Supplementary Methods

Immune phenotype in high- versus low-NRF2 high grade serous ovarian cancer and the impact on prognosis

Samera H. Hamad[1,2,3]*, Chelsea Katz[1#], Helen Toma[1#], Kosuke Murakami[4,5], Nasrine Bendjilali[6], Gord Zhu[1,3], Hadi Shojaei[1,3], Lanlan Fang[7,8], Samuel Leung[7,8], Martin Koebel[9], Huseyin Karaduman[10], Oliver Abinader[11], Ramkrishna Mitra[11], Lauren Krill[1,2,3], Christina Chu[1,2,3], David P. Warshal[1,2,3], Yemin Wang[7,8]

[1]Department of Surgery, Cooper University Health Care, Camden, New Jersey, United States.

[2]Department of Surgery, Cooper Medical School of Rowan University, Camden, New Jersey, United States.

[3]MD Anderson Cancer Center at Cooper, Camden, New Jersey, United States.

[4]Department of Obstetrics and Gynecology, Kindai University Faculty of Medicine, Osaka, Japan

[5]The Bloomberg-Kimmel Institute for Cancer Immunotherapy, Johns Hopkins University, Baltimore, Maryland, United States.

[6]College of Science and Mathematics, Rowan University, Glassboro, New Jersey, United States.

[7]Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada

[8]Ovarian Cancer Research Centre, Vancouver Coastal Health Research Institute, BC, Canada

[9]Department of Pathology, University of Calgary, Calgary, AB, Canada

[10]North Carolina State University, Raleigh, North Carolina, United States.

[11]Division of Biostatistics and Bioinformatics, Department of Pharmacology, Physiology, and Cancer Biology, Thomas Jefferson University, Philadelphia, PA, United States.

 # Equally contributed to this work

***To whom the correspondence should be addressed**

Samera Hamad, PhD: Cooper Medical School of Rowan University, Camden, NJ 08103, United States. Email. hamad@rown.edu or samera3h@gmail.com, Ph. 608-217-3839

Single cell RNA-seq data analysis of human HGSOC.

Seven human HGSOC samples from the GSE184880 [1] dataset were downloaded from The National Center for Biotechnology Information website (https://www.ncbi.nlm.nih.gov/). Data processing and core analysis were performed using Seurat v4.9.9 [2]. Cells with gene count below 200 and mitochondrial gene expressions above 25% were excluded. Normalization and variance stabilization were performed by sctransform v2, following the instructions from the Satijalab website (https://satijalab.org/seurat/). Then, Principal Component Analysis (PCA) dimension reduction and Uniform Manifold Approximation and Projection (UMAP) dimensional reduction were performed based on the first 30 principal components. Cluster identification was performed, and a resolution was optimized for fine separation. The estimated doublets were excluded by the DoubletFinder v2.0 package following the instructions on the website found here https://github.com/chris-mcginnis-ucsf/DoubletFinder. Each preprocessed sample data was integrated by Seurat v4.9.9. The datasets produced were used for downstream analysis. The core analysis was performed using Seurat v4.9.9. The annotation of each cluster was performed using SingleR v2.0.0 package (https://github.com/dviraran/SingleR). In SingleR, “MonacoImmuneData” was applied, and the resulting scores were displayed in hierarchical clustering. Single-sample GSEA was performed using the escape v1.8.0 package (https://github.com/ncborcherding/escape).

Data Analysis of human HGSOC RNA-seq from TCGA

To identify the major mutations in key genes associated with HGSOC samples used in this study, the mutation profile of ovarian cancer samples available at the Cancer Genome Atlas (TCGA) was analyzed using the cBioPortal database [3]. The mutational analysis included a cohort of 365 samples, among which mutations in the queried genes were observed in 286 samples. The analysis of NRF2 and immune gene score was based on well-established NRF2-gene signature and immune-gene signature [4-6]. A single sample gene-set enrichment analysis (ssGSEA) was used to calculate NRF2 score and immune score per each sample**.** To identify pathways related-hallmarks of cancer, we downloaded gene sets from MSigDB (https://www.gsea-msigdb.org/gsea/msigdb) [7], performed single-sample ssGSEA, and used R package GSVA to employ ssGSEA**.** We compared pathways in $NRF2^{High}$ versus $NRF2^{Low}$ group.

To identify the differentially expressed genes in $NRF2^{High}$ versus $NRF2^{Low}$ samples, raw count data from RNA sequencing were analyzed using the *DESeq2* R package [8]. Genes with low counts across all samples were filtered out by retaining only those with a total count of at least 10 across all samples. To account for differences in sequencing depth, normalization analysis was performed using the median-of-ratios method. Differential expression was assessed using a negative binomial generalized linear model with Wald test statistics. The resulting *P* values were adjusted for multiple testing using the Benjamini–Hochberg method to control the false discovery rate (FDR) [9]. Genes with an adjusted *P* value ≤0.05 and $|\log_2 FC| > 1$ were considered significantly differentially expressed.

Survival analysis was performed using the “survival” package in R. To evaluate differences in survival outcomes, we employed Kaplan–Meier method using the “survdiff” function to estimate survival probabilities over time. Separate analyses were conducted for 1) Overall Survival (OS), 2) Disease-Specific Survival (DSS), 3) Disease-Free Interval (DFI), and 4) Progression-Free Interval (PFI). For survival related immune markers, patients were categorized into two groups

"high" and "low" immune expression levels of each immune marker, in each ($NRF2^{High}$ and $NRF2^{Low}$) group. Immune expression levels were based on the median of normalized expression values where values above the median were classified as high, and those below the median were classified as low. Group differences were assessed using the log-rank test. For visualization of survival curves, the “survminer” R package was used. All analyses were performed using R 4.3.1 (https://cran.r-project.org/).

Tissue processing, Hematoxylin & Eosin (H&E), and immunohistochemistry (IHC) labeling of human tissue microarray (TMA)

The use of patient tumor tissue and outcome data was approved by the University of British Columbia ethical committee (H18-01652). Cases from patients who had undergone surgery and/or treatment at Vancouver General Hospital (2000-2008) with a final diagnosis of HGSOC were selected for second review by Dr. M Kobel. For tissue microarray (TMA) construction, 2 x 0.6 mm cores were taken from each case of the formalin-fixed paraffin-embedded (FFPE) blocks. Each TMA slide was stained for Hematoxylin and Eosin (H&E) or immunohistochemistry (IHC) as previously described [10].

IHC staining of P53 (Dako, DO-7, 1:500), WT1 (Dako, 6F-H2, ready for use), Vimentin (VIM) (Dako, V9, ready for use) and CD8 (Abcam, ab227707, 1:200) was performed at the Anatomical Pathology Department of the Vancouver General Hospital using the Ventana Benchmark XT stainer. H&E staining and IHC-PAX8 labeling (Novus, NB100-1065, 1:500) were performed in the MAPCORE lab of UBC using the Ventana Discovery XT platform. IHC labeling of NRF2 (Abcam, #ab62352, 1:150); FOXP3 (Novus, NB100-39002, 1:200); CD68 (DAKO, MO814, 1:3000), was done in the BASIC lab of BC Cancer Research Institute using the Ventana Discovery XT platform. IHC-PDCD1/PD1 and CD274/PD-L1 labeling was performed at the Anatomical Pathology Department of BC Cancer Agency-Vancouver. IHC-PDCD1/PD1 labeling (Clone NAT105, Cell

Marque,1:100) was done on Dako Omnis platform, while CD274/PD-L1 labeling (22C3 pharmDx assay, Dako, ready for use) was done on Dako Autostainer Link48 platform.

IHC slides were examined using an Olympus BX43 Light Microscope. Cells that labeled yellow to brown was considered IHC positive cells. All markers were scored using both 100X and 400X magnifications for the intensity and percentage of the labeled cells. Two variable factors were considered in scoring (1) the number of positively labeled cells; and (2) the intensity of the staining (Supplementary Table 7a) [11]. While scoring the IHC slides, our certified pathologists Drs. Zhu and Shojaei were blinded from the NRF2 status.

Referenced Cited

1 Xu J, Fang Y, Chen K, Li S, Tang S, Ren Y *et al*. Single-Cell RNA Sequencing Reveals the Tissue Architecture in Human High-Grade Serous Ovarian Cancer. *Clinical Cancer Research* 2022; 28: 3590-3602.

2 Hao Y, Hao S, Andersen-Nissen E, Mauck WM, 3rd, Zheng S, Butler A *et al*. Integrated analysis of multimodal single-cell data. *Cell* 2021; 184: 3573-3587.e3529.

3 Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA *et al*. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2012; 2: 401-404.

4 Singh A, Boldin-Adamsky S, Thimmulappa RK, Rath SK, Ashush H, Coulter J *et al*. RNAi-mediated silencing of nuclear factor erythroid-2-related factor 2 gene expression in non-small cell lung cancer inhibits tumor growth and increases efficacy of chemotherapy. *Cancer Res* 2008; 68: 7975-7984.

5 Cescon DW, She D, Sakashita S, Zhu CQ, Pintilie M, Shepherd FA *et al*. NRF2 Pathway Activation and Adjuvant Chemotherapy Benefit in Lung Squamous Cell Carcinoma. *Clin Cancer Res* 2015; 21: 2499-2505.

6 Xie Q, Chu H, Yi J, Yu H, Gu T, Guan Y *et al*. Identification of a prognostic immune-related signature for small cell lung cancer. *Cancer Med* 2021; 10: 9115-9128.

7 Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 2015; 1: 417-425.

8 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014; 15: 550.

9 Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* 2001; 125: 279-284.

10 Yamano S, Gi M, Tago Y, Doi K, Okada S, Hirayama Y *et al*. Role of deltaNp63(pos)CD44v(pos) cells in the development of N-nitroso-tris-chloroethylurea-induced peripheral-type mouse lung squamous cell carcinomas. *Cancer Sci* 2016; 107: 123-132.

11 Hamad SH, Montgomery SA, Simon JM, Bowman BM, Spainhower KB, Murphy RM *et al*. TP53, CDKN2A/P16, and NFE2L2/NRF2 regulate the incidence of pure- and combined-small cell lung cancer in mice. *Oncogene* 2022; 41: 3423-3432.