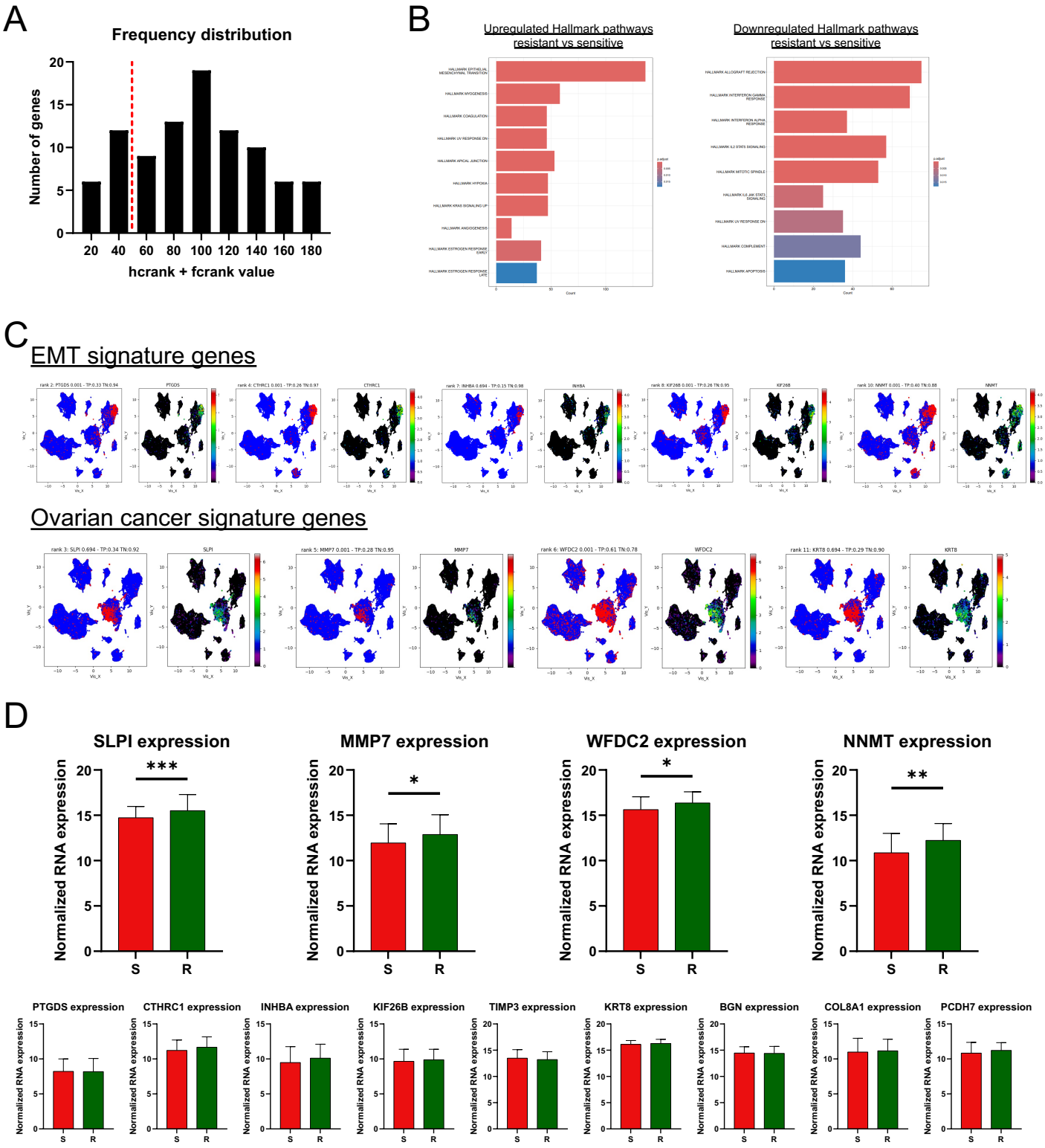
A

Cell type	Growth rate (day^{-1})	Carrying capacity (counts/day)
Tyk-nu standard media	1.463	257793.504
Tyk-nu M2 media	1.337	175425.532
Tyk-nu cp.r standard media	1.066	277519.584
Tyk-nu cp.r M2 media	0.729	370045.979

B



Supplementary Figure S9. A. Table containing logistic growth rates and carrying capacities for Tyk-nu (sensitive) and Tyk-nu cp.r (resistant) ovarian cancer cell grown in co-culture *in vitro*. Tyk-nu cp.r cells under M2 tumor-associated macrophage (TAM) conditioned media showed the lowest growth rate but the highest carrying capacity. **B.** Bar plots showed that the total populations at day 14 were the same size, however the ratio of sensitive to resistant cells was different.



Supplementary Figure S10. A. Histogram of hcrank + fcrank values across the genes output by COMET. Red dashed line represents the top 15th percentile of these values resulting in 14 top genes. **B.** Bar plots of pathways upregulated (left) and downregulated (right) in resistant cells compared to sensitive cells only. The EMT pathway is the most upregulated while immune pathways are the most downregulated. **C.** UMAPs showing the expression of the top 14 genes output by COMET showing both the binary and XL-mHG test outputs grouped by characteristic EMT genes and characteristic ovarian cancer genes. **D.** Expression of resistant markers identified by COMET in the TCGA bulk RNA-seq cohort. The expression of SLPI, MMP7, WFDC2, and NNMT was higher in resistant samples compared to sensitive samples. * p-value < 0.05; ** p-value < 0.01; p-value 0.001