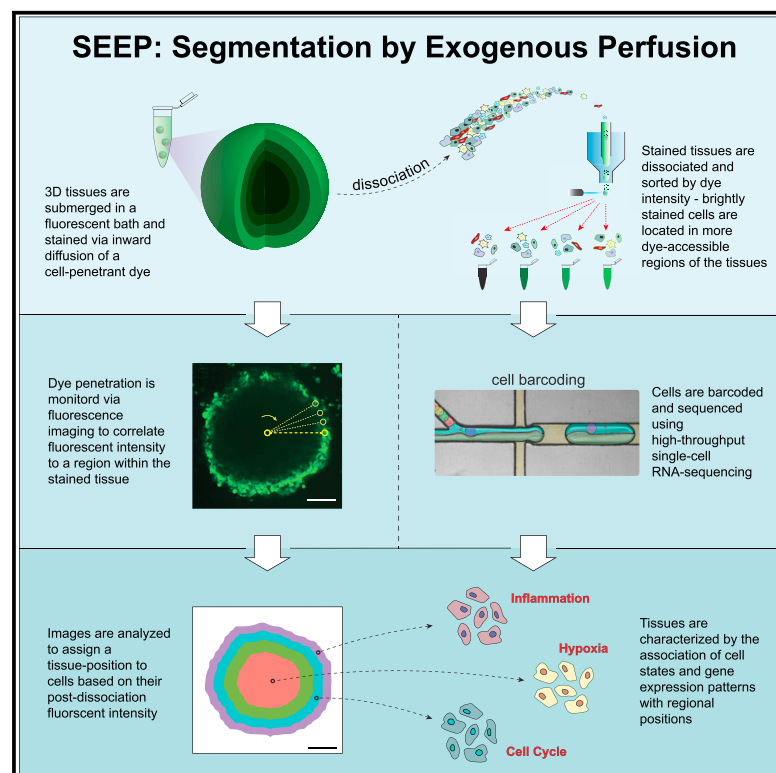


Positional influence on cellular transcriptional identity revealed through spatially segmented single-cell transcriptomics

Graphical abstract



Authors

David B. Morse,
Aleksandra M. Michalowski,
Michele Ceribelli, ...,
Jason D. Buenrostro,
Tuomas P.J. Knowles, Craig J. Thomas

Correspondence

tpjk2@cam.ac.uk (T.P.J.K.),
craigt@mail.nih.gov (C.J.T.)

In brief

Spatially resolved single-cell RNA sequencing was achieved by segmenting 3D tissues into concentric layers via the uptake and steady-state distribution of a small-molecule dye. Ovarian cancer models were analyzed to reveal spatial variations in gene expression across tissue depths. The method is rapid, highly adaptable, and optimized for non-specialized laboratories.

Highlights

- Perfusion and uptake of a small-molecule dye enables spatially resolved transcriptomics
- Engineered punch biopsies retain spatial transcriptomic information from harvested tissue
- Immunoreactive gene expression patterns associate with surfaces of ovarian cancer models
- Spatial expression patterns are cross-identified in a published ovarian cancer dataset



Article

Positional influence on cellular transcriptional identity revealed through spatially segmented single-cell transcriptomics

David B. Morse,^{1,2,3,4} Aleksandra M. Michalowski,⁵ Michele Ceribelli,² Joachim De Jonghe,⁶ Maria Vias,⁷ Deanna Riley,⁸ Theresa Davies-Hill,⁸ Ty Voss,² Stefania Pittaluga,⁸ Christoph Muus,^{3,4} Jiamin Liu,⁹ Samantha Boyle,⁷ David A. Weitz,^{3,10} James D. Brenton,⁷ Jason D. Buenrostro,^{4,11} Tuomas P.J. Knowles,^{1,12,*} and Craig J. Thomas^{2,13,14,*}

¹Cavendish Laboratory, Department of Physics, University of Cambridge, J J Thomson Ave, Cambridge, UK

²Division of Preclinical Innovation, National Center for Advancing Translational Sciences, National Institutes of Health, Bethesda, MD, USA

³John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA

⁴Broad Institute of MIT and Harvard, Cambridge, MA, USA

⁵Laboratory of Cancer Biology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

⁶Department of Biochemistry, University of Cambridge, Cambridge, UK

⁷Cancer Research UK, Cambridge Research Institute, Li Ka Shing Centre, Robinson Way, Cambridge, UK

⁸Laboratory of Pathology, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

⁹Advanced Imaging and Microscopy Resource, National Institutes of Health Clinical Center, Bethesda, MD, USA

¹⁰Department of Physics, Harvard University, Cambridge, MA, USA

¹¹Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA, USA

¹²Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, UK

¹³Lymphoid Malignancies Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

¹⁴Lead contact

*Correspondence: tpjk2@cam.ac.uk (T.P.J.K.), craigj@mail.nih.gov (C.J.T.)

<https://doi.org/10.1016/j.cels.2023.05.003>

SUMMARY

Single-cell RNA sequencing (scRNA-seq) is a powerful technique for describing cell states. Identifying the spatial arrangement of these states in tissues remains challenging, with the existing methods requiring niche methodologies and expertise. Here, we describe segmentation by exogenous perfusion (SEEP), a rapid and integrated method to link surface proximity and environment accessibility to transcriptional identity within three-dimensional (3D) disease models. The method utilizes the steady-state diffusion kinetics of a fluorescent dye to establish a gradient along the radial axis of disease models. Classification of sample layers based on dye accessibility enables dissociated and sorted cells to be characterized by transcriptomic and regional identities. Using SEEP, we analyze spheroid, organoid, and *in vivo* tumor models of high-grade serous ovarian cancer (HGSOC). The results validate long-standing beliefs about the relationship between cell state and position while revealing new concepts regarding how spatially unique microenvironments influence the identity of individual cells within tumors.

INTRODUCTION

High-throughput single-cell RNA sequencing (scRNA-seq) is used to describe complex tissues by characterizing the transcriptional states of individual cells. scRNA-seq yields unparalleled granularity with regards to understanding cellular identity and function in complex tissues as well as informing on mechanisms of pathology in disease.^{1,2} Common high-throughput scRNA-seq methods, however, require tissue dissociation prior to sequencing and, consequently, decouple cells from their original positions within tissues.^{1–3} Defining a cell location, relative to tissue margins, vascularization networks, and social context, is essential for understanding the variables that influence its transcriptional identity. In cancer, significant spatial genomic hetero-

geneity exists within tumors, arising both from distinct clones and the response of these clones to their environment.⁴ Tissue architectures impact disease induction and progression and are used for morphological classifications of disease pathology via immunohistochemistry and immunocytology.^{5,6}

Ideally, methods for correlating spatial cell position with transcriptional identity would couple methods accessible to most laboratories with next-generation sequencing technologies. Methods are emerging to extend the flexibility, breadth, and resolution of spatially resolved transcriptomic profiling. Technologies that integrate scRNA-seq data onto *in situ* hybridization (ISH) gene expression data of similar tissues have been successfully applied to tissues with organized morphology but are less applicable for tumors.^{7,8} Slide-based methods capture mRNA



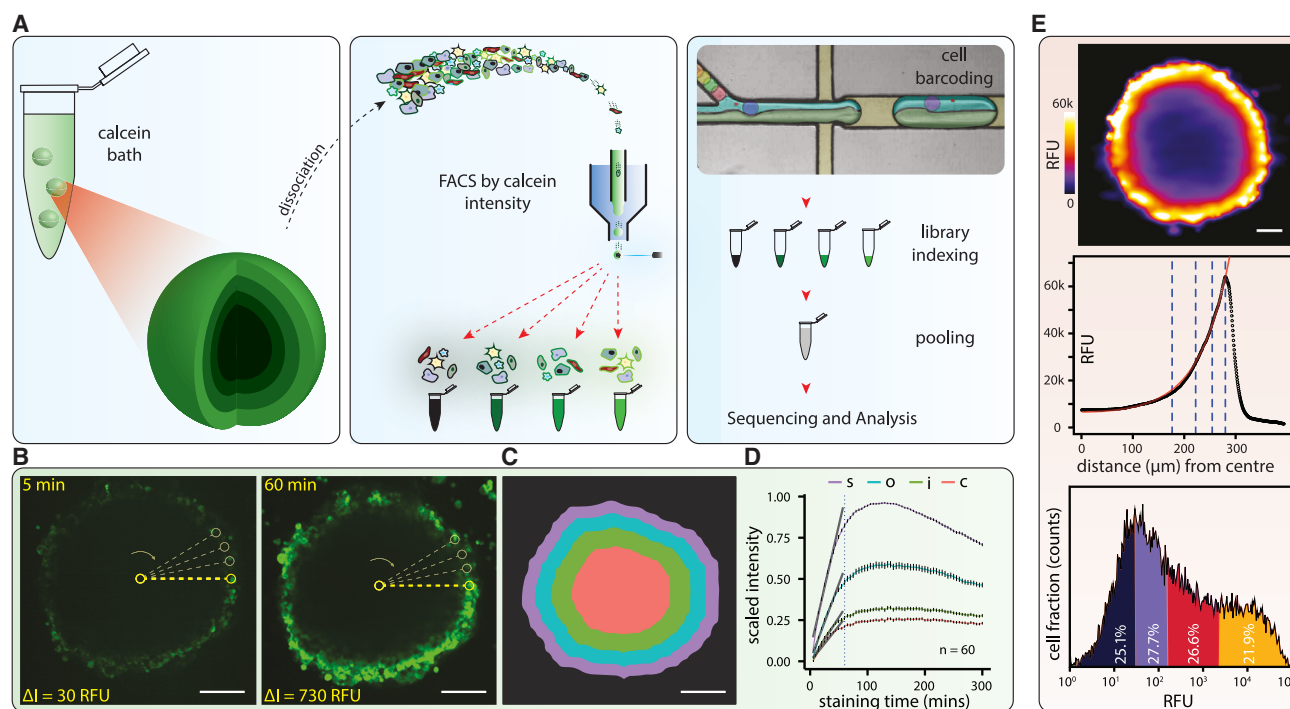


Figure 1. Segmentation by exogenous perfusion overview

(A) Cartoon schematic of the SEEP workflow for spheroid cultures. Visualized steps include calcein AM bathing, dissociation, sorting, binning, sequencing, and analysis.

(B) Example image showing the difference in fluorescent intensity of a representative HGSOc spheroid across a time course (5 and 60 min) resolved using confocal microscopy. At $t = 60$ min, cells on the surface are over 700 relative fluorescent units (RFU) brighter than core cells.

(C) At $t = 60$ min, spheroids were segmented into four concentric shells using a convoluted neural network.

(D) The fluorescence intensity of each layer ($n = 60$ spheroids) was monitored in time. At $t = 60$ min, each layer's mean fluorescence differed by over 5% and remained stable for over 100 min. Linear fits from $0 \rightarrow 50$ min (gray, R^2 values ≥ 0.97) show a linear accumulation of calcein as predicted by the integrated form of Equation (3) $C_f \propto At + B$ where $A = k_1 C$. $A = (0.015, 0.0093, 0.0054, 0.0044)$ and $B = (0.053, -0.0012, -0.0046, -0.0034)$ for the surface, outer, inner, and center layers of the spheroids respectively.

(E) Example image showing the fluorescent intensity of a representative HGSOc spheroid at $t = 60$ min resolved using confocal microscopy (top). An intensity profile of an individual spheroid (middle) and a FACS profile of 192 dissociated spheroids (bottom) show the distribution of fluorescence intensities across individual cells and the thresholds (dashed blue lines) used for segmentation. A hyperbolic sine fit to the middle panel (red, $R^2 > 0.99$) shows a fluorescence distribution profile predicted by Equation (2). The fit corresponds to the equation, $C = AC_R \frac{R \sinh(\phi_1 r/R)}{r \sinh(\phi_1)} + B$, where A and B are the scaling factors 119501 RFU and 4150 RFU respectively. Scale bars, (B and C) 150 μm , (E) 100 μm .

on spatially patterned grids of pre-defined barcodes and are useful for spatially resolving 2D tissue slices with increasing fine resolution.^{9–11} These methods can be performed iteratively on consecutive tissue slices to reveal three-dimensional (3D) tissue architectures. Other innovative methods include multiplexed error-robust fluorescence *in situ* hybridization (MERFISH) and fluorescent *in situ* sequencing (FISSEQ) that enable highly multiplexed ISH and *de novo* sequencing in intact tissues.^{12,13} These approaches provide high-resolution spatial maps but are limited in their ability to interrogate 3D tissue architectures and require significant laboratory specialization.

Here, we describe Segmentation by Exogenous Perfusion (SEEP), an integrated method for correlating environmental accessibility of cells within 3D disease models with scRNA-seq data. SEEP utilizes the diffusion kinetics of the small-molecule dye, calcein acetoxymethyl (AM), to establish a stable dye gradient within 3D tissues that varies in intensity according to radial-spatial cell position as a function of accessibility to the tissue surface. This gradient is used with fluorescence-activated

cell sorting (FACS) to sort and bin dissociated cells by diffusion distance to tissue surfaces. Single cells are then barcoded and sequenced using any standard scRNA-seq protocol (Figure 1A). The method can define the surface accessibility and radial-spatial positions of cells in both symmetric and asymmetric cell cultures (e.g., spheroids and organoids) and linear tissue samples (e.g., punch biopsies). We use SEEP to study the transcriptional architecture of high-grade serous ovarian cancer (HGSOc) cells located at defined layers of three different 3D HGSOc models including spheroids, organoids, and patient-derived xenograft (PDX) tumors. We explore how transcriptional identity is dependent on cells' surface accessibility in HGSOc disease models arising from both individual clones (e.g., cultured spheroids) and ascites-derived cell populations (e.g., organoid and PDX models). In doing so, we show how cellular identity varies along an axis defining tissue margins and nutrient accessibility.¹⁴

HGSOc, the most lethal gynecological cancer, is characterized by genomic instability from ubiquitous TP53 mutations and a high percentage of BRCA1/2 mutations.^{15,16} Despite

advancements in the genomic characterization of HGSOc, significant improvements in disease outcomes have not emerged over the past 30 years. In particular, metastatic disease remains a significant clinical challenge and is associated with resistance to therapy and decreased life expectancy.^{15–17} HGSOc cells disseminate into patients' ascites and natively cluster into 3D spheroid-like bodies that are believed to be responsible for metastatic disease found in the peritoneum and beyond.¹⁸ HGSOc cell clusters persist in the peritoneal cavity years after a primary tumor is removed and are capable of surviving both immune recognition and chemotherapy.¹⁷ Understanding the transcriptional architectures of these 3D cell clusters and the metastatic lesions they form will provide insights into HGSOc disease progression and potentially inspire new therapeutic concepts to combat this disease.

RESULTS

Diffusion-mediated accessibility staining for spatial segmentation

To preserve surface and environmental proximity information from 3D HGSOc models throughout a high-throughput scRNA-seq assay, we developed a perfusion method capable of segmenting tissues into concentric shells and an inner core based on cellular accessibility to a cell-permeable fluorescent dye. In this way, cells could be sorted and binned as a function of their diffusion distance from a tissue surface (Figures 1A–1D). Although SEEP can be performed using a variety of small-molecule dyes, we established the method using calcein AM, a non-fluorescent analog of calcein that utilizes an AM ester moiety to allow for passive diffusion across cell membranes.^{19,20} Once inside a cell, intracellular esterases convert calcein AM to the pH-independent fluorescent calcein and trap the dye inside cells.²¹ To properly enable the SEEP methodology, it was critical to define the diffusion parameters of each model upon exposure to calcein AM. With no barriers to diffusion, calcein AM and calcein (diffusivity, $D_e = 260$ and $500 \mu\text{m}^2/\text{s}$ in media, respectively) would diffuse into a spherical cavity in time, $t = (r^2)/6D_e$ (the equivalent of 58 and 30 s to diffuse into a $600 \mu\text{m}$ diameter sphere). 3D cell cultures and tissue samples, however, create heterogeneous and complex barriers to diffusion which complicate the parameters that govern inward flux and dye uptake.²² For calcein AM, uptake into a spherical cell cluster is dependent on both its effective diffusion through the complex barrier and the reaction rate for converting calcein AM to calcein. The reaction-diffusion equation for C (calcein AM) undergoing diffusional transport and conversion into the fluorescent product calcein in a spherically symmetric system is:

$$\frac{\partial C}{\partial t} = \frac{D_e}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial C}{\partial r} \right) - k_1 C. \quad (\text{Equation 1})$$

Given the large excess of calcein AM in the media surrounding our samples, we estimate the external dye concentration to be fixed at the surface of the sphere, $r = R$. That is, we impose $C(R) = C_R$, and Equation (1) will attain a steady state given by:

$$C = C_R \frac{R \sinh\left(\frac{\varphi_1 r}{R}\right)}{r \sinh(\varphi_1)} \quad (\text{Equation 2})$$

and illustrated by the hyperbolic sine curve fit to the radial fluorescent profiles observed in the spheroid dye accumulation (Figure 1E, red line, $R^2 > 0.99$). In equation (2), $\varphi_1 = R\sqrt{k_1/D_e}$. This is a special case (for the first-order reaction $n = 1$ of calcein AM $\rightarrow$ calcein) of the Thiele modulus, φ_n^2 , where $\varphi_n^2 = k_n R^2 C_{AS}^{n-1}/D_e$. Here, k is the rate constant of an n order reaction, R is the radius of the sphere, C_{AS} is the concentration of calcein AM at the surface of the spherical cell cluster, and D_e is the effective diffusivity.^{23–25} φ_n^2 describes the ratio of a system's rate of reaction to rate of diffusion. When the reaction is slower than the effective diffusion, $\varphi_n^2 \ll 1$, calcein (the product) fills the entire volume of the spherical cell cluster. When the reaction is faster than the effective diffusion, $\varphi_n^2 \gg 1$, calcein cannot reach the core of the spherical cell cluster and a non-uniform steady state is reached. Achieving a steady state in dye distribution is critical for the SEEP method and facilitates the correlation of cell brightness with exogenous accessibility and, by extension, radial-spatial cell positioning. Experimentally, we found that for each of the ascites-derived HGSOc samples analyzed, $\varphi_n^2 \gg 1$ and a steady state was achieved and maintained for time scales appropriate for tissue processing (Figures S1 and S2). Non-fluorescent calcein AM constantly diffuses into the system at a steady state and is converted into calcein. Because of this, the accumulation of fluorescent dye, C_f is linear and satisfies the following equation:

$$\frac{\partial C_f}{\partial t} = k_1 C. \quad (\text{Equation 3})$$

Therefore, once the steady state of C is realized, the model predicts a linear accumulation in the fluorescent form, calcein. Indeed, we observe a linear accumulation of calcein from $0 \rightarrow 50$ min (Figure 1D, gray lines, $R^2 > 0.97$). Above ~ 50 min, we attribute the fall of fluorescent calcein accumulation to calcein degradation within the live cells. This decrease in fluorescence is clearly observed from $160 \rightarrow 300$ min; however, the mean fluorescent stratification between layers remains resolved throughout this time period. By fitting Equation (2) to the fluorescence distribution across the spheroids as a hyperbolic sine curve (Figure 1E), we found the ratio of k_1/D_e to be $0.0004 \mu\text{m}^{-2}$ and φ_1^2 to be 31.36 corresponding to the fluorescence profile of Figure 1E. The uptake and conversion of calcein AM into calcein were determined by fitting the integration of Equation (3) to the linear accumulation of fluorescence on the surface layer of the spheroids (Figure 1D).

Of practical importance, SEEP relies on the distance to the environmental interface of tissue rather than on the distance to the center of a tissue to segment populations into concentric layers. This allows the assay to be performed on irregularly shaped tissues rather than relying on tissue symmetry for accurate segmentation. In addition, the method's extension via Punch-SEEP enables spatial segmentation to be performed on biopsy punches harvested from a variety of tissue sizes and geometries.

Calibrating diffusion kinetics

SEEP requires an imaging-based calibration step for observing dye penetration kinetics and a measurement step for sorting and sequencing cells based on their fluorophore accessibility.

Each of the different 3D HGSOc cell models examined in this study (spheroids, organoids, and PDX solid tumors) required a unique collection of calibrating steps to define the specific diffusion kinetics associated with each model. We exposed each model to calcein AM dye and monitored dye penetration using light sheet or confocal microscopy (Figures 1B–1E, S1A–S1C, and S2A–S2E). The time that each model system needed to reach a maximum change in the fluorescence intensity between the center and surface was monitored and recorded across multiple replicates.

Fluorescence images acquired during the calibration step were normalized to signal attenuation incurred with imaging depth. High-content confocal microscopy imaged 100s of well-defined and homogeneously stained (ex/em: 353/466 nm, 492/517 nm, and 630/650 nm) spheroid samples to calculate the attenuation parameters (Figures S1D and S1E). Signal decay for homogeneously stained spheroid samples was calculated in depth and described by the exponential $y \approx 1.11e^{-0.21x} - 0.018$ (residual standard error: 0.029 on 27° of freedom) (Figures S1F and S1G). Attenuation was subtracted up to $\sim 100 \mu\text{m}$. Beyond this point, signal normalization was no longer an appropriate estimate for dye accumulation, and levels beyond this threshold were not used in the segmentation calculations. After sample attenuation was calculated, the fluorescence intensity measurements of the diffusion kinetics were normalized for signal attenuation. We then divided the 3D cell and tissue models into concentric rings, or sections, of equal volume; the average intensity ($\bar{I}$) of each layer and the central core were calculated (Figures S1H–S1K). To ensure adequate fluorescence separation between layers, volumes were segmented so that the $\Delta \bar{I}$ between adjacent layers exceed 5% (Figure 1D). This parameter also determined the binning process for equal volumetric sorting based on fluorescent intensity. Similar methods were used to define the calibration kinetics and binning parameters for the organoid and biopsy models (Figure S2).

Binning, sorting, sequencing, and clustering

For the measurement step, pre-calibrated 3D models were stained, sorted based on fluorescent shell thresholds, and single-cell barcoded for RNA sequencing. Briefly, following exposure to a fixed concentration of calcein AM for the calibrated time, the sample was washed of excess dye, dissociated using Accutase and gentle agitation (spheroid and organoid) or the gentleMACS dissociator (PDX biopsies), and sorted via FACS into a pre-defined number of bins (Figures 1D and 1E). No calcein diffusion between cells was observed post-dissociation, and imaging post-sorting confirmed accurate binning (Figures S1L and S1M). We also reviewed potential confounding factors like cell size and intracellular esterase expression patterns and found that these variables were consistent across layers (Figure S1M). We collected the sorted cells as aliquots of individual layers and immediately processed the layers through the inDrop platform.³ We fabricated and quality controlled the necessary microfluidics and barcoded hydrogel beads in house (Figure S1F; Tables S1A–S1G). We indexed each layer independently, pooled the libraries, and sequenced across 7 NextSeq 500/550 High-Output v2 Kits (75 cycles). To control for possible technical bias incurred during SEEP, we examined the effect of staining and sorting (via FACS) on single-cell transcriptional patterns of HGSOc PEO1 cells (Fig-

ure S3). We subjected cells from a PEO1 monolayer to conditions imposed by SEEP and examined if these conditions affected single-cell transcriptional patterns. We found that sorting, staining, and sorting + staining had a negligible effect on global gene expression patterns (Figures S3A–S3C), average gene expression among clustered transcriptional groups (Figures S3G–S3I), and single-cell expression patterns across genes of significant variability in our HGSOc model systems (Figures S3J). Additionally, we found that significant gene expression changes existed between clustered transcriptional states within a single condition (Figures S3D and S3E) but not between technical replicates of a single condition (Figures S3F). The only gene found to have significantly variable mean expression between conditions was HIST1H4C that encodes for a replication-dependent histone of the H4 family. HIST1H4C was found to be depleted in sort + stain cells compared with control cells; however, HIST1H4C was not one of the variable genes analyzed in this study. In addition, HIST1H4C was not a member of the key gene sets used to define cells in this study.

To enable the identification of transcriptional state variations with respect to positioning and environmental accessibility, we first identified highly variable genes for each model and used them to drive principal component analysis (PCA) for dimensional reduction. Dimensionally reduced cell identities were embedded in a k-nearest neighbor (KNN) graph of Euclidean distances defined by the PCA. Jaccard similarity was used to optimize the KNN graph, after which we used the Louvain algorithm to cluster cells by optimizing the standard modularity function (Figure S4).^{26–28} T-distributed stochastic neighbor embedding (t-SNE) and uniform manifold approximation and projection (UMAP) were used to visualize cell clusters in two dimensions.^{29–31}

We identified differentially expressed genes within each cluster and used them to define transcriptional states via over-representation analysis (ORA) and gene set enrichment analysis (GSEA) (Tables S2A–S2f). We incorporated the environmental accessibility information with the gene expression data by performing a chi-square test for independence to determine the degree of association between each spatially defined layer and transcriptionally defined cell cluster. As a visualization aid, we overlaid the positional information on top of t-SNE plots that represented gene expression data. This revealed correlations between cell positions and transcriptional patterns.

SEEP-enabled mapping of transcriptional states in HGSOc spheroids

Although 2D (monolayer) cell culture methods remain a dominant experimental platform, cells compelled into 3D (spheroid or near-spheroid) cultures have re-emerged over the past decade thanks to new methods and compelling evidence that 3D cultures are superior models of many disease phenotypes.^{32,33} Although spheroid cultures lack key elements of *in vivo* disease (e.g., cell heterogeneity, stromal components), they have found particular utility in modeling solid tumors including HGSOc.^{34,35} Here, we utilized the HGSOc cell line PEO1 to generate spheroids to assess the SEEP methodology and explore positional influence on cell identity. The PEO1 cell line was derived from a malignant effusion from peritoneal ascites and revealed through Tagged Amplicon deep sequencing (TAM-Seq) to have an allele

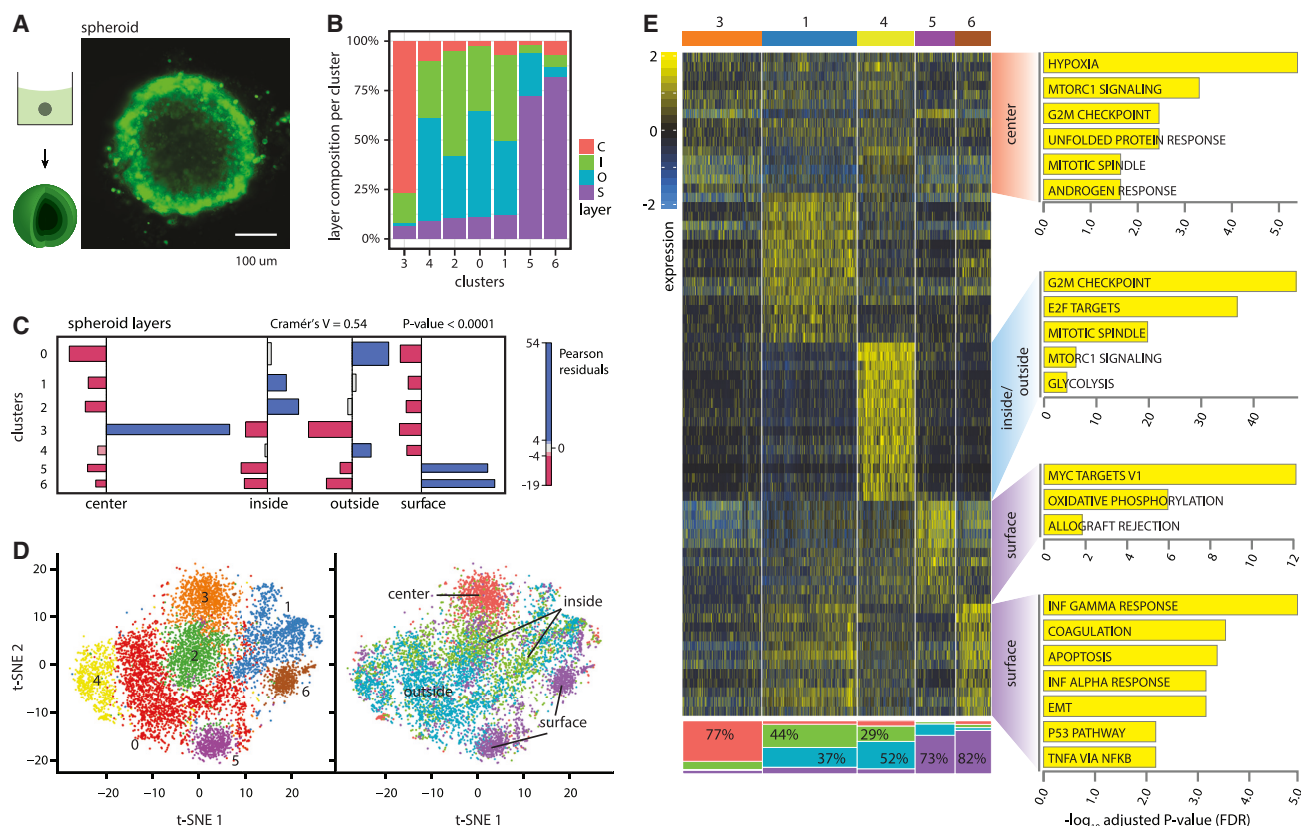


Figure 2. Positional analysis of HGSOC cultured PEO1 spheroids

(A) Cartoon schematic of the dye perfusion of a solitary spheroid in a calcein AM bath and a confocal image of a spatially segmented HGSOC spheroid. (B) Composition chart showing the seven transcriptionally defined cell clusters and the retrospective, layer-specific composition of each resolved cluster. Red, core; green, inside; blue, outside; purple, surface. (C) Bar-chart visualization of chi-square test of independence defined associations between gene expression clusters and spheroid layers. Blue, over-representation; red, under-representation; gray, null hypothesis in the cross tabulation. Bar width illustrates relative cluster size ($n = 1,178$ center cells, $n = 2,471$ inside cells, $n = 2,736$ outside cells, and $n = 2,667$ surface cells). (D) t-SNE visualization of gene expression clusters (left) and radial cell position (right) ($n = 7,908$ cells). (E) Heatmap visualization of selected gene expression clusters and their layer-specific composition (by %). Significant transcriptional Hallmark gene signatures from gene ontology (GO) are highlighted for clusters #3 (77% core cells), #4 (29% inside, 52% outside cells), #5 (73% surface cells), and #6 (82% surface cells). Color scale is linear. (Full accounting of Hallmark GO signatures can be found in [Tables S3A–S3G](#), and Hallmark signatures from GSEA can be found in [Figure S3](#) and [Tables S3H–S3N](#).) Scale bar, (A) 100 μm . GEO: GSE157299.

fraction > 99% mutated TP53 and BRCA2 genes signifying a nearly pure tumor fraction.^{36,37} Spheroidal cultures of PEO1 offer potential insight into the transcriptional programs that support metastases, including dissemination, migration, invasion, and seeding. We hoped that SEEP would reveal the radial-spatial distribution of transcriptional states across these spheroids.

We compelled PEO1 cells into compact spheroids and processed the spheroids via SEEP by sorting cells into 4 different layers based on the calibration parameters (surface, outer, inner, and core) ([Figures 2A](#) and [S1A](#)). Following sequencing, we aggregated the data from the 4 spheroid layers and identified highly variable genes to establish principal components (PCs) of transcriptional variation within the entire, aggregate spheroid sample ([Figure S4A](#)). Once processed, the scRNA-seq data revealed 7 clusters when grouped by functional state, each occupying a distinct proportion of each spheroid layer ([Figures 2B–2D](#)).

These data revealed multiple associations between selected transcriptional states and spheroid layers via a chi-square test of independence (e.g., strong association between transcriptional cluster #3 and the spheroid core) ([Figure 2C](#)). In addition, t-SNE dimensional reduction revealed overlap of transcriptional clusters with a positional layer of origin ([Figure 2D](#)). ORA, GSEA, and gene set variation analysis (GSVA) were used to infer transcriptional signatures and active pathways within individual layers and clusters.^{38,39} We observed variable enrichments of the Hallmark gene sets across layers ([Figures 2E](#) and [S5](#); [Tables S3A–S3N](#)). Clusters significantly associated with specific layers via the chi-square test of independence are highlighted on the heatmap. Enrichment in transcriptional programs associated with hypoxia, mTORC1 signaling, and the unfolded protein response (UPR) was found for cluster #3 which, as noted, had a strong association with spheroidal core cells. Cluster #4, which had a strong association with middle layer (inside/outside) cells,

showed strong enrichments for cell cycle pathways including G2/M checkpoints, E2F targets, and mitotic spindle. Histological analysis of intact PEO1 spheroids confirmed the enrichment of cycling cells in the outside layer using both Hematoxylin and Eosin (H&E) and Ki67 staining (fig. S6). Surface cells existed primarily in clusters #5 and #6 and possessed transcriptional programs associated with the epithelial to mesenchymal transition (EMT), suggesting a higher metastatic potential at spheroid surfaces. Surface cells also displayed immune-response-related programs including TNF α via NF κ B and interferon alpha and gamma (IFN α , IFN γ) responses. We utilized histological analysis to orthogonally confirm key SEEP outcomes. We first explored antileukoproteinase (secretory leukocyte protease inhibitor, SLPI), an emerging target of interest in metastatic cancer.⁴⁰ SLPI expression patterns visualized using a validated RNAscope probe and detailed image analysis of multiple spheroids demonstrated enriched expression on spheroid surface cells, consistent with the SEEP profiling (Figures S7A–S7D; Table S4). A second histological survey of the IFN γ responsive gene CXCL10 confirmed limited focal expression patterns as predicted by SEEP (Figures S7E and S7F). Taken together, these data support the accuracy of the SEEP method.

Spheroid cultures represent a moderately “controlled” cell culture format, and the presence of hypoxia and UPR characteristics in core cells supports long-held beliefs that core cells, separated from oxygen and nutrients, exist with relatively heightened oxidative and proteolytic stress. Enrichments in cell cycling signatures among outside cells suggest that growth and turnover are supported in this region of the spheroid culture at the time of analysis. The inside/outside spheroid regions also showed enrichment for gene targets of the transcriptional regulator YAP1 (Figure S8). YAP1 is regulated by the Hippo signaling pathway and is a key mediator of cell proliferation. Its activation within the intermediate spheroid layers sheds light on the aberrant and broader activation of YAP1 in cancer models and *in situ* disease.⁴¹ Interestingly, surface cells demonstrate an upregulation in apoptotic programs suggesting surface-related cell death events that would require replenishment from those layers directly underneath. Furthermore, the presence of EMT, inflammation, and immune response signatures found on the surface layer suggested that these cells, even when devoid of the host microenvironment, are primed for possible dissemination and broader exposure to the immune system.

SEEP-enabled mapping of transcriptional states in HGSOC organoids

To expand our initial SEEP findings to a more physiologically relevant disease model, we collected sphere-forming tumor cells from the ascitic fluids of HGSOC patients and grew them in Matrigel-media for several passages to establish HGSOC organoids. Tam-Seq revealed a > 99% allele fraction of TP53 mutated HGSOC within the sample, confirming a highly pure tumor fraction within the ascites-derived organoids. We gently bathed individual organoids in calcein AM for 55 min at 37°C (Figure 3A). For organoid samples, pre-calibration supported the collection of 3 concentric shells of equal volume (surface, middle, and core). As in the spheroid samples, we processed the organoids via SEEP and analyzed the transcriptional states of single cells on 3D radial positions. Organoid

samples separated into 7 natural clusters when grouped by gene expression (Figures 3B and S4C).

Using the chi-square test of independence, we found strong associations between distinct transcriptional clusters and organoid layers. Clusters #3, 4, and 5 were associated with surface cells, cluster #2 with middle cells, and cluster #1 with core cells (Figure 3C). t-SNE dimensional reduction visualizations revealed an overlap between transcriptional clusters and a positional layer of origin (Figure 3D). Furthermore, we used ORA, GSEA, and GSVA to reveal key transcriptional signatures enriched in each layer (Figures 3E and S9; Tables S5A–S5N). Several spatial transcriptional characteristics paralleled those found in the spheroid model. We found the expression of mTORC1 genes was enriched in the organoid core, and all organoid surface clusters included signatures involving inflammatory and immune-response-related programs. These included TNF α via NF κ B, IFN α /IFN γ responses, and estrogen responses (early and late). Compared with the spheroid model, however, there was less variance between the middle and core of the organoid model, with both regions showing enrichments among MYC targets V1 and cell cycle signatures (e.g., G2M checkpoint in core cells, E2F targets, and mitotic spindle genes in middle cells).

The organoids, unlike the spheroids, possibly retained elements of their *in vivo* origin. Due to the smaller size and increased complexity of this model, the clustering was performed using a 3-shell layering and each cluster was more homogenous relative to those determined for the spheroid samples. Nevertheless, there were coherent signature enrichments found for each positional domain. The organoid surface demonstrated compelling evidence of inflammation, stress, and immune responsiveness traits. This may relate to a cellular “memory” of exposure to the host microenvironment. However, the existence of similar characteristics in the spheroid models, where no such ‘memory’ would exist, suggests that the stress/inflammation/immune responsiveness traits were cell intrinsic and derived solely from surface positioning.

SEEP-enabled mapping of transcriptional states in HGSOC PDX biopsies

Finally, we applied SEEP to punch biopsies from HGSOC PDX models with the aim of understanding how transcriptional states vary according to surface accessibility in HGSOC *in vivo* tumor models. In addition to extending the applicability of SEEP and exploring the transcriptional architecture of a solid tumor grown from ascites-derived cells, we wanted to develop a method capable of preserving spatial information from samples collected in a clinical setting while adding negligible experimental burden. The collection and characterization of biopsy samples is a standard first-line approach to gain insight into abnormal or diseased tissues. When anatomically possible to collect, biopsies provide physical samples for histopathology and genomic analysis. Although histopathology retains positional relationships via visual analysis, the diagnostic breadth of data is limited. Conversely, the various genomic analyses done on biopsy tissues offer more information while sacrificing the spatial context of the sample. In hopes of combining elements of both methods, we evaluated the ability of SEEP to provide spatial context to transcriptional evaluations of punch biopsies.

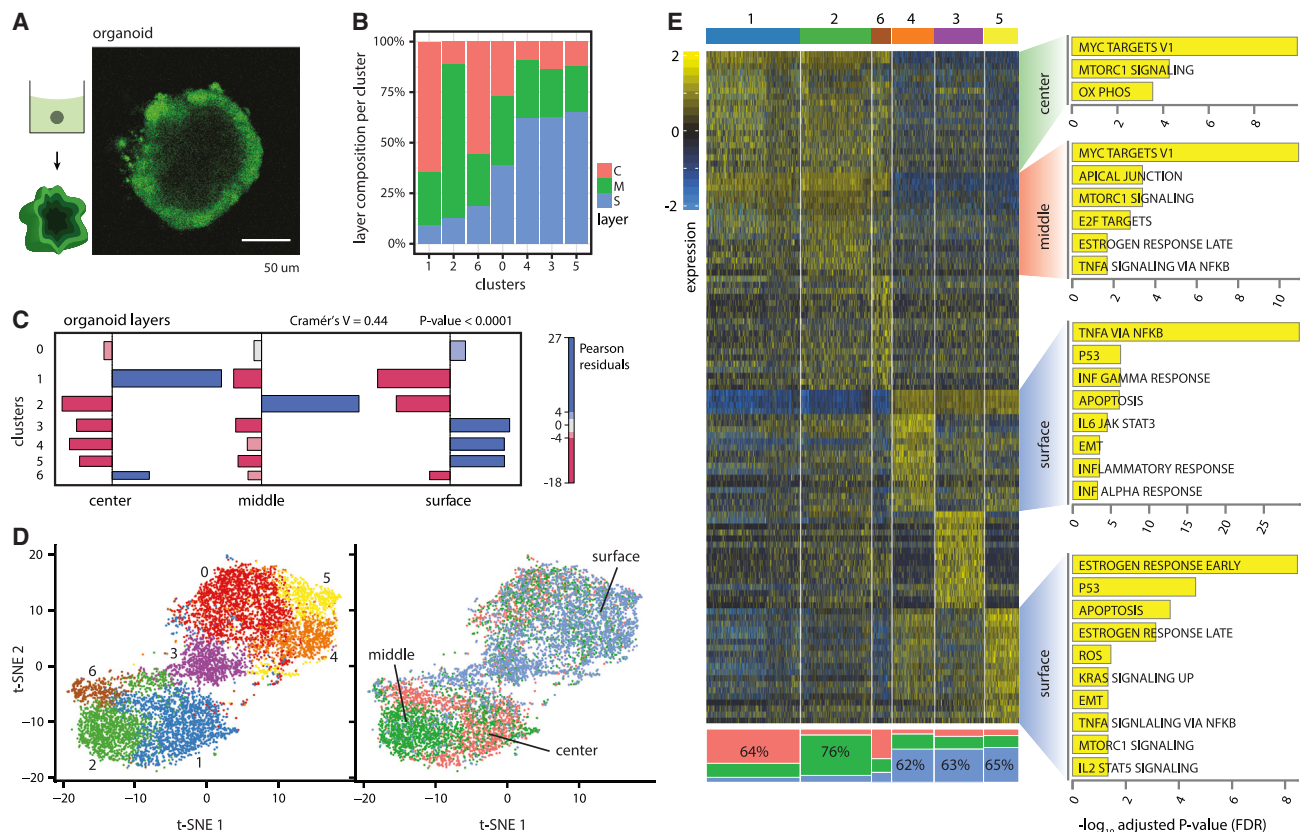


Figure 3. Positional analysis of HGSOC cultured organoids derived from patient ascites

(A) Cartoon schematic of the dye perfusion of a solitary organoid in a calcein AM bath and a confocal image of a spatially segmented HGSOC organoid. (B) Composition chart showing the seven transcriptionally defined cell clusters and the retrospective, layer-specific composition of each resolved cluster. Red, core; green, middle; blue, surface. (C) Bar-chart visualization of chi-square test of independence defined associations between gene expression clusters and organoid layers. Blue, over-representation; red, under-representation; gray, null hypothesis in the cross tabulation. Bar width illustrates relative cluster size ($n = 2,285$ center cells, $n = 2,819$ middle cells, and $n = 2,667$ surface cells). (D) t-SNE visualization of gene expression clusters (left) and radial cell position (right) ($n = 7,771$ cells). (E) Heatmap visualization of selected gene expression clusters and their layer-specific composition (by %). Significant transcriptional Hallmark gene signatures from GO are highlighted for clusters #1 (64% core cells), #2 (76% middle cells), #4 (62% surface cells), and #5 (65% surface cells). Color scale is linear. (Full accounting of Hallmark GO signatures can be found in [Tables S5A–S5G](#), and Hallmark signatures from GSEA can be found in [Figure S4](#) and [Tables S5H–S5N](#).) Scale bar, (A) 50 μm . GEO: GSE157299.

Several up-front challenges were encountered, including the reality that biopsies are often taken from large, asymmetric tissue masses. The subcutaneous HGSOC PDX tumors we used for the Punch-SEEP method were derived from ascites samples and revealed to have a >99% TP53 mutated allele fraction via TAM-Seq. Biopsies were taken as a single plug from the surface of carefully resected tumor samples to explore the radial-spatial distribution of cells states within solid tumors formed from ascites-derived cells. Simple immersion of biopsy samples in a calcein AM bath would expose core cells and surface cells simultaneously, failing to produce a dye gradient descriptive of cells' original positions. To mimic the radial-like diffusion of spheroid and organoid methods (i.e., surface exposure of the tissue model) we preloaded a calcein AM-saturated gelatin plug into a punch biopsy and allowed the plug to melt following contact with the tumor sample. Thus, exposure of the tumor surface to the calcein AM payload followed by diffusion through the sample

provided an imitation of the spheroid and organoid protocols. The dye loading was performed immediately prior to tissue harvesting, and after biopsy-punching and dye exposure the sample was incubated for 75 min at 37°C (Figure 4A). Gelatin plugs melted completely during the incubation period which was followed by tissue ejection from the punch biopsy, washing, dissociation, sorting, and sequencing.

The accuracy associated with the biopsy pre-calibration was coarser relative to the spheroid and organoid samples, likely due to inhomogeneities in vasculature and tissue density. Furthermore, because of necrosis within the tumors we analyzed, the number of cells that passed scRNA-seq filtering criteria was variable per layer. Thus, although we segmented tumors to contain equivalent numbers of cells per layer, the number of cells analyzed per layer was different. Specifically, cells characterized as surface more often passed scRNA-seq filtering and, therefore, appeared to be more prevalent than cells in other

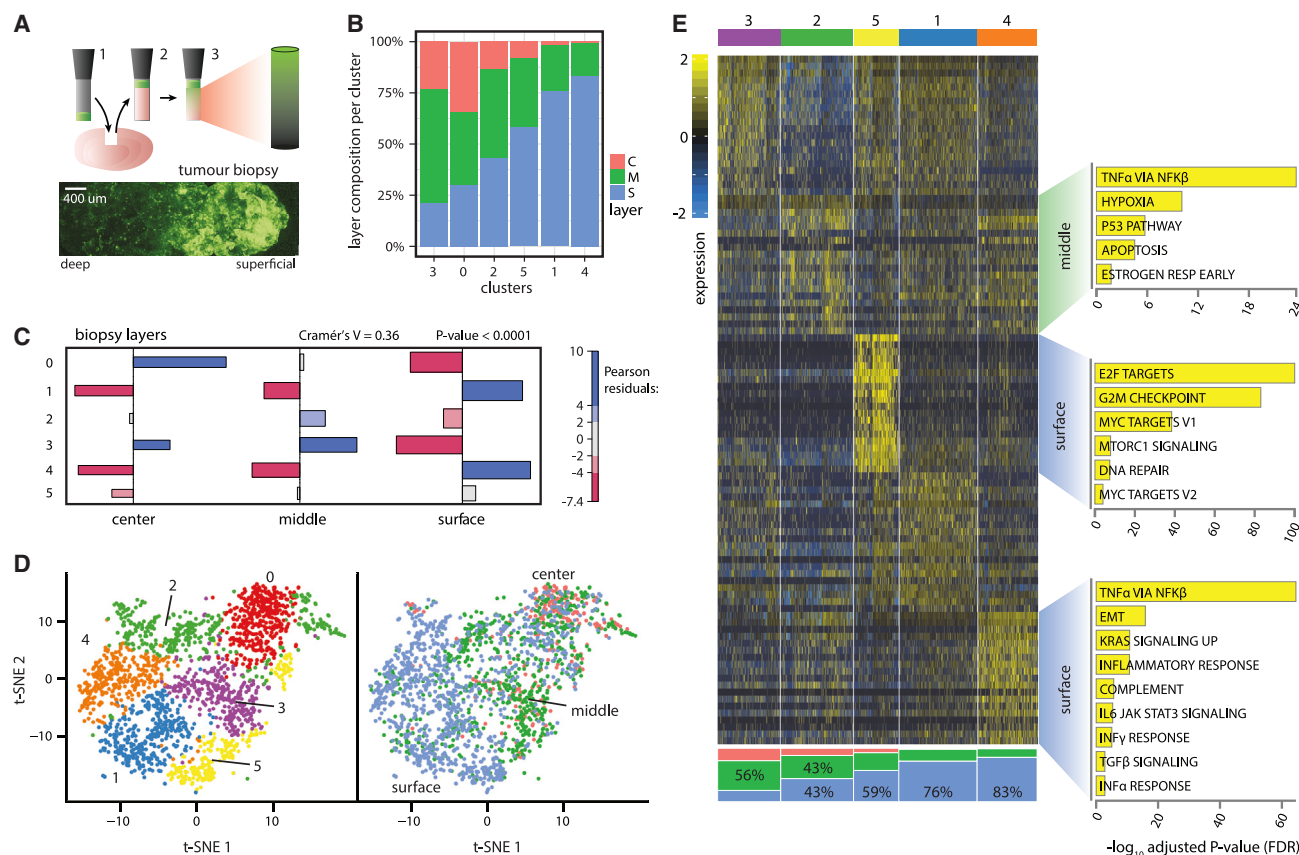


Figure 4. Positional analysis of HGSOC PDX-derived biopsy samples

(A) Cartoon schematic of the dye perfusion of a punch biopsy in a calcein AM bath and a confocal image of a spatially segmented HGSOC biopsy sample. (B) Composition chart showing the six transcriptionally defined cell clusters and the retrospective, layer-specific composition of each resolved cluster. Red, core; green, middle; blue, surface. (C) Bar-chart visualization of chi-square test of independence defined associations between gene expression clusters and biopsy layers. Blue, over-representation; red, under-representation; gray, null hypothesis in the cross tabulation. Bar width illustrates relative cluster size (n = 280 center cells, n = 677 middle cells, and n = 1,010 surface cells). (D) t-SNE visualization of gene expression clusters (left) and radial cell position (right) (n = 1,967 cells). (E) Heatmap visualization of selected gene expression clusters and their layer-specific composition (by %). Significant transcriptional Hallmark gene signatures from GO are highlighted for clusters #2 (43% middle cells), #5 (59% surface cells), and #4 (83% surface cells). Color scale is linear. (Full accounting of Hallmark GO signatures can be found in Tables S6A–S6G and Hallmark signatures from GSEA can be found in Figure S5 and Tables S6H–S6N.) Scale bar, (A) 400 μm. GEO: GSE157299.

layers. In addition, most oxygen-deprived cells collected from the center of the biopsies did not pass scRNA-seq filtering due to necrosis. This led to depleted hypoxia signatures within the biopsy cores. SEEP pre-calibration supported a three-shell segmentation (surface, middle, and core) that revealed 6 clusters with organized gene expression patterns (Figures 4B and S5C). Mouse cells and mouse-human doublets were identified and removed from downstream analysis. Using the chi-square test of independence, we found associations between distinct transcriptional clusters and the biopsy layers. Clusters #1, 4, and 5 were associated with surface cells, clusters #2 and 3 with middle cells, and clusters #0 and 3 with core cells (Figure 4C). Despite the coarse segmentation, t-SNE dimensional reduction visualizations revealed an overlap between transcriptional clusters and positional layer of origin (Figure 4D). As with the spheroids and organoids, we used ORA, GSEA, and GSVA to reveal key transcriptional signatures enriched in each layer (Figures 4E and S11; Tables S6A–S6L). Surface cluster #4 (83% surface

cells), like those in spheroid and organoid models, was enriched for EMT and inflammatory and immune-response-related programs, including TNFα via NFκB, IFNα/IFNγ responses, and IL6 JAK/STAT3 signaling gene sets.

The Punch-SEEP data, derived from the most sophisticated and relevant HGSOC model, was also the most complex to interpret due to inhomogeneities in bin size. Surface-associated clusters #1 and 4 possessed similar traits to the surface cells noted in both the spheroid and organoid cultures. These included inflammatory-, immune-, and EMT-related gene sets. Genes associated with apoptosis also emerged on the surface of the biopsy sample akin to both the spheroid and organoid surfaces.

Position-defined transcriptional states conserved across 3D HGSOC models

To examine transcriptional states that were consistently associated with a radial position in all three models (spheroid, organoid, and PDX biopsy), we compared segmented gene expression

data from all layers in each model directly, irrespective of expression-defined clustering. We hoped to explore how the spatial distribution of functional states might be retained across similar positional microenvironments in different settings. This analysis confirmed the presence of individual genes and gene signatures consistently enriched in the surface and core layers of all the models analyzed. A comparison of the log-normalized average gene expression values between surface and core cells from each model revealed several similarities (Figures 5A–5C, and S10; Table S7). Surface cells in all models possessed upregulation of genes associated with immune and inflammation-related signatures including Hallmark gene sets for TNF α via NF κ B and EMT. There was also a strong overlap in IFN α /IFN γ leading edge genes on the surfaces of the spheroid and organoid cells (Figures 2E, 3E, S5, S9, and S10). Although core cells had fewer coherent trends, there was consistent increased expression of several genes involved in mTORC1 and KRAS signaling (from spheroid and organoid models, respectively) and genes involved in the cell cycle process (Figures 2E, 3E, S5, and S9). Although core-enrichment for a hypoxia gene set was noted in the spheroid and organoid models only, we found upregulation of selected hypoxia-related genes in the core cells of each model.

A more detailed examination of EMT-related genes demonstrated a remarkable anti-correlation of epithelial and mesenchymal markers, as expected. This is most prominently demonstrated by the model-agnostic expression of the epithelial-related E-cadherin (*CDH1*) in core cells relative to the expression of the mesenchymal-related N-cadherin (*CDH2*) on surface cells (Figures 5D–5F). This trend was maintained for other epithelial-related genes [e.g., desmoplakin (*DSP*) and tight-junction protein 1 (*TJP1*)] and mesenchymal-related genes [e.g., matrix metalloproteinases 3 and 9 (*MMP3/9*) and zinc-finger protein *SNAI1* (*SNAI1*)]. The biopsy model demonstrated an expanded breadth of mesenchymal-related genes on surface cells including fibronectin 1 (*FN1*), vimentin (*VIM*) and twist-related protein 1 (*Twist1*). We hypothesized that EMT marker genes were enriched across tissue surfaces in response to the stress and inflammatory signatures found to be upregulated on these surfaces. This hypothesis was supported by *in situ* hybridization data including the aforementioned surface layer enrichment of antileukoproteinase (*SLPI*) (Figure S7).⁴⁰

To better understand the pan-model relationship of enriched gene sets, we employed single-cell GSVA across all samples. Useful for visualizing small changes in gene expression but significant changes in gene set expression, GSVA allowed us to explore cell states independently from cluster identity and explore correlative relationships among gene sets (Figures 5G–5I, S12, S13, and S14). Several trends were revealed across each HGSOC model including correlations between IFN γ and EMT Hallmark gene sets and between inflammatory response and apical surface Hallmark gene sets. Examining gene set correlations across MSigDB gene sets, we found a variety of strong correlations across single cells.^{42,43} Of note, in the spheroid model, we found an inverse correlation ($R = -0.63$) between Cell Cycle and IFN γ Response gene sets (Figure S12B), a relationship found to be demonstrated in the slowly cycling surface-residing spheroid cells that expressed many IFN α and IFN γ response genes. In the organoid model, we found a positive correlation ($R = 0.87$) between Cellular Response to Stress and

Positive Regulation of EMT gene sets (Figure S13B). This relationship was found on surface-residing cells across all three models with cells co-expressing both stress response genes and EMT-associated genes. In the PDX-biopsy models, we found a positive correlation ($R = 0.58$) between IL6 JAK/STAT3 signaling and TNF α via NF- κ B signaling gene sets (Figure S14B). This relationship was also noted in a survey of disseminated tumor cell clusters in the ascites of HGSOC patients.⁴⁴ In the PDX-biopsy models, we also found a positive correlation ($R = 0.60$) between the P53 Pathway and a Positive Regulation of EMT. As P53 was ubiquitously mutated across all our model systems, this relationship suggests a possible connection between aberrant P53 signaling (known to cause broad genomic instability in HGSOC) and EMT.⁴⁵ GSVA also confirmed that inflammatory response, TNF α via NF κ B, and EMT signatures were enriched in surface cells, whereas mTORC1 and angiogenesis signatures were enriched in core cells (Figures 5J–5L). These correlations allowed us to explore the relatedness of gene expression signatures found within various regions of the tumor models. Although we did not perform experiments to reveal the causal nature of these correlations, we hoped that this analysis could generate hypotheses about the local effect of upregulated gene sets on HGSOC tumors.

Next, we integrated each dataset to further extract spatially resolved expression features conserved across the model systems. We utilized the Seurat integration algorithm^{8,26} to identify pair-wise mutual nearest neighbors (MNN) in a low-dimensional space from canonical correlation analysis (CCA). We then computed the anchor transformation matrix to align the spheroid, organoid, and biopsy SEEP datasets and subsequently embedded the cells into a common space through UMAP.³¹ The 2D UMAP projections of the corresponding integrated model colored according to data source or layer source are shown in Figures 6A, 6B, and S15A. Applying the multiscale analysis methodology “Differentially Abundant-seq” (DA-seq) to this integrated model revealed a subcluster of cells with differential abundance between surface and center compartments (Figures 6C and S15A).⁴⁶ To reveal gene expression patterns in the most regionally distinct cell populations (surface and center cells), we isolated layer-specific differentially abundant (DA) subpopulations with a DA-seq score >0.8 and used the Seurat gene scoring method to identify enriched transcriptional signatures.²⁶

These analyses revealed three coherent surface layer cell populations and a single center layer cell population from the joined SEEP scRNA-seq models (spheroid, organoid, and biopsy data collections) (Figures 6C, 6D, and S15B). Hallmark pathway analysis highlighted a surface subpopulation with strong enrichment for IFN response and TNF α signaling via NF- κ B (surface cluster #1) and a second subpopulation with strong enrichment in cell cycle signatures (surface cluster #2). The third surface layer subpopulation showed enrichment for Hallmark Apical Junction, EMT, and estrogen responses. The center subpopulation was less resolved with coherent enrichments for only Oxidative Phosphorylation and MYC targets V1 (Figures 6D and S15B; Table S8).

SEEP-derived signatures are a conserved feature of ascites isolated from HGSOC cells

Finally, we sought to compare the SEEP-derived spatially dependent signatures with published observations from primary

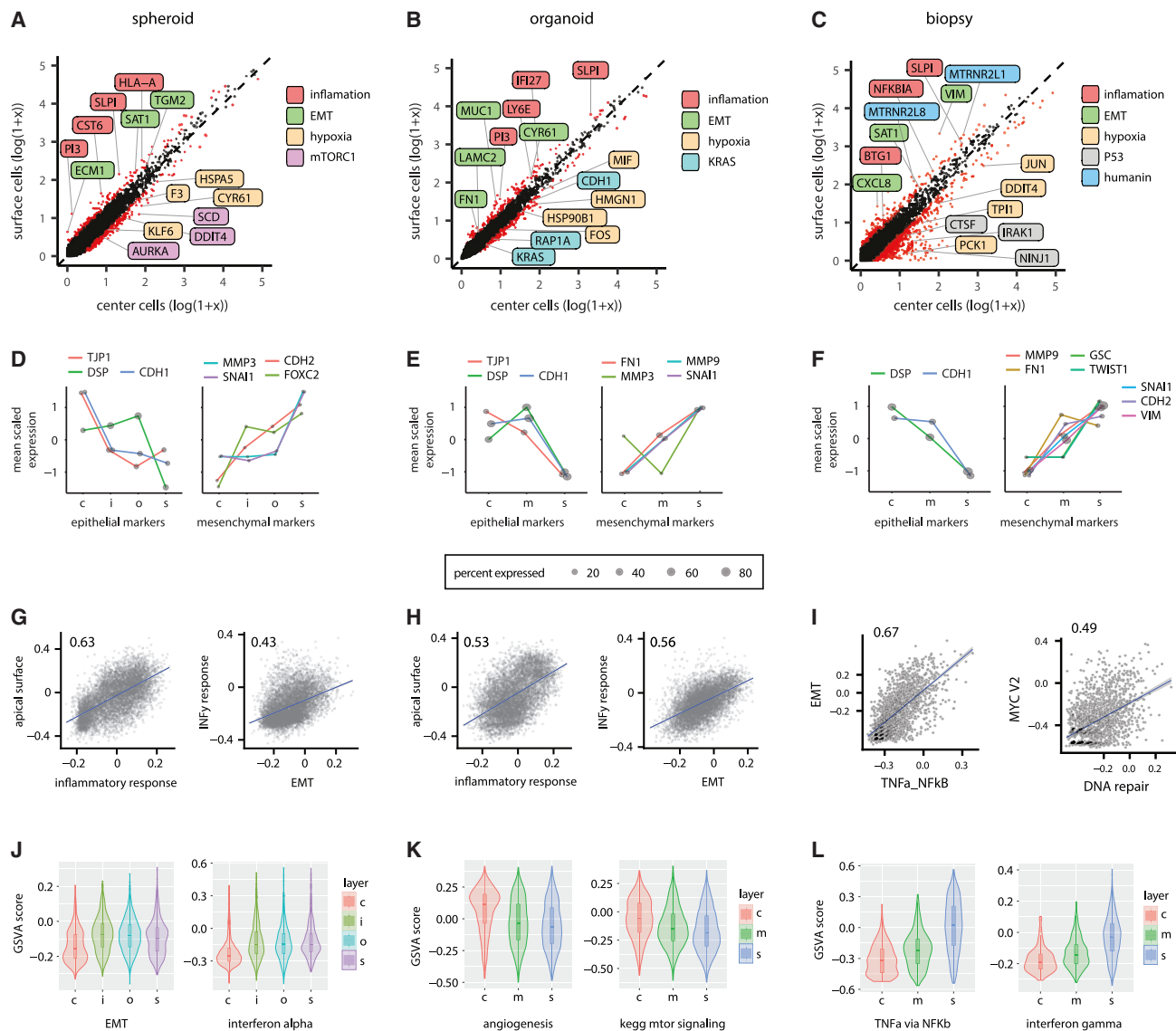
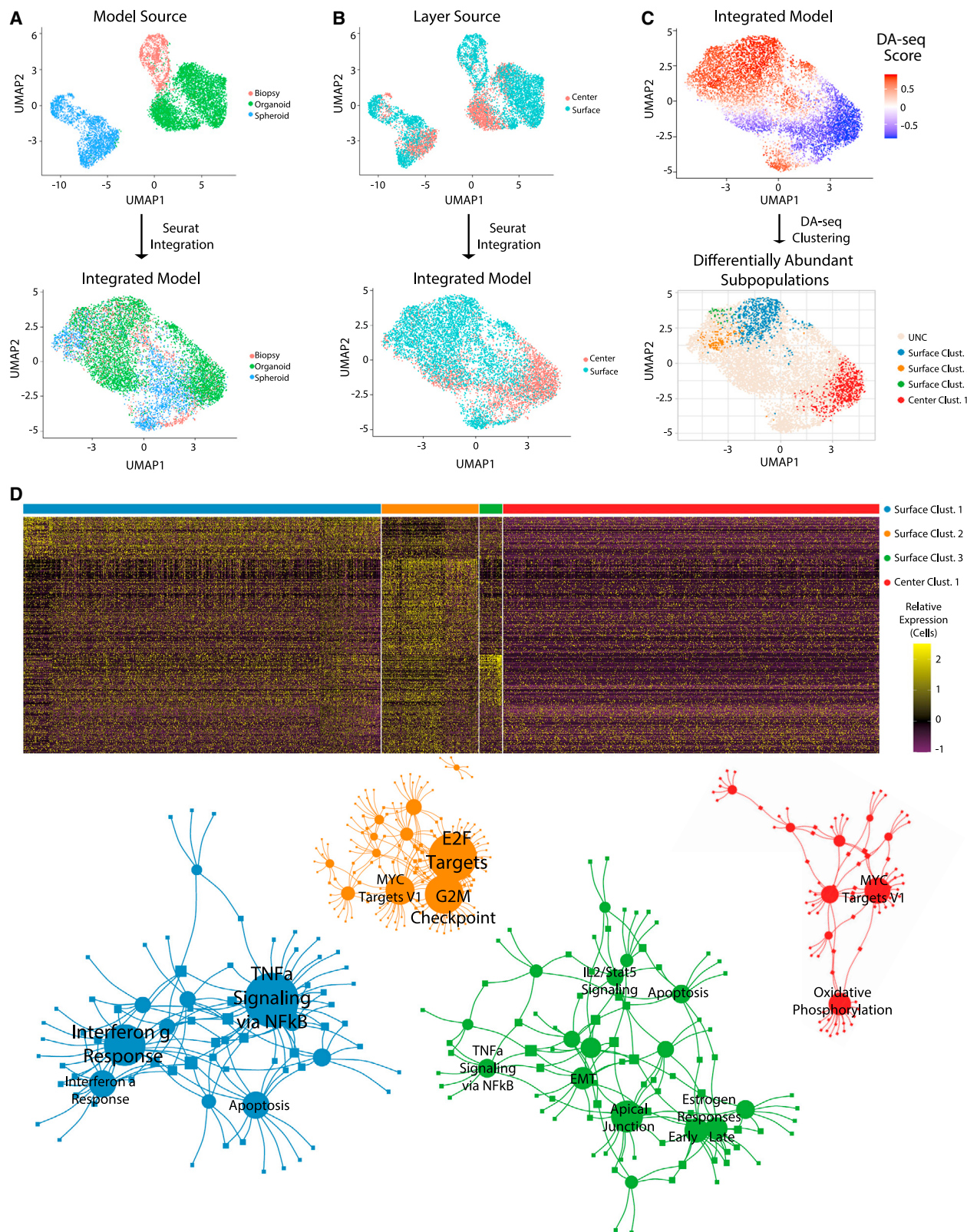


Figure 5. Consistency in regional transcriptomics profiles across HGSC models

(A) Correlation of log-normalized gene expression traits between surface and core cells in HGSC spheroid models.
 (B) Correlation of log-normalized gene expression traits between surface and core cells in HGSC organoid models.
 (C) Correlation of log-normalized gene expression traits between surface and core cells in biopsy of HGSC PDX models.
 (D) Expression trends for epithelial and mesenchymal marker genes across 3D layers in HGSC spheroid models.
 (E) Expression trends for epithelial and mesenchymal marker genes across 3D layers in HGSC organoid models.
 (F) Expression trends for epithelial and mesenchymal marker genes across 3D layers in biopsy of HGSC PDX models.
 (G) Correlation of key gene expression signatures across all cells using GSVA in HGSC spheroid models ($n = 7,908$ cells).
 (H) Correlation of key gene expression signatures across all cells using GSVA in HGSC organoid models ($n = 7,771$ cells).
 (I) Correlation of key gene expression signatures across all cells using GSVA in biopsy of HGSC PDX models ($n = 1,967$ cells).
 (J) Violin plots highlighting the distribution of all cells binned by 3D layer using GSVA signatures for EMT and IFN α response in HGSC spheroid models ($n = 1,178$ center cells, $n = 2,471$ inside cells, $n = 2,736$ outside cells, and $n = 2,667$ surface cells).
 (K) Violin plots highlighting the distribution of all cells binned by 3D layer using GSVA signatures for angiogenesis and mTOR signaling in HGSC organoid models ($n = 2,285$ center cells, $n = 2,819$ middle cells, and $n = 2,667$ surface cells).
 (L) Violin plots highlighting the distribution of all cells binned by 3D layer using GSVA signatures for TNF α signaling via NF κ B and IFN γ response in biopsy of HGSC PDX models ($n = 280$ center cells, $n = 677$ middle cells, and $n = 1,010$ surface cells).

HGSC samples. We mapped our signatures to scRNA-seq profiles of primary HGSC ascites samples to explore if these spatial profiles were present in primary patient ascites. This

served to both validate our consensus signatures and to impute spatial information onto a spatially naïve HGSC scRNA-seq dataset. To do this, we utilized a recent report by Izar et al. detailing



(legend on next page)

the scRNA-seq signatures of malignant and non-malignant cell clusters collected from the ascites of HGSOc patients.⁴⁴ In the report, a total of 22 ascites specimens were isolated from 11 patients. scRNA-seq analysis revealed both intra- and inter-tumor heterogeneity with subpopulations of cells enriched for immunoreactive and inflammatory pathway signatures. Spatial information was not collected in this study. To map and compare the pan-model SEEP signatures with this dataset, we refined the integrated HGSOc model to retain only highly DA subpopulations (Figures 6C and 7A) and performed reference-based single-cell data integration. This integration mapped our DA-seq subpopulations to the primary HGSOc ascites dataset reported by Izar et al. We integrated our subpopulation signatures onto Izar's "total ascites" dataset containing CD45⁺ depleted ascites cells (Figure 7B) and onto their "malignant ascites" dataset containing stringently enriched EPCAM⁺ and CD24⁺ cancer cells (Figure 7C).

Integration identified populations of cells within the total ascites and malignant ascites datasets which mapped to our subpopulations. Pathway analysis of these cells mirrored the SEEP pathway enrichment outcomes (Figure 7D). The strongest alignments were noted for the surface cluster #1 population that was strongly enriched for both IFN response and TNF α signaling via NF- κ B signatures in both the total and malignant ascites datasets. Surface cluster #2 also revealed a well-defined population of cells that were enriched for cell cycling signatures in both datasets. The alignment for surface cluster #3 revealed a small population of cells enriched for EMT and apical surface signatures, whereas the center cluster demonstrated a large population of cells sporadically enriched for oxidative phosphorylation, MYC targets, and MTORC1 signaling. A more granular comparative assessment of the top 30 gene markers from the DA-seq pan-model SEEP signatures reinforced a strong alignment with the primary ascites HGSOc cell populations (Figure 7E). A comparison of the expression patterns for TNF α induced protein 2 (TNFAIP2) in the four pan-model SEEP clusters and the total ascites and malignant ascites cell populations illustrates the alignment of these data. Each dataset shows pronounced expression in surface cluster #1 with progressively lower expression in surface clusters #3, #2, and center cluster #1 (Figure 7F). TNFAIP2 is a bona fide TNF α inducible gene with demonstrated associations to cell migration and metastasis. Utilizing the reference UMAP projections to visualize subpopulations associated with the Hallmark G2M-Checkpoint and E2F-Targets signatures further demonstrated the strong overlap between the SEEP

and the primary ascites datasets (Figures S16A and S16B). The comparative expression patterns for cell division marker ZW10 interacting kinetochore protein (ZWINT) are illustrative of this overlap with strong enrichment within surface cluster #2 in the SEEP pan-model dataset and both the total ascites and malignant ascites cell populations (Figure S16C). Extracting the Hallmark signature for IFN γ response also exhibited a prominent overlap between these datasets (Figure S16D).

Given the significant enrichment of IFN response and other immunoreactive pathways in our 3D HGSOc models and the primary HGSOc ascites samples, we sought to understand whether immunoreactive expression was generally enriched in 3D models or a unique product of HGSOc models. To interrogate this question, we explored the well-studied IFN γ gene product indoleamine 2,3-dioxygenase 1 (IDO1) using qPCR analysis of monolayer and spheroid models of HGSOc (PEO1, CIOV1, and CIOV3), melanoma (A375), and pancreatic cancer (PANC1). These data showed pronounced IDO1 expression in each of the 3D HGSOc spheroid models, whereas the melanoma and pancreatic cell models did not display this trend (Figures S16E). In addition, none of the monolayer cultures displayed pronounced IDO1 expression. These data suggest that immunoreactive cell populations may be a unique feature of 3D models of HGSOc and *in situ* disease and that these signatures are driven by selected cell subpopulations that are enriched on the surface of the 3D cell/tissue clusters. Collectively, these results show that recognizable cell populations within the primary ascites samples reasonably align with SEEP-revealed spatially distinct cell populations across model systems.

DISCUSSION

Modeling cancer using cultured cells or *in vivo* models offers a non-invasive approach to study these diseases and gain insights into their origin, progression, and treatment. Genomic characterizations, including single-cell sequencing technologies, provide a powerful tool to understand both *in vitro* and *in vivo* models and compare them with human samples. It remains challenging, however, to associate the transcriptional, proteomic, and metabolic identity of cells with their contextual environment. To overcome this challenge, various methods and technologies have emerged, such as physically selecting cells from defined coordinates within tissue sections, multiplexed fluorescence ISH,^{7,13} targeted *in situ* sequencing of RNA fragments,⁴⁹ spatial reconstruction of scRNA-seq data from ISH patterns,⁸ and multimodal

Figure 6. Conservation of positional transcriptomics profiles across HGSOc models

(A and B) UMAP projections of cultured PEO1 spheroids, patient-derived organoids, and PDX-derived biopsy samples before (top) and after (bottom) scRNA-seq integration by data source⁴⁷. Cell embeddings were color coded by data source (A) and spatial segmentation (B). Total cells from all models (center and surface): n = 8,805.

(C) DA cell subpopulations identified in the integrated HGSOc model. Top: UMAP embeddings of cells were colored by DA-seq score⁴⁶; small/large values indicate a high abundance of cells from the center/surface layers, respectively. Bottom: layer-specific DA subpopulations were detected by clustering cells with absolute DA-seq score greater than 0.8. Total cells from all models in DA regions: n = 2,105 (n = 1,019 center cells and n = 1,086 surface cells). Color scale is linear.

(D) Heatmap of layer-specific genes conserved across the three HGSOc models. The markers were selected using the *FindConservedMarkers* method⁴⁷ and gene functional over-representation⁴⁸ in the MSigDB Hallmark collection (statistical test results are included in Table S8). Heatmap dimension: 364 genes $\times$ 2,002 cells. Color scale is linear.

(E) Gene-pathway network of Hallmark gene sets enriched in the integrated model (adjusted p value < 0.05). Size of the nodes depends on the number of conserved markers overlapping a pathway. Gene set nodes are labeled circles, and gene nodes are squares without label. Nodes and edges are colored according to the DA subpopulation clusters: blue, surface cluster 1; orange, surface cluster 2; green, surface cluster 3; red, center cluster 1. Details of the enrichment analysis are included in Table S8.

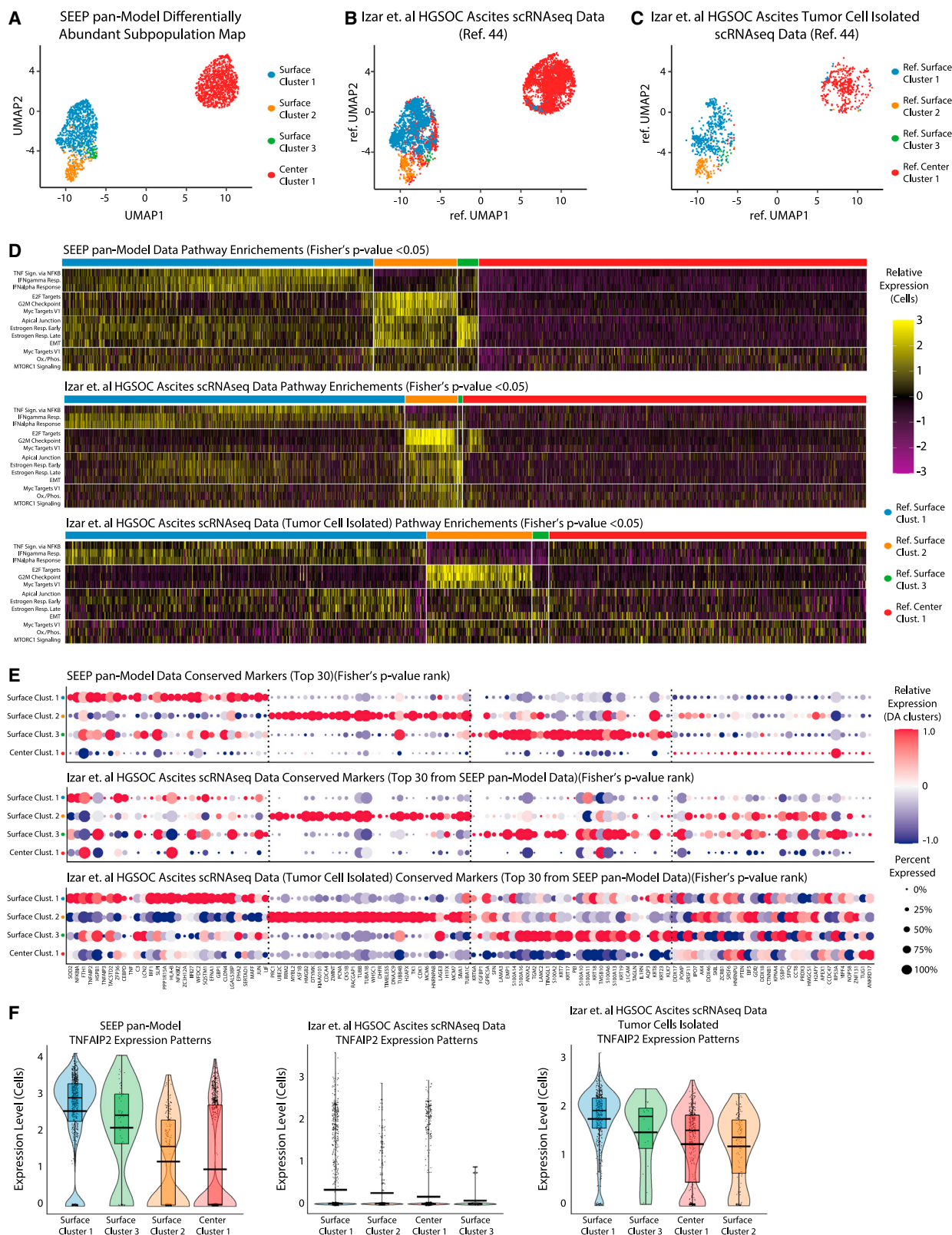


Figure 7. Conservation of integrated HGSOC spatial models in primary HGSOC ascites samples

(A–C) Reference-based transfer was used to map subpopulations from our integrated HGSOC model ($n = 2,083$; 812 surface cluster 1, 218 surface cluster 2, 42 surface cluster 3, 1,011 center cluster) (A) to primary single-cell data from total ascites collections ($n = 7,144$; 1,878 surface cluster 1, 247 surface cluster 2, 20

(legend continued on next page)

spatial profiling.⁵⁰ Each of these pioneering techniques have yielded insights into the models they explore.

Here, we endeavored to add to these approaches using a ‘paint, sort, sequence’ method that combines the power of scRNA-seq with positional stratification based on environmental accessibility. Our method, called SEEP, is enabled by a cell penetrant, pro-fluorescent dye (calcein AM) and a series of pre-calibration steps that accurately and robustly define positional layering of both *in vitro* culture models and *in vivo* biopsy samples. Utilizing traditional FACS methods, volumetrically defined cell layers can be collected and analyzed via scRNA-seq.

We used SEEP to characterize the positional/transcriptional heterogeneity of cells in three distinct HGSOc models (spheroids, organoids, and PDX tumor biopsies). Position-associated transcriptional signatures were examined for each model individually and through comparative analyses across model types. The spheroid model yielded the greatest level of accuracy in terms of spatial resolution. A total of four distinct layers (surface, outside, inside, and core) were captured and each possessed a unique transcriptional identity. Several previously established concepts in spheroid biology were confirmed including the hypoxic nature of core cells. Cells in the middle domains (outside and inside) were noted to be enriched for multiple gene sets related to cell proliferation. Finally, surface cell identity was enriched for transcriptional traits associated with inflammation (TNF α signaling via NF κ B), immune responsiveness (IFN α , IFN γ), and genes associated with EMT. Inflammation has been shown to be beneficial for tumor seeding, invasion, and metastasis in ovarian and other cancers but, to our knowledge, has not been reported to be associated with tumor surfaces.^{51,52} These features were conserved among surface cells on an organoid and biopsy model, suggesting that the surface microenvironment alone is sufficient to drive these characteristics in an intrinsic fashion. As cells from uniform spheroids cultures are driven into divergent functional states based on cell position, it seems likely that intratumoral position can determine cell state, and spatial heterogeneity across tumors cannot be fully explained by the migration of pre-determined cell types to preferred positions.

Transcriptional signatures enriched in specific layers were found to be correlated with each other including associations of EMT with IFN γ responsive genes and apical surfaces with inflammatory response genes. Analysis of EMT-related genes highlighted the epithelial-like nature of core cells and the mesenchymal-like nature of surface cells across each model. This suggests that surface cells, with primed transcriptional states and fewer physical barriers to dissemination, may play critical roles in metastasis. Nonetheless, it is remarkable that the surface cells

of HGSOc models surveyed in this study natively adopt this identity, particularly in spheroid models where there are no host-related factors involved. Pan-model analysis added additional validation that cell subpopulations exist in all three models including distinct surface subpopulations of either immunoreactive cells or cells undergoing cycling and division. Many of the pan-model transcriptional signatures found in this study broadly align with previously reported bulk-sequencing and pathway characteristics.^{53,54} Izar et al. recently reported on a collection of scRNA-seq data from ascites-derived HGSOc cells that revealed a widespread inflammatory program among the most common signatures associated with captured malignant cells.⁴⁴ We utilized this study as a comparator to validate the accuracy of the SEEP method, although also probing further the inflammatory and immunoreactive nature of HGSOc cell clusters. Our comparative analyses revealed distinct cell populations in the captured ascites HGSOc cells similar to the pan-model cell populations identified using the SEEP method. The results suggest the presence of inflammatory/immunoreactive, cell cycle machinery, and cellular differentiation and organization pathways on the surfaces of HGSOc tumors from patient ascites samples. These data validate the SEEP method and demonstrate the ability of complex 3D models to replicate complex disease phenotypes, such as intratumoral heterogeneity, although also highlighting the existence of distinct druggable phenotypes across different cell subpopulations.

We believe that SEEP will be applicable to a variety of different models not only as a stand-alone method but also as a complement to previously described techniques that relate transcriptional identity to cell position. Furthermore, as SEEP does not rely on transcriptional variation to infer position, it will complement and refine computational methods for defining spatial genomic relationships. Unlike ISH methods, SEEP retains a link between cDNA and cell barcodes enabling genotyping or the identification of gene fusions using long-read sequencing. In addition, the method can be performed at scale and is compatible with high-throughput pharmacological testing (e.g., in conjunction with CITE-seq to define screening conditions). Using existing spatial methods, these experiments are arduous to conduct across whole structures and many conditions. SEEP does require extra steps in the binning, sorting, and indexing stage. However, because these steps do not require niche equipment or expertise, the SEEP method lowers the barrier for non-specialized labs to acquire spatially resolved scRNA-seq data. Thus, the SEEP method can enable studies designed to identify variation in cell states as a function of environmental accessibility and adaptation to perturbation as influenced by the tissue shape. However, it is important to note that the

surface cluster 3, 4,999 center cluster) (B) and malignant-enriched ascites collections (n = 1,015; 463 surface cluster 1, 128 surface cluster 2, 20 surface cluster 3, 404 center cluster) (C) from HGSOc patients. From each data source, top 99% cells with the strongest contribution to the first UMAP dimension were used. In addition, the ascites cells with maximal transfer score greater than 0.5, and a difference from the next largest score of at least 0.25 were selected. Data were embedded and visualized across two UMAP dimensions, and the top 99% of cells with the highest contribution to the first UMAP dimension were used.

(D) Signature scores (69) calculated for the top Hallmark gene sets enriched in the integrated HGSOc model (p value rank ≤ 3). Up to 2,000 top cells from each cluster were used (the top 95% of cells with the strongest contribution to the first UMAP dimension). The signature scores displayed are scaled and centered to zero mean and one SD across cells. Color scale is linear. (E) Dot plot showing the relative expression of the top 30 markers (p value rank) from the DA-seq pan-model SEEP signatures. Expression is shown across the integrated SEEP data, the total ascites data, and the malignant ascites data. Color scale is linear.

(F) Violin plots overlaid with boxplots and averages (horizontal segments) of TNFAIP2 normalized expression from the integrated SEEP data, the total ascites data, and the malignant ascites data.

SEEP method does not reveal how transcriptional heterogeneity is influenced solely by tissue shape. SEEP's intrinsic ability to permanently stain cells is advantageous for optimizing dissociation methods. Dissociation of complex tissue types is a limitation for scRNA-seq experiments, and methods with stable spatial demarcation strategies will be advantageous for optimizing the dissociation protocol of a given tissue. The use of a multicolor system may be able to extend the scope of SEEP where pulsed additions of dye could allow for finer radial segmentation. Grid-segmentation along cartesian coordinates may be possible by two-dimensional staining (e.g., exposing a 3D tissue to dye fronts from a perpendicular axis) and/or the embedding of point-source dyes throughout a tissue (e.g., embedding fluorophore releasing hydrogels throughout a 3D tissue). The diffusion characteristics of calcein AM dye may be limiting for certain applications. Fortunately, a multitude of bio-orthogonal dyes exists that greatly expand the range of the SEEP method when used creatively to solve distinctive experimental goals.⁵⁵ For instance, the coupling of the SEEP method with the field of "activity-based diagnostics" may greatly expand the data capture capacity of the probes used in that emerging research platform.⁵⁶ Exchanging colored dyes for synthetic DNA oligos could be used to segment 3D tissues and would eliminate the need for FACS as synthetic oligo concentrations could be quantified during sequencing. In addition, dissociating and reforming spheroids stained via SEEP may add insights into whether surface-primed cells would retain their previous functional state, adopt a new state, or migrate to the "new" spheroid surface. Expanding the SEEP methods into more complex tissue and/or whole tumor systems will be complicated by extensive vascularization networks that will alter dye accessibility. This reality may, however, offer a means to correlate phenotypes associated with oxygen, nutrient, or drug accessibility to cell transcriptional identity and response following perturbation. Despite the inherent limitations discussed above, the feasibility and flexibility of the SEEP method can enable the creative expansion of spatial scRNA-seq queries. For appropriate systems, the accuracy and robust nature of SEEP should enable a variety of explorations and reveal insights into environmental accessibility, spatial identity, heterogeneity, and the unique relationship of single cells with the broader system.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
 - Lead contact
 - Materials availability
 - Data and code availability
- **METHOD DETAILS**
 - Cell culture and tissue handling
 - SEEP calibration staining
 - Imaging dye penetration
 - SEEP controls for examining the effect of staining and sorting on transcription
 - SEEP measurement, dissociation, and FACS

- Indrop deployment for scRNA-sequencing
- Data preprocessing and filtering
- Immunohistochemical evaluation of formalin-fixed paraffin-embedded spheroids (FFPE)
- In situ hybridization-based detection of SLPI and CXCL10
- Spheroid image analysis
- Data analysis and visualization
- Patient HGSO Mapping
- GSEA analysis
- GSVA analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cels.2023.05.003>.

ACKNOWLEDGMENTS

This work was supported by the Division of Preclinical Innovation, National Center for Advancing Translational Sciences, and the Center for Cancer Research, National Cancer Institute. The work at Harvard was supported by the NSF (DMR-1708729) and through the Harvard MRSEC (DMR-2011754). The work at Cambridge was supported by the BBSRC, the Newman Foundation, the Wellcome Trust, and the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) through the ERC grant PhysProt (agreement no. 337969). D.B.M. was supported by the National Institute of Health Oxford-Cambridge Scholars Program and the Certara Biomedical Research Scholarship.

AUTHOR CONTRIBUTIONS

Conceptualization: D.B.M., M.C., D.A.W., T.P.J.K., and C.J.T.; methodology: D.B.M., A.M.M., M.C., J.D.J., C.M., M.V., S.B., D.A.W., J.D.B., J.B., T.P.J.K., and C.J.T.; investigations: D.B.M., A.M.M., M.C., J.D.J., C.M., M.V., D.B., T.D.-H., T.V., S.P., and S.B.; data curation: D.B.M., A.M.M., T.V., and M.C.; writing – original draft preparation: D.B.M., M.C., T.P.J.K., and C.J.T.; writing – review and editing: D.B.M., A.M.M., M.C., J.D.J., C.M., M.V., S.P., S.B., D.A.W., J.D.B., J.B., T.P.J.K., and C.J.T.; visualizations: D.B.M., A.M.M., M.C., T.P.J.K., and C.J.T.; supervision: M.C., D.A.W., J.D.B., J.B., T.P.J.K., and C.J.T.; project administration: T.P.J.K. and C.J.T.; funding acquisition: D.B.M., D.A.W., T.P.J.K., and C.J.T.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

While citing references scientifically relevant for this work, we also actively worked to promote gender balance in our reference list. We support inclusive, diverse, and equitable conduct of research.

Received: December 29, 2021

Revised: January 22, 2023

Accepted: May 17, 2023

Published: June 21, 2023

REFERENCES

1. Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 161, 1202–1214. <https://doi.org/10.1016/j.cell.2015.05.002>.
2. Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017).

- Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* 8, 14049. <https://doi.org/10.1038/ncomms14049>.
3. Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A., and Kirschner, M.W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* 161, 1187–1201. <https://doi.org/10.1016/j.cell.2015.04.044>.
4. Gerlinger, M., Rowan, A.J., Horswell, S., Math, M., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., et al. (2012). Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* 366, 883–892. <https://doi.org/10.1056/NEJMoa1113205>.
5. Bissell, M.J., Kenny, P.A., and Radisky, D.C. (2005). Microenvironmental regulators of tissue structure and function also regulate tumor induction and progression: the role of extracellular matrix and its degrading enzymes. *Cold Spring Harb. Symp. Quant. Biol.* 70, 343–356. <https://doi.org/10.1101/sqb.2005.70.013>.
6. Hayat, M.A. (2004). *Handbook of Immunohistochemistry and In Situ Hybridization of Human Carcinomas, First Edition* (Elsevier Academic Press).
7. Achim, K., Pettit, J.B., Saraiva, L.R., Gavriouchkina, D., Larsson, T., Arendt, D., and Marioni, J.C. (2015). High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nat. Biotechnol.* 33, 503–509.
8. Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* 33, 495–502. <https://doi.org/10.1038/nbt.3192>.
9. Vickovic, S., Eraslan, G., Salmén, F., Klughammer, J., Stenbeck, L., Schapiro, D., Åijö, T., Bonneau, R., Bergenstråhle, L., Navarro, J.F., et al. (2019). High-definition spatial transcriptomics for in situ tissue profiling. *Nat. Methods* 16, 987–990. <https://doi.org/10.1038/s41592-019-0548-y>.
10. Rodriques, S.G., Stickels, R.R., Goeva, A., Martin, C.A., Murray, E., Vanderburg, C.R., Welch, J., Chen, L.M., Chen, F., and Macosko, E.Z. (2019). Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. *Science* 363, 1463–1467. <https://doi.org/10.1126/science.aaw1219>.
11. Ståhl, P.L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J.F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J.O., Huss, M., et al. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* 353, 78–82. <https://doi.org/10.1126/science.aaf2403>.
12. Lee, J.H., Daugherty, E.R., Scheiman, J., Kalhor, R., Yang, J.L., Ferrante, T.C., Terry, R., Jeanty, S.S.F., Li, C., Amamoto, R., et al. (2014). Highly multiplexed subcellular RNA sequencing in situ. *Science* 343, 1360–1363. <https://doi.org/10.1126/science.1250212>.
13. Chen, K.H., Boettiger, A.N., Moffitt, J.R., Wang, S., and Zhuang, X. (2015). RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* 348, aaa6090. <https://doi.org/10.1126/science.aaa6090>.
14. Halpern, K.B., Shenhav, R., Matcovitch-Natan, O., Tóth, B., Lemze, D., Golan, M., Massasa, E.E., Baydatch, S., Landen, S., Moor, A.E., et al. (2017). Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature* 542, 352–356. <https://doi.org/10.1038/nature21065>.
15. Nieman, K.M., Kenny, H.A., Penicka, C.V., Ladanyi, A., Buell-Gutbrod, R., Zillhardt, M.R., Romero, I.L., Carey, M.S., Mills, G.B., Hotamisligil, G.S., et al. (2011). Adipocytes promote ovarian cancer metastasis and provide energy for rapid tumor growth. *Nat. Med.* 17, 1498–1503. <https://doi.org/10.1038/nm.2492>.
16. Matulonis, U.A., Sood, A.K., Fallowfield, L., Howitt, B.E., Sehouli, J., and Karlan, B.Y. (2016). Ovarian cancer. *Nat. Rev. Dis. Primers* 2, 16061. <https://doi.org/10.1038/nrdp.2016.61>.
17. Lengyel, E. (2010). Ovarian cancer development and metastasis. *Am. J. Pathol.* 177, 1053–1064. <https://doi.org/10.2353/ajpath.2010.100105>.
18. Ahmed, N., and Stenvers, K.L. (2013). Getting to know ovarian cancer ascites: opportunities for targeted therapy-based translational research. *Front. Oncol.* 3, 256. <https://doi.org/10.3389/fonc.2013.00256>.
19. Kouno, T., Moody, J., Kwon, A.T.-J., Shibayama, Y., Kato, S., Huang, Y., Böttcher, M., Motakis, E., Mendez, M., Severin, J., et al. (2019). C1 CAGE detects transcription start sites and enhancer activity at single-cell resolution. *Nat. Commun.* 10, 360. <https://doi.org/10.1038/s41467-018-08126-5>.
20. Weston, S.A., and Parish, C.R. (1990). New fluorescent dyes for lymphocyte migration studies. Analysis by flow cytometry and fluorescence microscopy. *J. Immunol. Methods* 133, 87–97. [https://doi.org/10.1016/0022-1759\(90\)90322-m](https://doi.org/10.1016/0022-1759(90)90322-m).
21. Braut-Boucher, F., Pichon, J., Rat, P., Adolphe, M., Aubery, M., and Font, J. (1995). A non-isotopic, highly sensitive, fluorimetric, cell-cell adhesion microplate assay using calcein AM-labeled lymphocytes. *J. Immunol. Methods* 178, 41–51. [https://doi.org/10.1016/0022-1759\(94\)00239-S](https://doi.org/10.1016/0022-1759(94)00239-S).
22. Achilli, T.M., McCalla, S., Meyer, J., Tripathi, A., and Morgan, J.R. (2014). Multilayer spheroids to quantify drug uptake and diffusion in 3D. *Mol. Pharm.* 11, 2071–2081. <https://doi.org/10.1021/mp500002y>.
23. Thiele, E.W. (1939). Relation between catalytic activity and size of particle. *Ind. Eng. Chem.* 31, 916–920. <https://doi.org/10.1021/ie50355a027>.
24. Leedale, J., Herrmann, A., Bagnall, J., Fercher, A., Papkovsky, D., Sée, V., and Bearon, R.N. (2014). Modeling the dynamics of hypoxia inducible factor-1 α (HIF-1 α) within single cells and 3D cell culture systems. *Math. Biosci.* 258, 33–43. <https://doi.org/10.1016/j.mbs.2014.09.007>.
25. Fogler, H.S. (2016). *Elements of Chemical Reaction Engineering, Fifth Edition* (Prentice Hall).
26. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. *Cell* 177, 1888–1902.e21. <https://doi.org/10.1016/j.cell.2019.05.031>.
27. Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* 36, 411–420. <https://doi.org/10.1038/nbt.4096>.
28. McCarthy, D.J., Campbell, K.R., Lun, A.T., and Wills, Q.F. (2017). Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* 33, 1179–1186. <https://doi.org/10.1093/bioinformatics/btw777>.
29. Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.-C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* 14, 1083–1086. <https://doi.org/10.1038/nmeth.4463>.
30. van der Maaten, L., and Hinton, G. (2008). Visualizing data using t-SNE. *J. Mach. Learn. Res.* 9, 2579–2605.
31. McInnes, L., Healy, J., and Melville, J. (2018). Umap: uniform manifold approximation and projection for dimension reduction.. Preprint at arXiv. <https://arxiv.org/abs/1802.03426>.
32. Brugge, J.S. (2012). Into the deep: refocusing on 3D. *Nat. Cell Biol.* 14, 332. <https://doi.org/10.1038/ncb2470>.
33. Fennema, E., Rivron, N., Rouwkema, J., van Blitterswijk, C., and de Boer, J. (2013). Spheroid culture as a tool for creating 3D complex tissues. *Trends Biotechnol.* 31, 108–115. <https://doi.org/10.1016/j.tibtech.2012.12.003>.
34. Davidowitz, R.A., Selfors, L.M., Iwanicki, M.P., Elias, K.M., Karst, A., Piao, H., Ince, T.A., Drage, M.G., Dering, J., Konecny, G.E., et al. (2014). Mesenchymal gene program-expressing ovarian cancer spheroids exhibit enhanced mesothelial clearance. *J. Clin. Invest.* 124, 2611–2625. <https://doi.org/10.1172/JCI69815>.
35. Iwanicki, M.P., Davidowitz, R.A., Ng, M.R., Besser, A., Muranen, T., Merritt, M., Danuser, G., Ince, T.A., and Brugge, J.S. (2011). Ovarian cancer spheroids use myosin-generated force to clear the mesothelium.

- Cancer Discov. 1, 144–157. <https://doi.org/10.1158/2159-8274.CD-11-0010>.
36. Wolf, C.R., Hayward, I.P., Lawrie, S.S., Buckton, K., McIntyre, M.A., Adams, D.J., Lewis, A.D., Scott, A.R., and Smyth, J.F. (1987). Cellular heterogeneity and drug resistance in two ovarian adenocarcinoma cell lines derived from a single patient. *Int. J. Cancer* 39, 695–702. <https://doi.org/10.1002/ijc.2910390607>.
37. Beaufort, C.M., Helmi, J.C.A., Piskorz, A.M., Hoogstraal, M., Ruigrok-Ritstier, K., Besselink, N., Murtaza, M., van IJcken, W.F.J., Heine, A.A.J., Smid, M., et al. (2014). Ovarian cancer cell line panel (OCCP): clinical importance of in vitro morphological subtypes. *PLoS One* 9, e103988. <https://doi.org/10.1371/journal.pone.0103988>.
38. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* 102, 15545–15550. <https://doi.org/10.1073/pnas.0506580102>.
39. Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSEA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinform.* 14, 7. <https://doi.org/10.1186/1471-2105-14-7>.
40. Hoskins, E., Rodríguez-Canales, J., Hewitt, S.M., Elmasri, W., Han, J., Han, S., Davidson, B., and Kohn, E.C. (2011). Paracrine SLPI secretion up-regulates MMP-9 transcription and secretion in ovarian cancer cells. *Gynecol. Oncol.* 122, 656–662. <https://doi.org/10.1016/j.ygyno.2011.04.052>.
41. Serra, D., Mayr, U., Boni, A., Lukonin, I., Rempfler, M., Challet Meylan, L., Stadler, M.B., Strnad, P., Papasaikas, P., Vischi, D., et al. (2019). Self-organization and symmetry breaking in intestinal organoid development. *Nature* 569, 66–72. <https://doi.org/10.1038/s41586-019-1146-y>.
42. Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 1, 417–425. <https://doi.org/10.1016/j.cels.2015.12.004>.
43. Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 27, 1739–1740.
44. Izar, B., Tirosh, I., Stover, E.H., Wakiro, I., Cuoco, M.S., Alter, I., Rodman, C., Leeson, R., Su, M.-J., Shah, P., et al. (2020). A single-cell landscape of high-grade serous ovarian cancer. *Nat. Med.* 26, 1271–1279. <https://doi.org/10.1038/s41591-020-0926-0>.
45. Macintyre, G., Goranova, T.E., De Silva, D., Ennis, D., Piskorz, A.M., Eldridge, M., Sie, D., Lewsley, L.A., Hanif, A., Wilson, C., et al. (2018). Copy number signatures and mutational processes in ovarian carcinoma. *Nat. Genet.* 50, 1262–1270. <https://doi.org/10.1038/s41588-018-0179-8>.
46. Zhao, J., Jaffe, A., Li, H., Lindenbaum, O., Sefik, E., Jackson, R., Cheng, X., Flavell, R.A., and Kluger, Y. (2021). Detection of differentially abundant cell subpopulations in scRNA-seq data. *Proc. Natl. Acad. Sci. USA* 118. <https://doi.org/10.1073/pnas.2100293118>.
47. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., 3rd, Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. *Cell* 184, 3573–3587.e29. <https://doi.org/10.1016/j.cell.2021.04.048>.
48. Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M.N., and Sergushichev, A. (2021). Fast gene set enrichment analysis. Preprint at bioRxiv. <https://doi.org/10.1101/060012>.
49. Ke, R., Mignardi, M., Pacureanu, A., Svedlund, J., Botling, J., Wählby, C., and Nilsson, M. (2013). In situ sequencing for RNA analysis in preserved tissue and cells. *Nat. Methods* 10, 857–860. <https://doi.org/10.1038/nmeth.2563>.
50. Merritt, C.R., Ong, G.T., Church, S.E., Barker, K., Danaher, P., Geiss, G., Hoang, M., Jung, J., Liang, Y., McKay-Fleisch, J., et al. (2020). Multiplex digital spatial profiling of proteins and RNA in fixed tissue. *Nat. Biotechnol.* 38, 586–599. <https://doi.org/10.1038/s41587-020-0472-9>.
51. Coussens, L.M., and Werb, Z. (2002). Inflammation and cancer. *Nature* 420, 860–867. <https://doi.org/10.1038/nature01322>.
52. Jia, D., Nagaoka, Y., Katsumata, M., and Orsulic, S. (2018). Inflammation is a key contributor to ovarian cancer cell seeding. *Sci. Rep.* 8, 12394. <https://doi.org/10.1038/s41598-018-30261-8>.
53. Chen, G.M., Kannan, L., Geistlinger, L., Kofia, V., Safikhani, Z., Gendoo, D.M.A., Parmigiani, G., Birrer, M., Haibe-Kains, B., and Waldron, L. (2018). Consensus on molecular subtypes of high-grade serous ovarian carcinoma. *Clin. Cancer Res.* 24, 5037–5047. <https://doi.org/10.1158/1078-0432.CCR-18-0784>.
54. Cancer Genome Atlas Research Network (2011). Integrated genomic analyses of ovarian carcinoma. *Nature* 474, 609–615. <https://doi.org/10.1038/nature10166>.
55. Leary, E., Rhee, C., Wilks, B., and Morgan, J.R. (2016). Accurate quantitative wide-field fluorescence microscopy of 3-D spheroids. *BioTechniques* 61, 237–247. <https://doi.org/10.2144/000114472>.
56. Muir, R.K., Guerra, M., and Bogoy, M.M. (2022). Activity-based diagnostics: recent advances in the development of probes for use with diverse detection modalities. *ACS Chem. Biol.* 17, 281–291. <https://doi.org/10.1021/acscchembio.1c00753>.
57. Zilionis, R., Nainys, J., Veres, A., Savova, V., Zemmour, D., Klein, A.M., and Mazutis, L. (2017). Single-cell barcoding and sequencing using droplet microfluidics. *Nat. Protoc.* 12, 44–73. <https://doi.org/10.1038/nprot.2016.154>.
58. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. <https://doi.org/10.1038/nmeth.2089>.
59. Edmondson, R., Broglie, J.J., Adcock, A.F., and Yang, L. (2014). Three-dimensional cell culture systems and their applications in drug discovery and cell-based biosensors. *Assay Drug Dev. Technol.* 12, 207–218. <https://doi.org/10.1089/adt.2014.573>.
60. Mathews Griner, L.A., Zhang, X., Guha, R., McKnight, C., Goldlust, I.S., Lal-Nag, M., Wilson, K., Michael, S., Titus, S., Shinn, P., et al. (2016). Large-scale pharmacological profiling of 3D tumor models of cancer cells. *Cell Death Dis.* 7, e2492. <https://doi.org/10.1038/cddis.2016.360>.
61. Langdon, S.P., Lawrie, S.S., Hay, F.G., Hawkes, M.M., McDonald, A., Hayward, I.P., Schol, D.J., Hilgers, J., Leonard, R.C.F., and Smyth, J.F. (1988). Characterization and properties of nine human ovarian adenocarcinoma cell lines. *Cancer Res.* 48, 6166–6172.
62. Drost, J., and Clevers, H. (2018). Organoids in cancer research. *Nat. Rev. Cancer* 18, 407–418. <https://doi.org/10.1038/s41568-018-0007-6>.
63. Rossi, G., Manfrin, A., and Lutolf, M.P. (2018). Progress and potential in organoid research. *Nat. Rev. Genet.* 19, 671–687. <https://doi.org/10.1038/s41576-018-0051-9>.
64. Zhang, S., Dolgalev, I., Zhang, T., Ran, H., Levine, D.A., and Neel, B.G. (2019). Both fallopian tube and ovarian surface epithelium are cells-of-origin for high-grade serous ovarian carcinoma. *Nat. Commun.* 10, 5367. <https://doi.org/10.1038/s41467-019-13116-2>.
65. Maenhoudt, N., Defraye, C., Boretto, M., Jan, Z., Heremans, R., Boeckx, B., Hermans, F., Arijis, I., Cox, B., Van Nieuwenhuysen, E., et al. (2020). Developing organoids from ovarian cancer as experimental and preclinical models. *Stem Cell Rep.* 14, 717–729. <https://doi.org/10.1016/j.stemcr.2020.03.004>.
66. Maru, Y., Tanaka, N., Itami, M., and Hippo, Y. (2019). Efficient use of patient-derived organoids as a preclinical model for gynecologic tumors. *Gynecol. Oncol.* 154, 189–198. <https://doi.org/10.1016/j.ygyno.2019.05.005>.
67. Kopper, O., de Witte, C.J., Löhmußaar, K., Valle-Inclán, J.E., Hami, N., Kester, L., Balgobind, A.V., Korving, J., Proost, N., Begthel, H., et al. (2019). An organoid platform for ovarian cancer captures intra- and inter-patient heterogeneity. *Nat. Med.* 25, 838–849. <https://doi.org/10.1038/s41591-019-0422-6>.

68. He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2020). Mask R-CNN. *IEEE Trans. Pattern Anal. Mach. Intell.* **42**, 386–397.
69. Briggs, J.A., Weinreb, C., Wagner, D.E., Megason, S., Peshkin, L., Kirschner, M.W., and Klein, A.M. (2018). The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. *Science* **360**, eaar5780. <https://doi.org/10.1126/science.aar5780>.
70. Andreatta, M., and Carmona, S.J. (2021). UCell: robust and scalable single-cell gene signature scoring. *Comput. Struct. Biotechnol. J.* **19**, 3796–3798. <https://doi.org/10.1016/j.csbj.2021.06.043>.
71. Sergushichev, A.A. (2016). An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. Preprint at bioRxiv. <https://doi.org/10.1101/060012>.
72. Mootha, V.K., Lindgren, C.M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., et al. (2003). PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273. <https://doi.org/10.1038/ng1180>.
73. Hillje, R., Pelicci, P.G., and Luzi, L. (2020). Cerebro: interactive visualization of scRNA-seq data. *Bioinformatics* **36**, 2311–2313. <https://doi.org/10.1093/bioinformatics/btz877>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
HGSOC ascites-derived organoid samples	Laboratory of James Brenton	N/A
Chemicals, peptides, and recombinant proteins		
Calcein AM	ThermoFisher	Cat#L3224
Ethidium Homodimer-1	ThermoFisher	Cat#L3224
RPMI media	Gibco	Cat#11-875-085
Pen Strep	Gibco	Cat#15140-122
FBS	Gibco	Cat#16000044
BSA	Gibco	Cat#15260037
TrypLE Select	Gibco	Cat#12563-011
Accutase	Millipore Sigma	Cat#A6964-100ML
L-Glutamine (200 mM)	Gibco	Cat#25030081
Sodium Pyruvate (100 mM)	Gibco	Cat#11360070
PDMS	Momentive Performance Materials	Cat#RTV615A
Trichloromethylsilane	Millipore Sigma	Cat#M85301-5G
solvents: SU-8 developer, AZ 300MIF developer, molecular biology grade 100% ethanol, IPA, acetone	Various	Various
common salts & detergents: KCl, MgCl ₂ , NaCl, TRIS-HCl, TRIS acetate, magnesium acetate, potassium acetate, sodium acetate, EDTA, Tween-20, IGEPAL	Various	Various
inDrop V2 reagents	Various	Zilionis et al. ⁵⁷
Deposited data		
PEO1 spheroid SEEP scRNA-seq reads	this study	GEO: GSE157299
HGSOC organoid SEEP scRNA-seq reads	this study	GEO: GSE157299
HGSOC solid tumor Punch-SEEP scRNA-seq reads	this study	GEO: GSE157299
PEO1 calcein staining controls scRNA-seq reads	this study	GEO: GSE157299
Experimental models: Cell lines		
PEO1 cells (ECACC 10032308)	MilliporeSigma	Cat#10032308-1VL
Experimental models: Organisms/strains		
Adult NOD.Cg-PrkdcscidIl2rgtm1Wjl/SzJ (NSG) mice	CRUK PDX biobank	N/A
Oligonucleotides		
inDrop v3 oligonucleotides	Laboratory of Allon Klein	Custom Orders
inDrop v3 single-cell barcodes	github.com/indrops/indrops/blob/master/ref/barcode_lists/	Custom Orders
Software and algorithms		
inDrop.py	Klein et al. ³	Code available on this GitHub repository: github.com/indrops/indrops
Seurat (V4.1.1)	Stuart et al. ²⁶ ; Hao et al. ⁴⁷	https://cran.r-project.org/web/packages/Seurat/index.html
Leica Application Suite X	Leica Microsystems	N/A
Harmony (4.9)	PerkinElmer	N/A
Fiji-win64	Schneider et al. ⁵⁸	https://imagej.nih.gov/ij/
FlowJo (10.6.1)	BD Biosciences	N/A

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Daseq v1.0.0	Zhao et al. ⁴⁶	v1.0.0 https://github.com/KlugerLab/Daseq
Python (2.7)	Python Software Foundation	N/A
R (v3.6.3)	The R Project for Statistical Computing	N/A
Additional R packages: dplyr, tidyr, Matrix, ggplot2, colorRamps, reshape2, ggextra, GSVA, parallel, ComplexHeatmap, circlize, GetoptLong, grid, gridExtra, stats, lattice, latticeExtra, fgsea, cogen, RColorBrewer, SDMTTools, ggrepel, scales, DT	CRAN & Bioconductor repositories	N/A
In-house code for image analysis (MATLAB, python, & R)	this study	All code is available on this study's; GitHub repository: https://github.com/davidbmorse/SEEP_image_processing ; https://doi.org/10.5281/zenodo.7765319
In-house code for sequencing data analysis (R)	this study	All code is available on this study's; GitHub repository: https://github.com/davidbmorse/SEEP ; https://doi.org/10.5281/zenodo.7765315
Other		
Photomasks (25400 DPI)	CAD/Art Services	Various designs
500 μ m thick test-grade silicon wafers of various diameters	University Wafer	Various
Microtubing (i.d. 0.38 mm $\times$ o.d. 1.09 mm)	Scientific Commodities	Cat#BB31695-PE/2
Disposable Biopsy Punches (various diameters)	Integra Miltex	Various

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Craig J. Thomas (craigt@mail.nih.gov).

Materials availability

This study did not generate new unique reagents.

Data and code availability

All data needed to evaluate the conclusions in the paper are present in the main text and/or the supplementary materials. The datasets generated in this study have been deposited in the Gene Expression Omnibus, a public functional genomics data repository, under the accession number GSE157299. Codes used within this study are available on GitHub at https://github.com/davidbmorse/SEEP_image_processing and <https://github.com/davidbmorse/SEEP>.

METHOD DETAILS

Cell culture and tissue handling

Several methods exist to create spheroidal cultures.^{33,59,60} In this study, PEO1 cells were cultured in a stem-cell based media in ultra-low attachment (ULA) plates to compel spheroid formation. PEO1 (ECACC 10032308) cells were grown and sustained in adherent cell culture flasks in RPMI media with 10% FBS before being transferred to ultralow attachment (ULA) 96-well plates for HGSOc spheroid formation.^{36,61} 10,000 cells were seeded per well and allowed to grow for 14 days in spheroid media prior to harvesting.

Organoid cultures have emerged over the past decade as highly valuable cancer cell models. While there are diverging definitions, most agree that an organoid model is an explant of cells resected from an in vivo model that self-organize into 3D cell clusters that retain key elements of the tumor microenvironment.^{62,63} Several organoid models of HGSOc have been reported utilizing a variety of methods including derivations from fallopian epithelial stem cells, surface epithelial cells, or cells collected from patient ascites.^{64–67} Organoids were harvested from sphere-forming cells present in HGSOc patient ascites collections and grown in a custom Matrigel solution for 14 passages prior to harvesting. Organoid size and cell number was dependent on natural ascites sphere formation. DNA sequencing of the organoids revealed chromosomal aberrations and instabilities characteristic of p53 mutated HGSOc.

HGSOC cells for PDX models were collected from patients undergoing interval debulking surgery or ascites drains at Addenbrooke's hospital in Cambridge, UK. PDX tumors were generated using cells that were rapidly thawed from cryopreservation, and 3 million cells were injected subcutaneously into NOD.CgPrkdc^{scid}Il2rg^{tm1Wjl}/SzJ (NSG) mice. Mice were monitored weekly, and the tumor volume determined using caliper measurements. Once reaching a maximum volume of 1500 mm³, mice were culled and the tumor harvested. Serial transplants were performed to generate maximum of 3-4 generations. Animal procedures were conducted in accordance with the local AWERB, NACWO and UK Home Office regulations (Animals (Scientific Procedures) Act 1986).

SEEP calibration staining

Symmetric staining on spheroids

HGSOC spheroids were imaged using both confocal and light sheet microscopy to quantify the kinetics of dye-uptake. High-content confocal microscopy was performed on 96-well spheroid plates as follows: We removed exogenous esterases from mature spheroid plates by performing a gentle media-exchange using fresh, serum free, and phenol-red free media, leaving a final volume of 50 μ L of fresh media per well after the exchange. We transferred spheroid plates to a high-content Opera Phoenix confocal microscope and gently added 50 μ L of a 2x concentration calcein-AM staining solution (in culture-matched media) to the wells using a multichannel pipette. The final calcein-AM concentration was 0.5 μ M. Spheroids were imaged every 5 min for 5 hours to monitor dye penetration. Z-slices, 12 μ m apart were taken from the objective-facing surface of the spheroid through the spheroid's midpoint. In addition, we performed light-sheet microscopy on individual spheroids suspended in a 0.5 μ M calcein-AM solution during imaging. 12 μ m spaced z-slices were taken throughout the volume of the spheroids for 2 hours to monitor dye penetration. In both cases a 488 nm laser was used to excite calcein, and fluorescence emission collection was centered at 522 nm. For the calibration step, imaging was performed at 37 °C and 5% CO₂; tissues were discarded after imaging.

Symmetric staining on organoids

HGSOC organoids were imaged using high content confocal microscopy to quantify the kinetics of dye uptake as follows: 4 wells of organoids, grown in 24 well plates, were gently washed 3 times in RPMI with 10% FBS and allowed to gravity settle for 3 min between washes to ensure Matrigel, debris, single-cells, and small cell clusters were removed. After the last wash, organoids were transferred to flat bottom 96-well plates in 100 μ L of fresh, serum free, and phenol-red-free media for imaging. We transferred organoid plates to a high-content Opera Phoenix confocal microscope and gently added 100 μ L of 2x concentration calcein-AM staining solution (in culture-matched media) to the wells using a multichannel pipette. The final calcein-AM concentration was 0.5 μ M. Organoids were imaged as a series of z-stacks every 10 min for up to 2 hours to monitor dye penetration.

Punch-SEEP on solid tumor biopsy

Punch-SEEP staining was performed using a prepared biopsy punch. Dye penetration was monitored at various time points by confocal microscopy by ejecting tissue biopsies, washing, and imaging in media-filled glass-bottom dishes. Dye-loaded biopsy punches were prepared using disposable, 1.5 mm wide punches, prefilled with a calcein-AM saturated hydrogel plug that melted in cell-culture conditions. A 10% gelatin solution in sterile PBS was prepared and stored at 4 °C. Prior to biopsy loading, the gelatin was melted in a warm water bath and 4 mL were pipetted into a small plastic weigh-boat to a height of 2 mm. The weigh-boat containing the gelatin was then lyophilized for 24 hours. The lyophilized gelatin was cut into 1 cm squares and submerged in a 250 μ L solution of 2 mM calcein AM. After a few minutes, full absorption of the dye solution into the gelatin pieces was observed. The gelatin + dye solution was kept on ice and in the dark until use. Just before biopsy punching, the gelatin squares (now very soft and fragile) were removed from the dye solution and placed within a weigh-boat or on an impermeable cutting surface. Biopsies were taken of the gelatin squares and visually inspected to ensure a gelatin plug was flush with the surface of the biopsy punch. The gelatin plug was then gently pushed 0.5 mm into the biopsy punch to ensure the circular blade had clear/direct access to the tumor surface. Biopsies were immediately used to punch the dissected tumors. Punch-biopsies taken from subcutaneous HGSOC tumors grown in PDX mouse models were imaged using confocal microscopy to quantify the kinetics of dye-uptake. Tissue acquisition, staining, and imaging were performed as follows: Mice were sacrificed, and tumors were dissected and stored in ice-cold RPMI. Tumors were cut in half and biopsies were taken from the tumor surface normal to the cross-sectional cut. Biopsies were taken using dye-loaded 1.5 mm diameter self-ejecting biopsy punches that were prepared in advance. After punching, the biopsies were not ejected from the biopsy puncher and the entire puncher was placed at roughly a 20° angle in a 15 cm cell-culture dish prefilled with a few mm of RPMI at 4 °C. The sharp tip of the punch was submerged in the media while the handle rested on the lip of the culture dish. The lid was placed gently over the punches and the entire dish was incubated in a cell culture incubator at 37 °C and 5% CO₂. We performed this procedure for multiple biopsies, and imaged dye propagation through the sample at various time points using confocal microscopy. We ejected biopsies into 40 mL of RPMI to wash away excess dye. We placed washed biopsies on glass-bottom imaging dishes filled with 2 mL of RPMI media and imaged the samples on a Leica SP5 confocal microscope. Multiple fields of view were acquired to image the entire biopsy (Figures 4A, S3D, and S3E). Z-slices were taken every 10 μ m through the sample.

Imaging dye penetration

Confocal and light-sheet microscopy were performed to quantify the penetration of dye into tissues. To account for the attenuation of signal with depth in the dense samples, attenuation was calculated and used to normalize measurements. We stained PEO1 cells with cell tracker dyes green, violet, and deep red prior to spheroid formation. We formed 10,000-cell spheroids in ULA 96-well plates as done in previous experiments. After spheroid maturation, signal attenuation was measured by taking z-slices 10 μ m apart beginning from the objective-facing spheroid surface. For cell tracker green, of equivalent excitation/emission parameters to calcein,

exponential decay was observed up to 100 μm into the spheroids, and a trend line, $y = 14,425e^{0.201}$ fit to the data with $R^2 = 0.999$. This attenuation factor was used to normalize the images monitoring dye penetration into both spheroids and organoids. After attenuation normalization, dye penetration images for spheroids and organoids were computationally divided into n shells of equal volume. To identify spheroid and organoid boundaries among hundreds of confocal images and segment them into concentric shells of equal volume, the open-source convoluted neural network, Mask R-CNN, was implemented across image stacks.⁶⁸ Once the tissue structures were localized, a custom image analysis pipeline was used to segment cell layers based on cells' Euclidian distances from tissue margins. An erosion strategy was used and implemented in MATLAB; the code can be found in the GitHub pages mentioned below. n was calculated based on the % fluorescence difference between adjacent layers for each time point. The time needed to reach the highest n , with the greatest change in intensity between surface and core fluorescence was noted. Spheroid optimal penetration time was 60 min while organoid optimal penetration time was 55 min. For the biopsy samples, attenuation was not calculated as change in dye concentration was calculated perpendicular to the direction of attenuation. For biopsies, z-planes were flattened into a single max-projection and change in fluorescence intensity between the surface of the biopsy and the core were observed in time. n , the number of layers the biopsy would be divided into, was calculated based on the % fluorescence difference between adjacent layers for each time point. At 75 min, change in intensity between biopsy surface and core had plateaued.

For all tissues, a maximum change in fluorescence intensity plateaued at specific times. Achieving a steady state of fluorophore acquisition proved ideal for consistently segmenting tissues based on dye uptake as harvesting times were flexible. Beyond a critical time, each tissue could be removed from its staining solution at a wide range of times without effecting segmentation performance. Due to the desire to limit cell handling, tissues were harvested at the earliest time possible for proper segmentation.

SEEP controls for examining the effect of staining and sorting on transcription

To control for the potential accumulation of technical bias via calcein staining and/or FACS sorting, HGSOC PEO1 cells were exposed to the staining and sorting conditions used in SEEP. PEO1 cells were grown as monolayers, gently dissociated, and either immediately sequenced (control group) or conditioned via sorting (sort only group), staining (stain only group), or staining and sorting (sort & stain group). All experiments were performed twice to garner two technical replicates for each condition. All replicates and conditions were processed via inDrop independently. All library preparation was performed independent. After library preparation, replicates and conditions were indexed according to their identity, pooled, and sequencing on a 75 bp Illumina Nextseq 400M High output kit with 5% PhiX as a spike-in control. In total, 4,574 cells passed filtering to explore possible technical bias. Cells of each condition were processed and analyzed computationally as described in the below sections titled *Indrop Deployment for scRNA-sequencing*, *Data Preprocessing and Filtering*, and *Data Analysis and Visualization*.

SEEP measurement, dissociation, and FACS

Samples were stained as previously described for optimal segmentation times (spheroid = 60 min, organoid = 55 min, and biopsy = 75 min). For each of the following systems, the experiments described below were repeated two times on different days and with different cells giving us two biologically distinct (i.e., cultured at different times under identical conditions) replicates. Each biologically replicate was further split into two technical replicates just prior to performing FACS. Overall, each system had 4 replicates (2 biological replicates with 2 technical replicates each).

Spheroids

192 spheroids, from 2 ULA 96-well plates, were pooled in a 15 ml conical flask and washed 2x in fresh spheroid media. Spheroids were allowed to settle to the bottom of the 15 ml conical for 2 min with no centrifugation before removing supernatant and cellular debris. After the last wash, 2 ml of serum-free spheroid media was added to the spheroids. 2 μL of freshly made 500 μM calcein AM in PBS + 0.05% BSA was gently added to the spheroid solution for a final calcein-AM concentration of 0.5 μM . The conical was gently flicked 5x to resuspend the spheroids, the lid was vented, and the conical was stored in a cell culture incubator (37 $^{\circ}\text{C}$ and 5% CO_2) for 60 min. Every 10 min the conical was gently flicked 5x to re-suspend the settled spheroids. After 60 min, the spheroids were washed 3x in 15 ml of PBS + 0.05% BSA and allowed to gravity settle between washes.

Samples were dissociated in 2 ml of Accutase at 37 $^{\circ}\text{C}$ for 30 min with intermittent pipetting every 10 min to encourage dissociation. Intermittent pipetting consisted of gently pipetting 20x using a 1000 μL pipette tip pre-blocked with 0.05% BSA. After 30 min and the final dissociative pipetting, the spheroids were no longer visible, and the solution appeared cloudy and homogeneous. 13 ml of RPMI with 10% FBS was added to the solution, inverted 5x, and centrifuged for 5 min at 400g. The cells were resuspended in 1 ml of ice-cold PBS + 0.05% BSA and passed through a 40 μm cell filter. The filter was gently rinsed with 10 ml of ice-cold PBS + 0.05% BSA. 11 ml cell suspension was centrifuged at 400 g for 5 min, resuspended in 1 ml of ice-cold PBS + 0.05% BSA, and moved to a polypropylene FACS tube. Cells were kept on ice and in darkness until FACS.

Organoids

4 wells of organoids grown in 24 well plates were gently washed 3x in RPMI with 10% FBS and allowed to gravity settle for 3 min between washes to ensure Matrigel, debris, single-cells, and small organoids are removed. After Matrigel removal, organoids were processed as spheroids, above, except for the following changes: the staining incubation time was 55 min and 2 mL of Tissue E Express (Thermo) was used instead of Accutase as the dissociation agent.

PDX Biopsies

Individual Biopsies ($n = 5$) from subcutaneous PDX tumors were punched using the aforementioned method and dye-loaded biopsy punches. They were incubated at 37 $^{\circ}\text{C}$ and 5% CO_2 for 75 min, and then ejected into 40 mL of room temperature RPMI + 10% FBS.

Biopsies were centrifuged for 3 min at 500 g to collect. The supernatant was removed, and a scalpel was used to break up firm biopsies at the bottom of the 50 mL conical. The biopsy was then dissociated and washed using a gentle-MACS dissociator according to the 37 °C *m_TDK_2* (mouse kit) protocol provided by the dissociation kit. *m_imptumor_01* protocol was performed if visible biopsy fragments remained after dissociation. Dissociated cells were centrifuge in a MAX tube at 350 g for 5 min and resuspended in DMEM + 10% FBS filtered through a 70 µm filter placed on a 50 mL conical. The strainer was washed in 50 mL DMEM + 10% FBS. Finally, the cell suspension was centrifuge at 350 g for 5 min, resuspended in 1 mL of ice-cold PBS + 0.05% BSA, and moved to a polypropylene FACS tube. Cells were kept on ice and in darkness until FACS.

FACS

Cells were sorted into 4 (spheroids) or 3 (organoids and biopsies) bins based on 515 nm emission from a 488 nm laser source. Debris, dead cells, and cell clumps were conservatively removed based on forward/side scatter profiles. The sample was allowed to run for a few minutes while laser power was adjusted to observe a Gaussian distribution of unstained cells followed by a leading tail of stained cells. Bin widths were set to receive equal proportions of cells. After adjusting laser power and bin width, samples were collected in 1.5 mL protein LoBind Eppendorf tubes prefilled with 250 µL of ice-cold PBS + 0.05% BSA. Samples were kept on ice and immediately barcoded via the inDrop protocol.³

InDrop deployment for scRNA-sequencing

For all tissues inDrop was performed as follows: Cells were suspended in a 1x PBS, 0.01% (v/v) BSA, and 15% (v/v) OptiPrep solution at a concentration of 120,000 cells/ml and processed via the inDrop protocol with v3 barcoding design.^{69,57} Polyacrylamide bar-coded hydrogel beads were fabricated in house and quality controlled using FISH and qPCR on the extended primers; the barcode diversity for all experiments was 147,456. Capture efficiency and cross-contamination levels were quantified for each batch of hydrogels by species mixing experiments done on human HEK293 and mouse 3T3 fibroblasts and by fast imaging of the microfluidic encapsulation events. To maintain a barcode collision rate of less than 1%, collection fractions were acquired at or below 2,949 cells. Microfluidic handling and library preparation were carried out according to published protocols.⁶¹ Microfluidic flow rates, cell capture rates, and optically observed doublet rates can be found in Table S1A. Libraries were amplified and barcoded via limited-cycle PCR and inspected for quality on a BioAnalyzer HS kit. Library size distribution and concentration was measured on a BioAnalyzer HS and Qubit HS respectively to inform sample pooling at equimolar ratios. The final library was purified on a 1.5x volumetric ratio of AmpureXP beads and quantified using the Kapa NGS library quantification kit and a BioAnalyzer HS. The oligonucleotide sequences used to barcode the final library can be found in Tables S1B and S1C. Libraries were sequenced using a 75 bp Illumina Nextseq 400M High output kit with 5% PhiX was used as a spike-in control. Oligonucleotide barcode sequences were recorded in Tables S1B and S1C, and sequencing parameters were recorded in Table S1D.

Data preprocessing and filtering

Fastq files were generated using Illumina's bcl2fastq script (Table S1E). The data were filtered, quantified, and sorted using the inDrop analysis pipeline. The pipeline execution parameters can be found in Table S1E and parameters for a *yaml* file can be found in Table S1F.

Raw counts matrices from the inDrop pipeline^{69,57} were converted into sparse matrices using a custom R script and loaded into the Seurat²⁶ tool using the *Read10x* function. The custom script for sparse matrix conversion can be found in Table S1E. Cells were further filtered from empty droplets and quality controlled by analyzing the distribution of the number of UMIs per cell, the number of genes per cell, and the fraction of mitochondrial genes per cell. Common scRNA-seq filtering criteria were assessed and implemented on the individual cell collections and sequencing runs. As sample type, sample quality, cell-capture rate, and sequencing depth varied across experiments, filtering thresholds sought to differentiate between high-quality and low-quality cells based on the relative abundance of reads allocated to each cell barcode. To do this, we fit a mixture model to the data following a procedure outlined by the Martin Hemberg Group (<https://github.com/hemberg-lab/scRNA.seq.course>) and identified where the higher and lower distributions intersected. For the spheroid samples, cells were selected that contained between 200 and 2,500 unique genes and between 0 and 18% mitochondrial reads. 7,908 cells remained after filtering. These were made up of 1,178 center cells, 2,471 inside cells, 2,736 outside cells, and 2,667 surface cells. For the organoid samples, cells were selected that contained between 580 and 5,000 unique genes, between 5 and 30% mitochondrial reads, and between 500 and 15,000 unique molecules. 7,771 organoid cells remained after filtering. These were made up of 2,285 center cells, 2,819 middle cells, and 2,667 surface cells. For the PDX-biopsy samples, cells were selected that contained between 200 and 3000 unique genes, between 0 and 40% mitochondrial reads, and between 350 and 15,000 unique molecules. 1,967 PDX biopsy cells remained after filtering. These were made up of 280 center cells, 677 middle cells, and 1,010 surface cells.

After filtering, cells were scored for cell cycle by the *CellCycleScoring* function. The data were then log normalized using a global-scaling normalization via the *NormalizeData* function and a scale factor of 10,000. For the biopsy sample containing both human and mouse cells, mouse cells (n = 151) and mouse-human doublets (n = 29) were identified and removed by aligning RNA-seq reads to both human and mouse transcriptomes. These data were removed from the overall analysis.

Immunohistochemical evaluation of formalin-fixed paraffin-embedded spheroids (FFPE)

All slides were baked prior to staining. MIB-1 immunohistochemical staining was performed on an automated immunostainer BenchMark Ultra (Roche). A 1:200 dilution of MIB-1 (Clone MIB-1, Dako, Cat No. M7240, RRID:AB_2142367) was applied and

detected with an ultraView Universal DAB Detection Kit (Roche, Cat No. 760-500). The slides were dehydrated in graded alcohols, treated with xylene (2 x 5 min), and cover slipped.

In situ hybridization-based detection of SLPI and CXCL10

The manual RNAscope® 2.5 High Definition (HD)- BROWN Assay kit (Advanced Cell Diagnostics, Cat No. 322300) was used according to manufacturer's instructions to perform ISH for Hs-SLPI Probe (Cat No. 531861) and Hs-CXCL10 (Cat No. 311851) in formalin-fixed paraffin-embedded (FFPE) spheroids. Slides were deparaffinized in xylene (2 x 5 min), dehydrated in 100% EtOH (2 x 1 min), air dried at RT, and incubated in RNAscope® Hydrogen Peroxide (Cat No. 322330, ACDBio) at RT for 10 min to quench endogenous peroxidases. Slides were washed in DI water twice before being submerged in 700 mL of fresh boiling RNAscope® 1X Target Retrieval Reagents solution (Cat No. 322000, ACDBio) for 15 min and then washed twice in DI water, once in 100% EtOH, and air dried at RT. A hydrophobic barrier was drawn around the tissue and RNAscope® Protease Plus (Cat No. 322330, ACDBio) was applied for 20 min in a HybEZ™ Oven at 40°C (ACDBio). Slides were then washed twice in DI water and incubated with the appropriate probe for 2 hours in the HybEZ™ Oven. RNAscope® signal amplification reagents (Cat No. 322310, ACDBio) AMP 1 (30 min), AMP 2 (15 min), AMP 3 (30 min), AMP 4 (15 min), AMP 5 (75 min), and AMP 6 (15 min) were applied and incubated in the HybEZ™ Oven. Before adding each AMP reagent, the slides were washed twice with RNAscope® 1X Wash Buffer (Cat No. 310091, ACDBio). RNAscope® DAB detection reagents (Cat No. 322310, ACDBio) were applied and incubated for 10 min in the HybEZ™ Oven. Sections were counterstained with Harris-Mayer's Hematoxylin, washed in tap water, placed in 0.02% ammonia water for 10 seconds, and washed again with tap water. Sections were then dehydrated in graded alcohols, treated with xylene (2 x 5 min), and cover slipped.

Spheroid image analysis

A semi-automated algorithm systematically analyzed gene expression in sub-regions of individual spheroids. An image analysis script was created and executed in the Matlab scientific computing environment (Matlab version 2021a, Mathworks, Inc.). The script functions performed RGB (red, green, blue) three-channel thresholding to isolate the image region pixels that contain the colorimetric dye marker of gene expression. The same thresholding settings were applied to all images. The sum area of the binary masks from these regions approximates the number of 'expression positive' cells, which exhibit detectable levels of gene expression. Intensity differences were measured using script functions converting the RGB color space to grayscale values and then inverted the grayscale image so that dark regions correspond to higher intensity values. The integrated intensity of the inverted grayscale image in the 'expression positive' cell binary mask was used to approximate total gene expression in these regions. The script applied a separate threshold function to isolate the 'whole spheroid' region as a binary mask. Binary morphological functions were applied to fill any internal gaps in the 'whole spheroid' region. The script distance transform function converted the pixels in the 'whole spheroid' binary mask region producing output pixel values that represent the distance (in pixel count) to the edge of the 'whole spheroid' region. Within each spheroid, the distance labeled pixels were sorted into radial band regions. The cut-off values for distance sorting produced four band regions per spheroid, and each of these bands had approximately the same width (in pixels). The 'expression positive' cell region was then measured in each of the radial band regions. The area of the 'expression positive' cell regions was also normalized to the total area in the corresponding radial band.

Data analysis and visualization

Pairwise anchor-based single-cell data integration implemented in Seurat v4.1.1⁴⁷ was used to generate the HGSOC model from the Center and Surface segmentation layers of cultured PEO1 spheroids, patient-derived organoids, and PDX-derived biopsy cells. The top 2000 features most variable across the data sources were identified by the *SelectIntegrationFeatures* function and subsequently used for data alignment with the *FindIntegrationAnchors* and *IntegrateData* methods. The default integration parameters were applied with log normalization and CCA dimension reduction. *ScaleData* (mitochondrial and cell cycle genes regressed out), *RunPCA*, and *RunUMAP* were performed on the integrated dataset. Subsequently, the Daseq v1.0.0 algorithm⁴⁶ was used to detect distinct cell subpopulations in the regions of DA between the Center and Surface layers of the integrated HGSOC model (absolute DA-seq score > 0.8 and cluster size of at least 50 cells). To characterize DA subpopulations, Seurat's *FindConservedMarkers* (Wilcoxon Rank Sum test) was performed to select genes upregulated in *cluster vs all* mode in each data source of the HGSOC model (maximum p-value less than 0.05). Functional ORA was then run with the Hallmark Collection of MSigDB v6.2 using fgsea's v1.18 *fora* (Fisher's exact test) and *CollapsePathwaysORA* (an algorithm retaining over-represented non-redundant gene sets).⁴⁸ Ucell v1.99.7⁷⁰ and the *AddModuleScore_Ucell* function was used to calculate single-cell scores for the Hallmark signatures. The enriched (adjusted p-value less than 0.05) and non-redundant gene sets, together with the significant conserved markers overlapping these gene sets, defined the transcriptional signatures distinguishing the DA subpopulations. The reference HGSOC model was generated with Seurat's *RunPCA* (8 PCs) and *RunUMAP* dimension reductions using the integrated model and the identified DA subpopulations and applied in the downstream analysis for mapping the single-cell patient HGSOC data.

Patient HGSOC Mapping

HGSOC single-cell RNA sequencing datasets published by Izar et al.⁴⁴ were downloaded from the Gene Expression Omnibus (GEO) database (GSE146026) and analyzed with Seurat v4.1.1^{8,47} after transforming the normalized count scale into Seurat's default scale using a 10k scale factor and the natural log. Seurat's reference-based transfer of single-cell cluster information^{8,26,47} was applied to

predict Center and Surface DA subpopulation labels in the ascites cells collected from the HGSOC patients. *FindTransferAnchors*, *TransferData*, and *MapQuery* functions were run with the *k.anchors* parameter set to 20, CCA dimension reduction, and the UMAP projection model. A subset of ascites cells with maximal transfer scores greater than 0.5 and a difference from the next largest score of at least 0.25 were used for validation of the transcriptional signatures' characteristic of the reference DA subpopulations.

GSEA analysis

Pre-ranked GSEA was performed using the R Bioconductor *fgsea* package and the hallmark collection of the Molecular Signature Database (MSigDB version 6.2).^{38,42,71,72} Genes were ranked based on the average expression in a cell cluster relative to the global average. Permutation p-values for the enrichment scores were calculated based on 10000 gene set-wise runs and significantly enriched gene sets were identified with false discovery rate adjusted p-values less than 0.05. Common GSEA leading edge (LE) gene sets were extracted for comparison of functional enrichment across ovarian spheroids, organoids, and biopsies.

GSVA analysis

GSVA implemented in the R *cerebroApp* package was applied to determine gene set activation (MSigDB version 6.2 collections) in each cell.^{38,39,73} The GSVA scores were calculated with Gaussian kernel and the maximum difference parameters on Seurat log normalized data and highly variable genes. The analysis utilized the computational resources of the NIH HPC Biowulf cluster (<http://hpc.nih.gov>).