Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Multi-omics analysis of distinct hybrid epithelial/mesenchymal states in ovarian carcinosarcoma infers functional diversity and therapeutic vulnerabilities

**Holly Barker**

barker.h@wehi.edu.au

The Walter and Eliza Hall Institute of Medical Research https://orcid.org/0000-0002-9350-5520

**Anthony Hadla**

The Walter and Eliza Hall Institute of Medical Research

**Xueyi Dong**

The Walter and Eliza Hall Institute of Medical Research https://orcid.org/0000-0003-1136-3117

**Andrew Farrell**

The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC, Australia

**Casey Anttila**

Walter and Eliza Hall Institute https://orcid.org/0000-0003-1391-7160

**Ruvimbo Mishi**

The Walter and Eliza Hall Institute of Medical Research https://orcid.org/0000-0002-8166-1155

**Ling Ling**

The Walter and Eliza Hall Institute of Medical Research

**Daniela Zalcenstein**

Walter and Eliza Hall Institute https://orcid.org/0000-0003-4928-1846

**Jiayu Wang**

The Walter and Eliza Hall Institute of Medical Research https://orcid.org/0009-0003-1973-4346

**Vineet Vaibhav**

Walter and Eliza Hall Institute of Medical Research

**Jumana Yousef**

Walter and Eliza Hall Institute of Medical Research https://orcid.org/0000-0002-2364-6299

**Laura Gianni**

The Walter and Eliza Hall Institute of Medical Research

**Gayanie Ratnayake**

The Royal Women's Hospital

**Cassandra Vandenberg**

The Walter and Eliza Hall Institute https://orcid.org/0000-0002-8891-8873

**Rory Bowden**
The Walter and Eliza Hall Institute of Medical Research https://orcid.org/0000-0001-8596-0366
**Yunshun Chen**
The Walter and Eliza Hall Institute of Medical Research https://orcid.org/0000-0003-4911-5653
**Clare Scott**
Walter and Eliza Hall Institute of Medical Research https://orcid.org/0000-0002-3689-5956

**Article**

**Keywords:**

**Posted Date:** November 16th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-11545684/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations: Yes** there is potential Competing Interest. C.L.S. declares Advisory Boards for AstraZeneca, Clovis Oncology, Roche, Eisai Inc, Sierra Oncology, Takeda, MSD and Grant/Research support from Eisai Inc, AstraZeneca, Boehringer Ingelheim and IDEAYA Biosciences. C. Vandenberg and H.E. Barker report research support (paid to institution) from AstraZeneca Pty Ltd and Boehringer Ingelheim outside the submitted work. Other authors declare no conflicts of interest.

## Title

# Multi-omics analysis of distinct hybrid epithelial/mesenchymal states in ovarian carcinosarcoma infers functional diversity and therapeutic vulnerabilities


## Authors and Affiliations

Anthony Hadla[1,2], Xueyi Dong[1,2], Andrew Farrell[1], Casey J. A. Anttila[1,2], Ruvimbo D. Mishi[1], Ling Ling[1], Daniela Amann-Zalcenstein[1,2], Jiayu Wang[1], Vineet Vaibhav[1], Jumana Yousef[1], Laura F. Gianni[1,2], Gayanie Ratnayake[3], Cassandra J. Vandenberg[1,2], Rory Bowden[1,2**], Yunshun Chen[1,2**], Clare L. Scott[1,3,4,5**] and Holly E. Barker[1,2**#]

[1] The Walter and Eliza Hall Institute of Medical Research (WEHI), Parkville, VIC, 3052, Australia

[2] Department of Medical Biology, The University of Melbourne, Parkville, VIC, 3052, Australia

[3] The Royal Women's Hospital, Parkville, VIC, 3052, Australia

[4] Department of Medical Oncology, Peter MacCallum Cancer Centre, Melbourne, VIC, 3050, Australia

[5] Department of Obstetrics, Gynaecology and Newborn Health, The University of Melbourne, Parkville, VIC, 3052, Australia

[**] senior authors contributed equally

[#] corresponding author: barker.h@wehi.edu.au



## Abstract

Ovarian carcinosarcoma (OCS) is a rare and aggressive malignancy characterised by epithelial (carcinomatous) and mesenchymal (sarcomatous) components. This biphasic nature arises through epithelial-to-mesenchymal transition (EMT), which is a continuum involving transient and reversible states intermediate between full epithelial and mesenchymal extremes. Hybrid epithelial/mesenchymal (E/M) cell states display elevated epithelial-to-mesenchymal plasticity, and are likely drivers of tumour initiation, metastasis, immune evasion, and drug resistance in many cancer types. Therapeutic targeting of highly drug-resistant mesenchymal cells has resulted in discordant outcomes, potentially due to the enrichment of these remarkably plastic and aggressive partial EMT (p-EMT) states, which may represent promising targets for novel therapies.

Using single-nucleus RNA sequencing and spatial transcriptomics, we characterised an extensive cohort of OCS patient-derived xenograft models to elucidate genes and signalling pathways associated with hybrid E/M cell states. We demonstrated that genes implicated in DNA replication and repair, cell cycle progression, and neurotransmitter signalling were over-represented in intermediate EMT cell states, underpinning their enhanced proliferative and dynamic nature. These p-EMT states were also enriched for genes involved in multiple signal transduction pathways previously implicated in EMP, including MAPK and Wnt signalling. As EMT regulation in cancer is complex and multi-layered, we used flow cytometry and cell surface proteomics to identify surface proteins associated with p-EMT states in OCS, including DPP4 as a prominent candidate for therapeutic targeting, alongside B7-H4/VTCN1, MUC15, and GGT1. We highlight novel therapeutic approaches for individuals with OCS and, potentially, other highly aggressive and drug-resistant cancers, where EMT is a key feature.

## Introduction

Ovarian carcinosarcoma (OCS), formerly termed malignant mixed Müllerian tumour (MMMT) of the ovary, is the most lethal ovarian malignancy, accounting for 1-5% of ovarian cancer diagnoses and associated with a median overall survival of less than two years[1,2]. OCS is unique amongst ovarian cancers, in that it is composed of epithelial (carcinomatous) and mesenchymal (sarcomatous) components, whereas other epithelial ovarian cancers (EOCs), like high grade serous ovarian carcinoma (HGSOC), are entirely carcinomatous[3]. This biphasic nature of OCS reflects a monoclonal origin, with the sarcomatous component arising from carcinomatous cells via epithelial-to-mesenchymal transition (EMT)[4].

EMT, and the reverse process, mesenchymal-to-epithelial transition (MET), underlie epithelial-to-mesenchymal plasticity (EMP), which describes the potential of cells to transition between epithelial and mesenchymal states[5]. These mechanisms are important for many biological processes, including gastrulation in early embryogenesis, wound healing, and tissue regeneration[6]. EMP is also commonly hijacked by cells that have undergone neoplastic transformation and can drive tumour progression, particularly metastasis and drug resistance. Previously, EMT was

considered a binary process, where cells displayed either an epithelial or mesenchymal phenotype; however, it is now appreciated that some cells acquire mesenchymal-like properties whilst maintaining a level of epithelial marker expression[7]. These hybrid epithelial/mesenchymal (E/M) cell states have been identified across a wide variety of cancer types, including gynaecological cancers, and generally display enhanced tumorigenic, metastatic, and immune evasive features compared to cells that have undergone a more complete EMT, whilst still contributing to chemoresistance[7].

Currently, there are few approved targeted therapies for OCS, with first-line treatment recommendations including cytoreductive surgery and platinum- and taxane-based chemotherapy, as for other EOCs[3]. Unfortunately, OCS often responds poorly to these treatments, with more women developing disease progression following chemotherapy, compared with HGSOC[2]. As the carcinomatous component of OCS frequently displays a high-grade serous pathology, these tumours are generally associated with similar genetic aberrations[3,8]. This implicates the unique sarcomatous component of OCS, and the process that drives its formation, as the major promoter of the elevated chemoresistance observed in these tumours. Further, the metastatic lesions of OCS are often enriched for carcinomatous cells, suggesting that the epithelial-like component facilitates dissemination of neoplastic cells and colonisation of secondary tumours[2,3]. Therefore, the biphasic components of OCS may drive distinct facets of tumour progression that are inextricably linked to EMP, making it an attractive therapeutic target.

Reversing EMT has proven difficult, with no drug directly targeting EMT in cancer having received FDA approval to date[9]. Previous attempts have focused on targeting mesenchymal cells or EMT-inducing factors, but these approaches may promote the enrichment of highly plastic partial EMT (p-EMT) states with unintended enhanced tumour-initiating capacity and metastatic potential[10]. Therefore, it may be prudent to target these hybrid E/M cell states directly; however, previous studies have mostly employed bulk transcriptomics to investigate EMP in cancer, which masks subpopulation-specific signals and fails to capture the true heterogeneity of neoplastic cells.

Consequently, this study used single-nucleus RNA sequencing (snRNA-seq) and spatial transcriptomics (ST) to investigate an extensive cohort of OCS patient-derived xenograft (PDX) models with distinct pathological and molecular characteristics. Single-cell transcriptomic analysis of OCS remains limited, with only a few studies published to date, each examining one or two

primary samples, and none exploring p-EMT[11-14]. Although these provided valuable insights into OCS biology, individual cases cannot capture the full heterogeneity observed in the clinic. Using our comprehensive dataset, we established previously unreported OCS-specific epithelial- and mesenchymal-like gene signatures, enabling the interrogation of gene expression profiles of hybrid E/M cell states in OCS. As EMP is a complex process regulated at transcriptional, post-transcriptional, translational, post-translational, and epigenetic levels, flow cytometry and cell surface proteomics were also used to identify proteins enriched in these hybrid E/M cell states[15]. Collectively, this multi-omics analysis identified genes and signalling pathways associated with p-EMT states, which represent promising targets for novel therapies against OCS and, potentially, other cancers in which hybrid E/M cells drive lethal metastasis and drug resistance.

## Materials and Methods

### Patient-derived xenograft (PDX) tumours

Animal experiments were conducted under the approval of the WEHI Animal Ethics Committee (AEC 2019.024, 2022.030, and 2022.038). OCS PDX models #105, #177, #233, #264, PH003, PH142, PH419, and PH592 have been described previously[4,16-18]. PH PDX models were established through transplanting fragments of cryopreserved tumour tissue subcutaneously from PDXs generated in the Mayo Clinic (USA) (Supplementary Table S1). The remaining PDX models (#105, #177, #233, #264, #436, and #718) were established through transplanting fragments of tumour tissue obtained from patients consented to the WEHI-Stafford Fox Rare Cancer Program (Melbourne, Australia) (Supplementary Table S1). Tumour fragments (1-4 $mm^3$) were transplanted subcutaneously into NOD-SCID-IL-2Rγ mice under anaesthesia. Tumour growth was monitored twice weekly and once a volume > 700 $mm^3$ was reached the mice were euthanised and tumours excised. PDX identity was confirmed regularly by short tandem repeat profiling.

### Histological staining

Formalin-fixed paraffin-embedded (FFPE) blocks from OCS PDX tumours were sectioned and stained with hematoxylin and eosin (H&E) and antibodies for immunohistochemistry (IHC), including anti-p53 (mouse, DO-7, Dako; #M700101), anti-PAX8 (rabbit, poly, Proteintech; #10336-1-AP), anti-WT-1 (rabbit, poly, Abcam; #AB15249), anti-Ki67 (mouse, MIB-1, Dako; #M7240), anti-pan-cytokeratin (mouse, AE1/3, Dako; #M3515), anti-vimentin (rabbit, D21H3,

Cell Signaling Technology; #5741), and anti-HMGA2 (rabbit, D1A7, Cell Signaling Technology; #8179). Stained sections were imaged with the SLIDEVIEW VS200 Universal Whole Slide Imaging Scanner (20X) and uploaded to 3DHISTECH SlideCenter (v3.3) or Aiforia for viewing.

Western blotting

Western blotting was performed as described previously[17]. Briefly, lysates were prepared from snap frozen tumour pieces from OCS PDX models. Proteins were separated on 4-12% NuPAGE™ Bis-Tris Protein Gels (Thermo Fisher Scientific) and transferred to PVDF membranes using the iBlot™ 2 Dry Blotting System (Thermo Fisher Scientific). Membranes were probed with the following antibodies: anti-E-cadherin (#3195, Cell Signaling Technology), anti-cytokeratin-18 (#SC-51582, Santa Cruz Biotechnology), anti-N-cadherin (#AB18203, Abcam), anti-vimentin (#5741, Cell Signaling Technology), anti-N-MYC (#51705, Cell Signaling Technology), anti-HMGA2 (#8179, Cell Signaling Technology), and anti-β-actin (#A5441, Sigma Aldrich).

Sample preparation for single-nucleus RNA sequencing (snRNA-seq)

After previous issues faced when preparing single nuclei from snap frozen solid OCS PDX tumours[19], we selected FFPE blocks from OCS PDX tumours for snRNA-seq. FFPE blocks were assessed for their RNA quality using the RNeasy® FFPE Kit for RNA Extraction (Qiagen) and the TapeStation (Agilent) (Supplementary Table S2). Two 25 µm scrolls from each FFPE block were deparaffinised, rehydrated, and both mechanically and enzymatically digested using the Chromium Next GEM Single Cell Fixed RNA Sample Preparation Kit (10x Genomics). Nuclei suspensions were filtered using a 30 µm pre-separation filter (Miltenyi Biotec) and counted with the Countess II Automated Cell Counter and DAPI nuclei stain.

snRNA-seq library construction and sequencing

All samples were processed using standardised protocols under optimised conditions to minimise technical variability between samples. Single-nuclei suspensions were processed according to the FLEX sample preparation protocol (10x Genomics) using the Chromium Fixed RNA Kit and the Dual Index Kit TS Set A. Gene expression libraries were purified with SPRIselect Bead-Based Reagent (Beckman Coulter) and sequenced using the Illumina NextSeq2000 Sequencing System.

Sample preparation for spatial transcriptomics (ST)

One 5 µm section from each FFPE block was placed on a Super Plus Microscope Slide (Fisher Scientific) and left to dry overnight in a desiccator. Using Visium CytAssist 6.5 mm tissue slides and cassettes and a Low Profile Thermocycler Adaptor, the slides were deparaffinised, H&E-stained, imaged using the Olympus V200 slide scanner (20X), and then de-crosslinked according to the Visium CytAssist Spatial Gene Expression for FFPE protocol (10x Genomics).

ST library construction and sequencing

Sections were processed further using the Visium Spatial Gene Expression for FFPE protocol (10x Genomics) with the Visium FFPE Reagent Kit, Visium Human Transcriptome Probe Kit, and Dual Index Kit TS Set A. Briefly, probes were hybridised to sections, ligated, and then released and extended using the Visium CytAssist instrument (10x Genomics). The cDNA was then amplified prior to the construction of the gene expression library, which was purified with SPRIselect Bead-Based Reagent (Beckman Coulter). Sequencing was performed using the Illumina NextSeq2000 Sequencing System.

Bioinformatics and statistical analysis of snRNA-seq and ST data

*Processing and quality control*

Reads from each sample were processed using 10x Genomics Cell Ranger software (v8.0.1) and BCL files were demultiplexed into FASTQ files. For snRNA-seq, sample-demultiplexed single-nuclei gene-count matrices were generated using the 10x Genomics pre-built GRCh38 reference transcriptome and Chromium Human Transcriptome Probe Set (v1.0.1). All subsequent analysis was performed in R (v4.4.1) with Bioconductor (v3.20). The DropletUtils package (v1.26.0)[20] was used to create objects from the Cell Ranger output directories and identify non-empty droplets. For ST, feature-barcode matrices were generated using Space Ranger software (v2.1.1) with manual CytAssist Microscope Image alignment by Loupe Browser (v8.1.1) used as an input for Space Ranger.

The Seurat R package (v5.3.0)[21] was used to process the data generated from snRNA-seq and ST analyses. Nuclei (snRNA-seq) or barcoded spots (ST) were filtered based on the number of features, unique molecular identifier counts (UMIs), and percentage of UMI counts from mitochondrial genes. Thresholds for each metric varied with the technology used and the sequencing depth (Supplementary Table S3). For each modality, raw-count Seurat objects from the

12 OCS PDX tumour samples were merged without batch integration and subset to retain only nuclei or spots passing sample-specific quality control. The merged objects were log-normalized, variable features selected (1,500 for snRNA-seq and 1,000 for ST), scaled, analyzed by principal component analysis (PCA), and embedded by uniform manifold approximation and projection (UMAP) using the first 30 principal components.

*Gene set analysis of pan-cancer epithelial and mesenchymal signatures*

Pan-cancer epithelial and mesenchymal signatures, generated from three independent publications in the literature[22-24], were validated using gene set analysis via the WEB-based Gene SeT AnaLysis Toolkit (WebGestalt; https://www.webgestalt.org/) (Supplementary Tables S4 and S5). Over-representation analysis (ORA) was performed with "*Homo sapiens*" as the organism of interest and "genome" as the reference set. For the gene ontology (GO) analysis, the "Biological Process noRedundant" and "Molecular Function noRedundant" functional databases were utilised, whereas for Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis, the "KEGG" functional database was used. The "weighted" set cover was also employed for both sets of analyses to further reduce redundancy. A GO term or KEGG pathway or network was considered significantly over-represented if associated with a false discovery rate (FDR) $\leq 0.05$. Pan-cancer epithelial and mesenchymal scores were calculated as the mean log-normalised expression of all genes within their respective signatures. The pan-cancer EMT score for each nucleus and barcoded spot were then generated by subtracting the epithelial score from the mesenchymal score.

*Derivation and validation of OCS-specific epithelial- and mesenchymal-like gene signatures*

The edgeR quasi-likelihood pipeline (v4.5.8)[25] was used to identify genes which displayed a linear relationship with the pan-cancer EMT scores. Genes that were found to have a negative correlation with this score were considered epithelial-linked, whereas those with a positive correlation were mesenchymal-linked. For a gene to be included in the OCS-specific epithelial- and mesenchymal-like signatures an FDR $< 0.01$, calculated using the Benjamini-Hochberg method, was required in all the samples which displayed a negative correlation between pan-cancer epithelial and mesenchymal scores. Any genes included in the OCS-specific signature also had to be detected in the nuclei and barcoded spots of all samples.

ORA of the OCS-specific signatures was performed in the same way as the pan-cancer signatures for validation. Loupe Browser (v8.1.2) was also used to visualise the spatial expression pattern of OCS-specific epithelial- and mesenchymal-like signatures on the ST sections stained with H&E. The level of enrichment for each signature was displayed as the average log-count-per-10,000 of all genes within individual signatures. The "AUCell" R package (v1.30.1)[26] was then used to calculate OCS-specific epithelial and mesenchymal-like gene signature scores based on log-counts per million for each gene contained within the signatures. OCS-specific EMT scores for each cell and barcoded spot were then calculated by subtracting the epithelial score from the mesenchymal score.

*Generation and pseudo-bulking of EMT bins*

Nuclei with < 3,000 UMIs were then removed before segregation of all nuclei and barcoded spots into five distinct EMT bins using OCS-specific EMT scores from low to high. Nuclei or barcoded spots in the same EMT bin across samples were pseudo-bulked by summing raw UMI counts for each gene. Samples where the library size of the nuclei or barcoded spots within a specific EMT bin were < 500,000 or < 3,000,000 UMI counts, respectively, were not considered for pseudo-bulking. Any gene in the pseudo-bulked snRNA-seq dataset with ≥ 10 UMI counts in at least six samples and a total count ≥ 15 UMIs was included in this analysis. In the pseudo-bulked ST dataset, any genes with ≥ 100 UMI counts in at least four samples and a total count ≥ 1,000 UMIs were also retained.

*Differential gene expression analysis of pseudo-bulked EMT bins*

Differential expression analysis was then performed in edgeR to compare gene expression in each pseudo-bulked EMT bin with the other four, with each appropriate pseudo-bulk sample contributing equally, regardless of the number of nuclei or barcoded spots associated with the specific EMT bin. "Marker genes" were defined as those with a positive $\log_2$-fold change and an FDR < 0.05 across all four comparisons. KEGG analysis for the marker genes of each distinct EMT bin was then performed using the "kegga" function in the R package limma (v3.64.3)[27]. Pathways or gene networks with an FDR < 0.05 were considered significantly over-represented.

Flow cytometry

Tumours from OCS PDX models were mechanically digested using surgical blades and enzymatically digested using the human Tumour Dissociation Kit in gentleMACS™ C tubes on a gentleMACS™ Dissociator (Miltenyi Biotec) according to manufacturer's instructions. The single-cell suspension was then filtered through a 100 µm cell strainer (Corning) before red blood cells were depleted via hypotonic lysis, using red blood cell lysis buffer (WEHI).

The flow cytometry screen was performed using the LEGENDScreen™ Human PE Kit (BioLegend), according to the manufacturer's protocol. Viably frozen cells from different digested tumours were first stained with BV421-MHC class 1 (H-2k[d]) (#742431, BD Biosciences), Alexa Fluor 647-E-cadherin (#324112, BioLegend), and PE-Cy7-N-cadherin (#350811, BioLegend) before adding to screen plates (Supplementary Table S6). Cells were then stained with 0.1 µg/mL Propidium Iodide (PI) (Sigma-Aldrich) and run on a Cytek® Aurora Spectral Cytometer. Data analysis was carried out using FlowJo software (v1.0.7.1).

Cell surface proteomics

Tumours from OCS PDX models were mechanically digested using surgical blades and filtered through a pre-wetted 100 µL cell strainer (Corning). Cell pellets were resuspended in oxidation buffer consisting of 1x DPBS, 1 mM sodium metaperiodate (Thermo Fisher Scientific), 200 µM aminooxy-biotin (Biotium), and 10 mM aniline (Sigma-Aldrich) to label surface proteins. Labelling efficiency was determined by staining with BV786-Streptavidin (#563858, BD Biosciences) and acquiring on a Cytek® Aurora Spectral Cytometer, with data analysis performed using FlowJo software (v1.0.7.1) (Supplementary Table S7).

The remaining cells were lysed, nuclei removed, and lysates enriched for biotinylated proteins using high-capacity streptavidin agarose resin (#20357, Thermo Fisher Scientific). Washes were carried out in 6.0 M urea buffer with 100 mM Tris-HCl at pH 8.5 before alkylating in 50 mM iodoacetamide. On-bead cell surface biotinylated proteins were enzymatically digested with a 50 mM ammonium bicarbonate buffer containing 1 µg trypsin (EMS0004, Sigma-Aldrich). Peptides were collected, dried, and reconstituted in 0.1% formic Acid (FA)/2% acetonitrile (ACN) for mass spectrometry analysis. Reconstituted peptides were separated using reverse-phase liquid chromatography on a 15 cm C18 fused silica column with an integrated emitter tip (IonOpticks, Australia) within a NewVanquish UHPLC coupled to an Orbitrap Astral mass spectrometer (Thermo Fisher Scientific). Data was acquired in data independent acquisition (DIA) mode with

the settings outlined in Supplementary Table S8. The global mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[28] partner repository with the dataset identifier PXD085406 [Reviewer Token: 7x8N0nUEbyOA].

Proteomics data analysis

Database searching of proteomic data files was performed with Spectronaut software (v18)[29] against the Human reviewed database (UniProt, August 2024) using the BGS standard settings. Cysteine carbamidomethylation was included as a fixed modification, and N-terminal acetylation and methionine oxidations were included as variable modifications. The FDR and the posterior error probability (PEP) cutoffs were set to less than 1% at the peptide and protein levels, and a minimum length of seven amino acids for peptides was specified. Enzyme specificity was set as C-terminal to arginine and lysine for trypsin proteases with a maximum of two missed cleavages. Proteotypic peptides were retained and filtered using a run-wise protein-group posterior error probability (PG.PEP (Run-Wise)) threshold of $\leq 0.05$. Protein-level abundances were inferred from the retained precursor intensities using the Accurate Protein-wide Label-free Quantification by Delayed Normalisation and Maximal Peptide Ratio Extraction (MaxLFQ) algorithm[30]. Proteins that were identified in 50% or more of the samples in at least one group were retained.

To address potential skewness in intensity data, a log2 transformation was applied to all protein intensity measurements. Subsequently, data normalisation was performed using cyclic loess method to correct for systematic technical variations across the experimental runs. Multivariate analysis, including PCA, was employed to identify potential outliers. Differential protein abundance was analysed using linear models implemented in limma (v3.62.1), with empirical Bayes moderation applied to the estimated variances. Adjustments for multiple comparisons were applied using the Benjamini-Hochberg procedure to control the FDR. A protein was deemed significantly differentially expressed if its FDR was $\leq 5\%$.

**Data availability**

Data is available upon request. A data transfer agreement is required.

**Results**

OCS PDX models recapitulated the pathological and molecular heterogeneity of OCS

OCS displays significant pathological and genetic heterogeneity. To ensure any findings from this study would be both biologically and clinically relevant, it was critical that the PDX models selected for gene expression profiling reflected this variability. As such, 12 OCS PDX models were characterised by pathological review, based on hematoxylin and eosin (H&E) staining and immunohistochemistry (IHC). The mutational profiles of the original patient samples and pre-clinical models where available have been presented previously[4,16-18], with a summary provided in Supplementary Table 9. Frequent pathogenic variants (PVs) were present in *TP53*, with less common PVs observed in other tumour suppressor genes, including *NF1*, *RB1*, *FBXW7*, and *PTEN*. Among genes implicated in DNA repair, there was a *BRCA1* germline variant in #264 and somatic *BRCA2* aberrations in #436 and PH142. Copy number changes were also present in genes associated with cell cycle progression, including amplification of *CCNE1* in #105 and a small copy number increase of *EGFR* (five copies) in PH142.

In conjunction with H&E staining, the epithelial and mesenchymal markers, pan-cytokeratin and vimentin, respectively, were used to estimate the proportion of carcinoma and sarcoma in each tumour. Pan-cytokeratin was clearly expressed in #105, #233, #264, #436, PH003mixed, PH142, and PH419, whereas vimentin was detected in all tumours, albeit at low levels in #233 and #436 (Figure 1A). From this, it was determined that #436 was entirely carcinomatous, whereas #177, #718, and PH003sarcoma were composed of only sarcomatous cells (Figure 1B). The PH592sensitive and PH592refractory tumours, which exhibited differing sensitives to cisplatin, were characterised as undifferentiated high-grade tumours and the remaining six models all had mixed morphology[4]. All tumours also displayed strong diffuse p53 positivity by IHC except for PH142, which contained a *TP53* PV that resulted in a truncated protein (null staining), and #233, the only model with genetically wild-type *TP53* (Figure 1A; Supplementary Table S9).

With EMT being a key feature of OCS, Western blotting for common epithelial and mesenchymal markers was performed (Figure 1C). All tumours expressed both epithelial and mesenchymal proteins except #718, which was reviewed as entirely sarcomatous (Figures 1B and C). The fully carcinomatous model, #436, displayed elevated expression of E-cadherin and Cytokeratin-18 and minimal expression of ZEB1, N-cadherin, and vimentin. Notably, the mesenchymal markers N-MYC and HMGA2 were expressed highly in this tumour. There were also models, like PH142 and

PH419, which exhibited significant proportions of carcinomatous and sarcomatous elements and expressed all the markers assessed. There did not appear to be a particular pattern, with some tumours expressing high levels of both epithelial and mesenchymal proteins, some tumours expressing low levels of both, and some expressing opposing levels of epithelial and mesenchymal markers as expected. Collectively, the 12 PDX tumours reflected the pathological and molecular heterogeneity of OCS and were thus suitable for snRNA-seq and ST analysis.

Interrogation of OCS PDX models using single-nucleus RNA sequencing (snRNA-seq) and spatial transcriptomics (ST)

FFPE blocks from each of the 12 OCS PDX tumours were then analysed using snRNA-seq (10x Genomics Flex) and ST (10x Genomics Visium CytAssist) (Supplementary Figure S1A). Pooled samples were de-multiplexed using the associated probes, and individual nuclei (snRNA-seq) or barcoded spots (ST) were filtered based on UMI counts, number of features, and percentage of UMIs from mitochondrial genes (Supplementary Table S3). This ensured that only high-quality regions of tissue were included for downstream analyses whilst also helping to exclude regions of mouse origin from the PDX models. This was important because, although the probes used for snRNA-seq and ST are human-specific, some overlap remains between human and mouse transcripts. In total, 112,533 cells were retained in the snRNA-seq dataset (1,163 – 22,117 cells per sample), whereas 42,624 barcoded spots (1,842 – 4,809 barcoded spots per sample) were retained in the ST dataset (Supplementary Table S3). One sample (PH003mixed) in the snRNA-seq dataset was removed from downstream analysis, as only 373 cells remained following filtering. UMAP analysis demonstrated that samples clustered largely according to tumour of origin across both datasets, with increased transcriptional heterogeneity observed between tumours than within tumours (Supplementary Figures S1B and C).

Pan-cancer epithelial and mesenchymal gene signatures were enriched for genes associated with epithelial and mesenchymal features

We used previously published pan-cancer epithelial and mesenchymal gene signatures to identify the most epithelial- and mesenchymal-like populations across the OCS PDX samples (Supplementary Tables S4 and S5)[22-24]. To confirm that the pan-cancer signatures reflected core features of epithelial and mesenchymal cells, over-representation analysis (ORA) was performed to identify gene ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG)

pathways or networks significantly associated with each signature. In the pan-cancer epithelial gene set there was over-representation of epithelial-linked terms, including “epithelial cell development” and “epidermis development”, and enrichment of genes implicated in cell junctions (“cell junction assembly”, “cell-to-cell junction organisation”, and “tight junctions”), which are a hallmark of epithelial cells (Supplementary Figure 2A - C). Contrastingly, the pan-cancer mesenchymal signature exhibited over-representation of GO terms and KEGG pathways or networks related to the extracellular matrix (ECM), such as “extracellular structure organisation”, “cell-substrate adhesion”, “proteoglycans in cancer”, and “ECM-receptor interaction” (Supplementary Figures S2D - F). There was also significant enrichment of genes involved in “gastrulation”, “response to transforming growth factor beta”, and “metallopeptidase activity”, all of which are associated with EMT. This confirmed that the combined pan-cancer signatures maintained core properties of epithelial and mesenchymal cells.

These signatures were then used to generate epithelial and mesenchymal scores for all nuclei (snRNA-seq) and barcoded spots (ST) in the tumours. Although the distribution of scores was similar across both datasets, there was an increased spread observed in the snRNA-seq dataset (Supplementary Figure S3). This reflected the single-nucleus resolution of these gene expression profiles, compared to ST on the Visium platform, which measures average gene expression of all cells within a 50 μm region. Generally, these scores aligned well with tumour pathology, with largely carcinomatous tumours (#264 and #436) associated with intermediate epithelial and low mesenchymal scores, whereas the fully sarcomatous models (#177, #718, and PH003sarcoma) displayed low epithelial and intermediate-to-high mesenchymal scores (Figure 1B; Supplementary Figure S2). There was also a negative association between epithelial and mesenchymal scores in #105, #233, PH142, and PH419, indicating that these were mixed samples composed of cells ranging from epithelial- to mesenchymal-like states. This supported findings of clearly defined carcinomatous and sarcomatous elements in these models (Figure 1B).

Pan-cancer EMT scores were then generated for every nucleus and barcoded spot across the samples, and a very stringent analysis pipeline was used to establish new OCS-specific gene signatures. This required a gene to be detected in all 12 tumours assessed via snRNA-seq and ST and exhibit a significant association with low or high pan-cancer EMT scores in all four mixed models (#105, #233, PH142, and PH419) across both datasets. From this, 432-gene epithelial-like

and 187-gene mesenchymal-like signatures specific to OCS were derived (Supplementary Tables S10 and S11).

OCS-specific gene signatures contained mostly distinct genes compared to pan-cancer signatures

The OCS-specific gene signatures were then validated using ORA, with enrichment of genes related to cell junctions identified in the epithelial-like gene set, including "cell-to-cell junction organisation", "cell junction assembly", "tight junctions", and "adherens junctions" as well as others related to adhesion ("positive regulation of cell adhesion") and polarity ("establishment or maintenance of cell polarity"), which are key features of epithelial cells (Figure 2A - C). Contrastingly, the OCS-specific mesenchymal-like gene signature included an over-representation of genes implicated in "extracellular structure organisation", "collagen metabolic process", "glycosaminoglycan binding", and "response to transforming growth factor beta", all of which were also enriched in the pan-cancer mesenchymal gene set (Figure 2D - F; Supplementary Figure S2D and E). New GO terms also arose, including "bone development", "ossification", and "connective tissue development" (Figure 2D). Additionally, genes associated with the "PI3K-Akt signalling pathway" were significantly associated with this signature, as well as others related to the ECM ("proteoglycans in cancer"), cytoskeleton ("regulation of actin cytoskeleton"), or interactions between both ("focal adhesion") (Figure 2F). This confirmed that the OCS-specific gene signatures reflected the typical characteristics of epithelial and mesenchymal cells.

To confirm that the OCS-specific gene signatures were not overly biased towards the pan-cancer gene sets used to derive them, the gene lists were compared. There were 63 genes common to the pan-cancer epithelial and OCS-specific epithelial-like signatures, including *CDH1*, *CLDN4*, *CLDN7*, *OCLN*, and *KRT8*, all of which play key roles in tight junctions and adherens junctions (Figure 2G; Supplementary Table S12). There were 37 genes shared between the pan-cancer mesenchymal and OCS-specific mesenchymal-like signatures, such as key EMT-inducing transcription factors (*SNAI1*, *TWIST1*, *ZEB1*, and *ZEB2*) and collagen-related genes (*COL1A1*, *COL1A2*, *COL3A1*, *COL5A2*, *COL6A1*, *COL6A2*, and *COL6A3*) (Figure 2G; Supplementary Table S13). Therefore, core epithelial and mesenchymal genes were maintained in the OCS-specific gene signatures, but the largely distinct gene sets confirmed that they were not biased and reflected the unique biology of OCS.

OCS-specific gene signatures aligned with cell morphology and identified regions of p-EMT

The OCS-specific epithelial- and mesenchymal-like gene signatures were used to spatially map carcinomatous and sarcomatous regions and compare results with morphology previously determined by pathological assessment. For #105, PH142, and PH419, which are composed of clearly demarcated carcinomatous and sarcomatous elements, there was close alignment with epithelial and mesenchymal cells identified with the OCS-specific gene signatures (Figure 2H and I). In more monomorphic tumours, such as #264 (largely carcinomatous) or #177 and PH003sarcoma (largely sarcomatous), there was enrichment of the epithelial- or mesenchymal-like signatures, respectively. Interestingly, #436 displayed a similar level of epithelial- and mesenchymal-like signature expression across the tumour despite histological features being consistent with an entirely carcinomatous classification, supporting the existence of p-EMT states.

Spatial definition of early, mid and late p-EMT cell states in OCS PDX

As EMT represents a spectrum of states, it is unlikely that only a single hybrid E/M population exists. To define early, mid and late p-EMT cells, EMT scores were used to sort each nucleus (snRNA-seq) and barcoded spot (ST) into distinct EMT bins. Five EMT bins were defined, with EMT 1 and EMT 5 being the most epithelial- and mesenchymal-like states, respectively, and EMT 2, EMT 3, and EMT 4 representing the early, mid and late p-EMT states, respectively. The distributions of OCS-specific EMT scores across the datasets were assessed and as expected, the mixed tumours (#105, #233, PH142, and PH419) displayed the broadest spread (Figure 3A and B). Notably, these distributions were shifted to the left in the snRNA-seq dataset compared to the ST dataset, indicative of lower EMT scores, likely due to differences in sample preparation, specifically the enzymatic dissociation required for the snRNA-seq samples. As PH142 exhibited the widest distribution of OCS-specific EMT scores across both datasets, it was selected to derive the threshold values used for the segregation of cells and barcoded spots into the five distinct EMT bins (Figure 3A and B).

The spatial location of barcoded spots representing the five EMT bins was also mapped for each sample. As each barcoded spot is 50 μm in diameter, it was possible to observe shifts in EMT across incredibly small regions of tissue, revealing spatial heterogeneity. For instance, #105 contained barcoded spots associated with EMT 1 at the outer edges of one piece of tumour tissue, whereas they were located more centrally in other pieces of tissue, indicating significant intratumoral spatial heterogeneity was present (Figure 3C). #233 and PH419 similarly displayed

barcoded spots associated with EMT 1 in both central and peripheral regions of the tumour sections. Across the samples, barcoded spots in EMT 1 were rarely directly adjacent to those in EMT 5, highlighting that there was a gradual shift in EMT scores in neighbouring barcoded spots. There were, however, transitions from EMT 1 to EMT 5 (via intermediate EMT bins) within five or six barcoded spots in some sections of the mixed tumours, indicating that this shift in EMT could occur over incredibly small regions of tissue. Therefore, distinct p-EMT cell states could be identified in OCS PDX samples using the OCS-specific gene signatures, and they displayed significant spatial heterogeneity.

Marker genes of distinct hybrid E/M cell states indicated these cells had different functions

Marker genes enriched in each pseudo-bulked EMT bin were identified by differential gene expression analysis, followed by KEGG analysis to identify over-represented pathways or gene networks (Figure 4; Supplementary Tables S14 - S23). We focused on the snRNA-seq dataset as this provided gene expression profiles at single-nucleus resolution; however, the same analysis was performed on the ST dataset for comparison (Supplementary Figure S4; Supplementary Tables S24-S33). To limit sample-specific biases, each pseudo-bulked sample contributed equally, regardless of the number of nuclei associated with each bin in a sample. In the EMT 2 bin, KEGG analysis identified unique over-representation of many pathways or networks associated with DNA replication and repair, such as “DNA replication” (17 genes), “homologous recombination” (10 genes), “base excision repair” (10 genes), “nucleotide excision repair” (10 genes), and “mismatch repair” (9 genes) (Figure 4B; Supplementary Table S20). In the EMT 3 bin, KEGG analysis revealed unique over-representation of neural-related pathways or networks, such as “neuroactive ligand-receptor interaction” (24 genes), “neuroactive ligand signalling” (14 genes), “cholinergic synapse” (nine genes), and “GABAergic synapse” (eight genes) (Figure 4C; Supplementary Table S21). *GNB5*, *ADCY2*, *CHRNA4*, *CHRNA7*, *GNG4*, *GNG10*, and *GABRG3* were implicated in three out of four of these pathways (Supplementary Table S16). Further, “hedgehog signalling” was uniquely over-represented in the EMT 3 bin and this was associated with five different marker genes, *GPR161*, *GLI1*, *DHH*, *SCUBE2*, and *PTCH2* (Supplementary Table S16). In the EMT 4 bin, KEGG analysis identified unique over-representation of “glucagon signalling”, with six different marker genes (*ATF2*, *FOXO1*, *PHKA1*, *PRKAG3*, *PYGL*, and *CAMK2A*) implicated in this pathway (Figure 4D; Supplementary Tables S17 and S22).

There were also KEGG pathways or networks significantly associated with multiple EMT bins, as marker genes of these distinct bins were implicated in common pathways. This included "cell cycle" (EMT 2 and EMT 3) "Fanconi anemia" (EMT 2 and EMT 3), "aldosterone synthesis and secretion" (EMT 3 and EMT 4), "insulin signalling pathway" (EMT 1, EMT 2, and EMT 4), "hormone signalling" (EMT 3 and EMT 5), "Wnt signalling pathway" (EMT 1, EMT 4, and EMT 5), "MAPK signalling pathway" (EMT 1, EMT 4, and EMT 5), "calcium signalling pathway" (EMT 4 and EMT 5), and "AGE-RAGE signalling pathway for diabetic complications" (EMT 4 and EMT 5) (Figure 4; Supplementary Tables S20 - S22). Overall, cells representing hybrid E/M states in OCS contained an over-representation of genes implicated in DNA replication and repair, cell cycle, hormone signalling, neurotransmitter signalling, and MAPK, Wnt, and calcium signalling.

Hybrid E/M cell states in OCS PDX models were enriched for some stem cell markers but not murine p-EMT markers

Next, we wanted to identify cell surface proteins specific to hybrid E/M cells. CD51/ITGAV, CD61/ITGB3, and CD106/VCAM, have previously been established as cell surface markers of hybrid E/M cells in a genetically-engineered mouse model (GEMM) of skin squamous cell carcinoma and MMTV-PyMT breast cancer[31]. These markers also define different p-EMT cell states in a GEMM of OCS we have described previously[32] (identification of p-EMT cell states is unpublished data). There has been increasing evidence of an association between hybrid E/M states in cancer and stemness, making common ovarian cancer stem cell markers, such as CD24, CD44, CD117/c-Kit, and CD133/PROM1, also worthy of investigation[7,10,31,33].

To assess whether these cell surface markers could define different p-EMT cell populations, a flow cytometry screen was performed using the LEGENDScreen$^{TM}$ Human PE kit. Single-cell suspensions from digested OCS PDX tumours were pre-stained with antibodies specific for E-cadherin and N-cadherin before being independently stained with antibodies specific for 354 cell surface markers, including these markers of interest (Supplementary Figure S5A; Supplementary Table S34). To account for inefficient staining with the N-cadherin antibody, marker enrichment was not assessed solely using percentage of positive cells but also the distribution of cells based on fluorescence intensity (Supplementary Figure S5B). CD51/ITGAV, CD24, and CD44 were almost ubiquitously expressed by all cell types (epithelial-like (E-cadherin$^{+}$/N-cadherin$^{-}$), hybrid

E/M (E-cadherin$^{+}$/N-cadherin$^{+}$), or mesenchymal-like (E-cadherin$^{-}$/N-cadherin$^{+}$)) isolated from the OCS PDX tumours (Figure 5A, 5E and F). CD61/ITGB3 was detected at lower levels across all cell populations, whereas CD106/VCAM1 was only expressed in the E-cadherin$^{-}$/N-cadherin$^{+}$ cells, indicative of being a mesenchymal marker (Figure 5B and C). Notably, CD117/c-Kit was enriched within E-cadherin$^{+}$ populations, with 76.1% and 70.3% positivity in E-cadherin$^{+}$/N-cadherin$^{-}$ and E-cadherin$^{+}$/N-cadherin$^{+}$, respectively, whilst only being detected in 30.8% of cells expressing N-cadherin alone (Figure 5F). This enrichment pattern was not as clear for CD133/PROM1, but there was a positive shift in fluorescence intensity observed in E-cadherin$^{+}$ cells (Figure 5G).

To validate these findings, cell surface proteomics using liquid chromatography with tandem mass spectrometry (LC-MS/MS) was performed on OCS PDX tumours that were largely carcinomatous ("epithelial-like"), sarcomatous ("mesenchymal-like"), or composed of significant proportions of both elements ("mixed"). PCA confirmed that tumours representing each group displayed distinct proteomic profiles, with PC1 and PC2 explaining 53.9% and 19.5% of variance between samples, respectively (Supplementary Figure S6A). Core epithelial and mesenchymal proteins were assessed, with E-cadherin and EpCAM highly expressed in the epithelial-like tumours and vimentin and N-cadherin displaying similar expression levels across all tumours (Supplementary Figure S6B and C).

In total, 4,395 proteins were included in the analysis (Supplementary Table S35). As this was bulk proteomic analysis, proteins enriched in mixed OCS PDX tumours were considered potential markers of hybrid E/M states. The murine p-EMT and human stem cell markers assessed by flow cytometry were detected by cell surface proteomics except for CD24 which has a small protein core of 31-35 amino acids, making it difficult to identify via LC-MS/MS (Figure 5H and I)[34]. Otherwise, CD51/ITGAV was clearly enriched in the epithelial-like tumours and expressed moderately in mixed tumours. CD61/ITGB3 was similarly upregulated in the epithelial-like tumours, particularly #233, but at a reduced overall protein intensity, supporting results from the flow cytometry screen (Figure 5B, 5H and I). CD106/VCAM1 exhibited peak expression in the #718 mesenchymal-like tumours, but moderate levels were also detected in #436 (Figure 5H and I). Notably, this marker was also lowly expressed in most mixed tumours. Among the stem cell markers, CD44 was expressed inconsistently across models, with peak expression in #233 and

PH142, and limited levels detected in #177. CD117/c-Kit and CD133/PROM1 displayed similar patterns of expression, as observed via flow cytometry, exhibiting elevated expression in the epithelial-like tumours (Figure 5F - I). These markers were also detected at intermediate levels in mixed tumours, with CD133/PROM1 being more highly expressed. Additionally, both markers were lowly expressed in the mesenchymal-like tumours, with CD117/c-Kit not detected at all in #177, which displayed minimal expression of all the stem cell markers assessed (Figure 5H).

Collectively, flow cytometry and cell surface proteomics analysis revealed that CD51/ITGAV, CD61/ITGB3, and CD106/VCAM1, which had previously been established as markers of intermediate EMT states in murine cancer models, may not be associated with hybrid E/M states in human OCS, with the latter likely representing a mesenchymal-linked marker. Some stem cell markers, particularly CD117/c-Kit and CD133/PROM1, were enriched in E-cadherin$^{+}$ cells, including hybrid E/M states, and in the epithelial-like and mixed OCS PDX tumours, suggestive of a potential link with intermediate EMT states in OCS.

Dipeptidyl peptidase 4 (CD26/DPP4) was associated with hybrid E/M states in human OCS

Using the flow cytometry screen and cell surface proteomics analysis, other novel surface markers associated with intermediate EMT states in OCS could also be identified (Supplementary Table S36). These markers may drive the establishment or maintenance of hybrid E/M states and thus would be potential therapeutic targets.

Seven markers were enriched in the E-cadherin$^{+}$/N-cadherin$^{+}$ cells of OCS, being CD19, CD26/DPP4, CD169/SIGLEC1, CD50/ICAM3, CD180/RP105, CD203c/ENPP3, and B7-H4/VTCN1 (Figure 6A-C; Supplementary Figure S7). Two of these (CD26/DPP4 and B7-H4/VTCN1) have previously been shown to be expressed by ovarian tumour cells, whilst the others were well-characterised immune markers[35,36]. These proteins were also detected by LC-MS/MS. Dual expression of E-cadherin and N-cadherin suggests this could be evidence of immune mimicry rather than maintenance of human leukocytes in the PDX tumours (Figure 6D and E). Notably, CD26/DPP4 was also one of the 34 proteins significantly enriched within the mixed OCS PDX models, relative to the epithelial- and mesenchymal-like tumours via cell surface proteomics and of those exhibited the highest protein intensity across these mixed tumours (Figure 6D - F).

Among these 34 proteins, were 21 transmembrane proteins, including GGT1, MUC15, AMIGO2, and CELSR1 (Figure 6F – H; Supplementary Table S36). Additionally, all except LYPD2 were associated with more than one detected peptide, which was important when assessing reliability of LC-MS/MS. Overall, flow cytometry and cell surface proteomics analysis identified several surface markers enriched in either E-cadherin$^{+}$/N-cadherin$^{+}$ cells or mixed OCS PDX tumours, respectively. Of these proteins, CD26/DPP4 was particularly compelling as it was identified as a marker of interest across both independent analyses.

**Discussion**

Tumour cells that undergo p-EMT display elevated plasticity, which is inextricably linked to their enhanced tumorigenic, metastatic, and drug-resistant potential[7]. These hybrid E/M cell states are thus attractive therapeutic targets in cancer. Due to the biphasic nature of OCS, p-EMT is a key feature and likely explains why this cancer is so aggressive. Therefore, OCS provides an ideal model for investigating hybrid E/M cell states. Here, we presented a comprehensive multi-omics analysis of an extensive cohort of OCS PDX models, reflective of the significant pathological and molecular heterogeneity associated with these tumours.

The OCS PDX models analysed in this study harboured molecular features previously reported to occur in OCS, including near ubiquitous inactivating PVs in *TP53* and less common PVs in genes implicated in DNA repair and cell cycle regulation[37]. Significant variation in the proportion of carcinomatous and sarcomatous elements was observed across the tumours, by pathological review and expression of EMP-related proteins, which supported the significant heterogeneity that exists in this cohort and increased confidence that any findings would be both biologically and clinically relevant.

As hybrid E/M cell states are difficult to identify histologically, with dramatic morphological changes associated with EMT often occurring late in the transition, snRNA-seq and ST were employed to derive OCS-specific signatures for the most epithelial- and mesenchymal-like states. The epithelial-like gene set was over-represented for genes implicated in tight junctions and adherens junctions. Tight junctions are composed of transmembrane proteins, including occludins, claudins, and junctional adhesion molecules, and help maintain close cell-to-cell interactions between epithelial cells[38]. They also establish a semi-permeable barrier which regulates paracellular transport and establishes apical-basal polarity, a key feature of epithelial cells that is

generally lost during EMT[39]. Adherens junctions also promote cell adhesion via strong interactions between E-cadherin proteins expressed on the surface of adjacent epithelial cells[40]. As E-cadherin is downregulated and N-cadherin is upregulated during EMT, bonds that form between N-cadherin proteins are far more dynamic, which leads to mesenchymal-like cells with enhanced migratory and invasive properties. In contrast, the mesenchymal-like signature was enriched for ECM-related terms, linked to extracellular structure, collagen, proteoglycans, and glycosaminoglycan binding, as expected for cells more dependent on ECM remodelling[41]. There was also over-representation of genes implicated in response to Transforming growth factor beta (TGF-β) signalling, a master regulator of EMT across many cancer types, including ovarian cancer[42,43]. These included genes involved in latent TGF-β regulation activation (*LTBP1*, *FBN1*, and *THBS1*), TGF-β co-receptor signalling (*ENG* and *TGFBR3*), downstream EMT transcriptional regulation (*ZEB1* and *ZEB2*), and ECM deposition and remodelling (*COL1A1*, *LOX*, and *DDR2*)[44,45]. Unique enrichment of genes associated with ossification, bone development, and connective tissue development in the OCS-specific mesenchymal-like signature also supported previous findings that the sarcomatous component of OCS can be heterologous and composed of cells not native to the ovary, including malignant bone and cartilage[46].

In general, OCS-specific epithelial- and mesenchymal-like gene signatures aligned well with the carcinomatous and sarcomatous elements of OCS PDX tumours, respectively. One exception was #436, which was entirely carcinomatous based on morphology but displayed similar levels of expression for both signatures. This provided further support for the conversion theory of OCS development which postulates that the carcinoma cells convert to sarcoma cells via EMT[3,47]. Our data indicate that cells are genetically primed to undergo this transformation prior to significant morphological changes associated with EMT being visible.

The OCS-specific signatures were used to generate EMT scores which enabled the segregation of cells and barcoded spots across tumours into distinct EMT bins. There was a skewing of the EMT scores towards lower values across the mixed samples in the snRNA-seq dataset, relative to the ST dataset. Although both technologies were performed on FFPE tissue, there is an additional enzymatic dissociation step during tissue processing for snRNA-seq. This step results in the digestion of collagen in the ECM. Mesenchymal cells are generally surrounded by a denser ECM, characterised by extensive collagen cross-linking, which could be more resistant to enzymatic

dissociation[48]. This could explain the increased representation of epithelial-like states (i.e. skewing of the data towards lower EMT scores) in the snRNAseq-dataset.

Distinct EMT states displayed significant spatial heterogeneity in the ST data. Previous studies have shown that EMT in cancer can be spatially-resolved based on the position of the leading edge[49]. Central cells generally retain epithelial-like properties associated with growth and differentiation, whereas cells at the leading edge often adopt a mesenchymal-like state to promote migration and invasion[50,51]. In our data, although the most epithelial (EMT 1) and mesenchymal (EMT 5) barcoded spots were always separated by barcoded spots from intermediate bins (EMT 2, EMT 3, and EMT 4), sometimes these epithelial barcoded spots were in the core and other times they were located at the edge.

Identifying marker genes of the early, mid and late p-EMT bins 2, 3 and 4, respectively, indicated enrichment of genes related to DNA replication, repair, and cell cycle regulation in the early p-EMT cell states. Therefore, cells in an early p-EMT state are likely to be the most proliferative in OCS, potentially driving tumour initiation and recurrence. With DNA repair mechanisms necessary to maintain genomic stability of highly proliferative tumour cells, targeting these pathways may be prudent. For instance, snRNA-seq analysis revealed WEE1, a negative regulator of cell cycle progression which inhibits entry into mitosis at the G2/M checkpoint to allow for DNA repair, as an EMT 2 bin marker gene[52]. Our lab has showed previously that the WEE1 inhibitor (WEE1i), PD-407824, increased OCS cell killing in the presence of cisplatin, one of the commonly used platinum agents in OCS therapy[17]. In human OCS organoids, the WEE1i, adavosertib, was also one of the most effective drugs in combination with cisplatin. Recent studies have also identified that ovarian cancer cells with Cyclin E1 over-expression display increased sensitivity to WEE1i, and as observed in our PDX cohort, *CCNE1* amplification was a relatively common genetic aberration in OCS[53].

The mid p-EMT cell state (EMT 3 bin) was enriched for genes implicated in neuroactive signalling pathways, supporting the highly dynamic and plastic nature of these p-EMT states. One EMT 3 marker gene in the ST dataset, *APLNR*, encodes a G protein-coupled receptor that binds Apelin and has been implicated in the modulation of angiogenesis, proliferation, migration, and invasion in cancer[54]. In HGSOC, over-expression of *APLNR* was significantly associated with a worse prognosis, whilst in cell lines under hypoxic conditions that induces EMT, there was an increased

expression of *APLNR*[55]. Further, in colon cancer, increased activation scores for the neuroactive ligand-receptor interaction pathway were significantly associated with a worse prognosis and reduced predicted responses to immunotherapy[56]. This supports the mid p-EMT cell states in OCS being highly plastic and immune evasive.

The late p-EMT cell state (EMT 4 bin), as well as EMT 1 and 5, were enriched for genes involved in Mitogen-activated protein kinase (MAPK) signalling, which regulates EMT via interactions with the TGF-β signalling pathway[57]. TGF-β signalling can induce the expression of ERK, and this contributes to the breakdown of adherens junctions in epithelial cells and promotes migration[58]. Other MAPKs like p38 can also be upregulated via TGF-β signalling and have been shown to maintain E-cadherin expression, which explains the association with both epithelial- and mesenchymal-like states[59]. MAPK marker genes specifically enriched in EMT 4 included *ATF2*, which encodes a transcription factor associated with numerous pathways linked to p-EMT states in OCS, including glucagon signalling, aldosterone synthesis and secretion, and cGMP-PKG signalling, as well as other EMP-related pathways, such as PI3K/AKT, Relaxin, and TNFα signalling. In HGSOC, p38 and ATF2 positivity were significantly correlated with elevated lymph node metastasis, high histopathological stage, and reduced overall survival, supporting ATF2 as a potential target in OCS[60]. Our lab has also recently shown that OCS is particularly reliant on MAPK signalling for growth and survival, and that KRAS-mutant OCS cell lines and organoids were sensitive to a combination of the MEK inhibitor, mirdametinib, with the dual EGFR/RAF inhibitor, BGB-283[17]. Notably, increased synergy was also observed with eribulin, a microtubule inhibitor which we have shown reverses EMT in preclinical models of OCS and is currently being evaluated in combination with pembrolizumab in the EPOCH clinical trial for relapsed ovarian and uterine carcinosarcomas (NCT05619913)[4].

In the late p-EMT states, there was also over-representation of genes involved in Wnt signalling, with *WIF1*, *FRZB*, and *NKD1* identified as EMT 4 marker genes in the snRNA-seq and ST datasets. These encode proteins which suppress Wnt signalling[61,62]. This was concordant with previous studies showing that hybrid E/M stability factors, such as OVOL2 and GRHL2, can also suppress Wnt signalling[63,64] Moreover, one of the few scRNA-seq analyses previously performed on OCS identified that *WIF1* was part of a mesenchymal sub-cluster of genes that differentiated OCS from HGSOC[12]. Therefore, WIF1 may have a unique role in maintaining hybrid E/M states in OCS that

are primed to undergo sarcomatous transformation. Pathway responsive genes for activity inference (PROGENy) analysis, which measures transcriptional activity of specific pathways, rather than just over-representation of associated genes, could be useful to confirm downregulation of Wnt signalling in the cells in the EMT 4 bin[65].

At a protein level, previously established murine hybrid E/M markers (CD51/ITGAV, CD61/ITGB3, and CD106/VCAM1) were not enriched in E-cadherin$^{+}$/N-cadherin$^{+}$ cells or mixed OCS PDX tumours. However, CD106/VCAM1 did seem to be associated with mesenchymal cells, supporting previous studies which demonstrated CD106 can induce an EMT in colorectal and ovarian cancer[66,67]. Analysis of ovarian cancer stem cell markers indicated that CD24 was expressed almost ubiquitously by OCS cells via flow cytometry and were not associated just with hybrid E/M cells. CD44, CD117/c-Kit, and CD133/PROM1, on the other hand, were elevated in epithelial-like and mixed tumours in the cell surface proteomics analysis, relative to the more sarcomatous models. Collectively, this opposes a previously long-held belief that stemness is more associated with the mesenchymal-like state. In fact, the now widespread use of single-cell transcriptomic methods to study EMP has revealed that stem cell-like features are more associated with hybrid E/M cell states in cancer[7,10,31,33]. Moreover, CD133$^{+}$ cells have been shown previously to represent highly tumorigenic hybrid E/M cell states in ovarian carcinomas[33]. Overall, the stem cell markers CD44, CD117, and CD133 are worthy of future exploration as potential cell surface markers of p-EMT states in OCS.

Other novel cell surface proteins associated with p-EMT states were also elucidated. Of the seven surface markers enriched in E-cadherin$^{+}$/N-cadherin$^{+}$ cells via flow cytometry, five were well-characterised immune markers, potentially representing a form of immunogenic mimicry. This has been proposed as a mechanism of tumour plasticity that contributes to immune evasion and drug resistance, supporting the marker gene findings from the snRNA-seq/ST data for the mid p-EMT cell state[68]. The other two markers, B7-H4/VTCN1 and CD26/DPP4, have previously been detected in both immune cells and ovarian tumour cells[35,36]. B7-H4/VTCN1 is an immune checkpoint molecule, which has been shown to inhibit T-cell proliferation and the secretion of cytokines, providing further support for a role of p-EMT cell states in facilitating immune evasion[69]. Moreover, CD26/DPP4, which was the only marker of interest shared across the flow cytometric and cell surface proteomics analyses, has a context-dependent role in EMP, with it

shown to promote EMT in papillary thyroid and endometrial carcinoma cells, whilst suppressing EMT in ovarian and breast cancer cells[70-73]. This is consistent with what would be expected of hybrid E/M cells which display enhanced EMP and the ability to freely transition between epithelial- and mesenchymal-like states[10]. Using a syngeneic mouse model of ovarian cancer, it has been demonstrated that the CD26/DPP4 inhibitor sitagliptin, which is currently approved for use against type II diabetes, led to reduced metastatic burden, increased activation, proliferation, and recruitment of $CD8^+$ T cells, and improved overall survival, relative to controls[74,75]. This recruitment of CD8+ T-cells may be mediated by CXCL10 binding to CXCR3, with CXCL10 a known target of the DPP4 proteolytic cleavage activity[76]. Sitagliptin is being assessed in a clinical trial for progressive gliomas (NCT07003542) and is worthy of further investigation in OCS. Other surface markers identified as enriched in the mixed OCS PDX tumours via cell surface proteomics included EMT-inducers (AMIGO2 and GPNMB), inhibitors of EMT (ACE2), and others, like CD26/DPP4, GGT1 and MUC15, which as mentioned above, exhibit a context-dependent role in EMP[77-80]. MUC15 was particularly interesting as it displayed peak protein expression in the four most mixed OCS PDX models (#105, #233, PH142, and PH419), whilst being minimally expressed in the other models. This included #233, which despite being categorised as an epithelial-like tumour, did contain some sarcomatous cells.

Future investigations should focus on exploring drugs that target some of these elements, including sitagliptin (CD26/DPP4) and PTX-100 (GGT1), in cell lines and PDX-derived organoids, as well as in OCS PDX models established in human immune system mice, to determine their impact on cell viability and functional potential while accounting for potential immune-mediated effects[75,81]. At least five B7-H4/VTCN1-directed antibody-drug conjugates (ADCs), including puxitatug samrotexan (AZD8205), are in active clinical development, dominated by TOP1 inhibitor payloads, as well as auristatin-based microtubule inhibitor payloads and could be evaluated in OCS as a means of targeting hybrid E/M cell states given the enrichment of B7-H4/VTCN1 in E-$cadherin^+$/N-$cadherin^+$ cells[82].

Collectively, this study used snRNA-seq, ST, flow cytometry, and cell surface proteomics to investigate hybrid E/M states in an extensive cohort of OCS PDX models. We revealed several genetic elements and signalling pathways associated with these intermediate EMT states in OCS, which have been shown to facilitate tumour initiation, immune evasion, drug resistance, metastasis

and tumour recurrence, in a wide range of cancer types, including ovarian cancer. This work provided a significant step towards identifying novel therapeutic strategies for people with OCS and, potentially, other cancers in which EMT supports or drives lethal metastasis and drug resistance.

## References


1 Mano, M. S. *et al.* Current management of ovarian carcinosarcoma. *International Journal of Gynecological Cancer* **17**, 316-324 (2007).

2 Hollis, R. L. *et al.* Ovarian carcinosarcoma is a distinct form of ovarian cancer with poorer survival compared to tubo-ovarian high-grade serous carcinoma. *British journal of cancer* **127**, 1034-1042 (2022).

3 Barker, H. E. & Scott, C. L. in *Seminars in cancer biology.* 110-120 (Elsevier).

4 Ho, G. Y. *et al.* Epithelial-to-Mesenchymal transition supports ovarian carcinosarcoma tumorigenesis and confers sensitivity to microtubule targeting with eribulin. *Cancer research* **82**, 4457-4473 (2022).

5 Haerinck, J., Goossens, S. & Berx, G. The epithelial–mesenchymal plasticity landscape: principles of design and mechanisms of regulation. *Nature Reviews Genetics* **24**, 590-609 (2023).

6 Kalluri, R. & Weinberg, R. A. The basics of epithelial-mesenchymal transition. *J Clin Invest* **119**, 1420-1428 (2009). https://doi.org/10.1172/JCI39104

7 Pastushenko, I. & Blanpain, C. EMT Transition States during Tumor Progression and Metastasis. *Trends Cell Biol* **29**, 212-226 (2019). https://doi.org/10.1016/j.tcb.2018.12.001

8 Zhao, S. *et al.* Mutational landscape of uterine and ovarian carcinosarcomas implicates histone genes in epithelial–mesenchymal transition. *Proceedings of the National Academy of Sciences* **113**, 12238-12243 (2016).

9 Bagheri, M. *et al.* Pharmacological induction of chromatin remodeling drives chemosensitization in triple-negative breast cancer. *Cell Rep Med* **5**, 101504 (2024). https://doi.org/10.1016/j.xcrm.2024.101504

10 Lüönd, F. *et al.* Distinct contributions of partial and full EMT to breast cancer malignancy. *Dev Cell* **56**, 3203-3221.e3211 (2021). https://doi.org/10.1016/j.devcel.2021.11.006

11 Regner, M. J. *et al.* A multi-omic single-cell landscape of human gynecologic malignancies. *Molecular cell* **81**, 4924-4941. e4910 (2021).

12 Xu, J. & Tu, M. Single-cell transcriptomics reveals tumor landscape in ovarian carcinosarcoma. *Journal of Zhejiang University-SCIENCE B* **25**, 686-699 (2024).

13 González-Martínez, S. *et al.* Single-nucleus RNA sequencing identifies a novel tenogenic heterologous differentiation in endometrial carcinosarcomas: implications for diagnosis and tumor classification. *The Journal of Pathology* **268**, 227-242 (2026).

14 Xie, K. *et al.* Dissecting the histological heterogeneity of ovarian carcinosarcoma and high-grade serous ovarian cancer in primary and metastatic tumors by single-cell transcriptomic analysis. *Chinese Medical Journal* **138**, 2195-2197 (2025).
15 Aiello, N. M. & Kang, Y. Context-dependent EMT programs in cancer metastasis. *Journal of Experimental Medicine* **216**, 1016-1026 (2019).
16 Nesic, K. *et al.* Farletuzumab ecteribulin and MORAb-109, folate receptor alpha and mesothelin targeting antibody–drug conjugates, show activity in poor prognosis gynaecological cancer models. *Clinical and Translational Medicine* **15**, e70274 (2025).
17 Farrell, A. *et al.* High-throughput drug screening identifies EGFR/MAPK pathway targeting sensitivities in organoid models of ovarian carcinosarcoma. *Journal of Experimental & Clinical Cancer Research* (2026).
18 Nesic, K. *et al.* BRCA1 secondary splice-site mutations drive exon-skipping and PARP inhibitor resistance. *Molecular cancer* **23**, 158 (2024).
19 Brown, D. V. *et al.* A risk-reward examination of sample multiplexing reagents for single cell RNA-Seq. *Genomics* **116**, 110793 (2024). https://doi.org/10.1016/j.ygeno.2024.110793
20 Lun, A. T. L. *et al.* EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol* **20**, 63 (2019). https://doi.org/10.1186/s13059-019-1662-y
21 Hao, Y. *et al.* Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol* **42**, 293-304 (2024). https://doi.org/10.1038/s41587-023-01767-y
22 Miow, Q. *et al.* Epithelial–mesenchymal status renders differential responses to cisplatin in ovarian cancer. *Oncogene* **34**, 1899-1907 (2015).
23 Mak, M. P. *et al.* A patient-derived, pan-cancer EMT signature identifies global molecular alterations and immune target enrichment following epithelial-to-mesenchymal transition. *Clinical Cancer Research* **22**, 609-620 (2016).
24 Gibbons, D. L. & Creighton, C. J. Pan-cancer survey of epithelial–mesenchymal transition markers across the Cancer Genome Atlas. *Developmental Dynamics* **247**, 555-564 (2018).
25 Chen, Y., Chen, L., Lun, A. T. L., Baldoni, P. L. & Smyth, G. K. edgeR v4: powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets. *Nucleic Acids Res* **53** (2025). https://doi.org/10.1093/nar/gkaf018
26 Aibar, S. *et al.* SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* **14**, 1083-1086 (2017). https://doi.org/10.1038/nmeth.4463
27 Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015). https://doi.org/10.1093/nar/gkv007
28 Perez-Riverol, Y. *et al.* The PRIDE database at 20 years: 2025 update. *Nucleic Acids Res* **53**, D543-D553 (2025). https://doi.org/10.1093/nar/gkae1011
29 Bruderer, R. *et al.* Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues. *Mol Cell Proteomics* **14**, 1400-1410 (2015). https://doi.org/10.1074/mcp.M114.044305

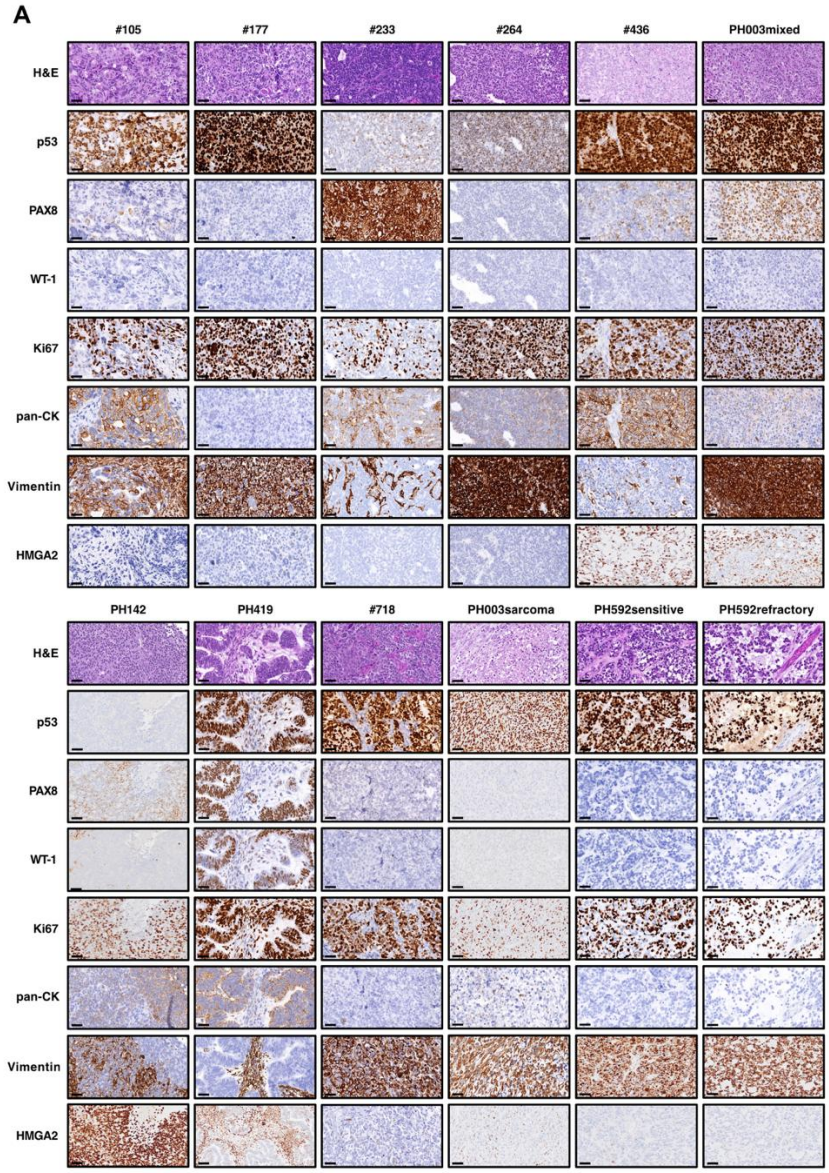
30 Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **13**, 2513-2526 (2014). https://doi.org/10.1074/mcp.M113.031591
31 Pastushenko, I. *et al.* Identification of the tumour transition states occurring during EMT. *Nature* **556**, 463-468 (2018).
32 Ho, G. Y. *et al.* Epithelial-to-mesenchymal transition supports ovarian carcinosarcoma tumorigenesis and confers sensitivity to microtubule-targeting with eribulin. *Cancer Res* (2022). https://doi.org/10.1158/0008-5472.CAN-21-4012
33 Strauss, R. *et al.* Analysis of epithelial and mesenchymal markers in ovarian cancer reveals phenotypic heterogeneity and plasticity. *PLoS One* **6**, e16186 (2011). https://doi.org/10.1371/journal.pone.0016186
34 Zhu, J., Nie, S., Wu, J. & Lubman, D. M. Target proteomic profiling of frozen pancreatic CD24+ adenocarcinoma tissues by immuno-laser capture microdissection and nano-LC–MS/MS. *Journal of Proteome Research* **12**, 2791-2804 (2013).
35 Gao, A. *et al.* Effect of VTCN1 on progression and metastasis of ovarian carcinoma in vitro and vivo. *Biomedicine & Pharmacotherapy* **73**, 129-134 (2015).
36 Moffitt, L. R. *et al.* Hypoxia regulates DPP4 expression, proteolytic inactivation, and shedding from ovarian cancer cells. *International Journal of Molecular Sciences* **21**, 8110 (2020).
37 Jarratt, A., Polidano, J., Scott, C. L. & Barker, H. E. Genomics of ovarian cancers and the potential of precision medicine. *Therapeutic Advances in Medical Oncology* **17**, 17588359251396651 (2025).
38 Shin, K., Fogg, V. C. & Margolis, B. Tight junctions and cell polarity. *Annu. Rev. Cell Dev. Biol.* **22**, 207-235 (2006).
39 Citi, S. *et al.* A short guide to the tight junction. *Journal of cell science* **137**, jcs261776 (2024).
40 Coopman, P. & Djiane, A. Adherens Junction and E-Cadherin complex regulation by epithelial polarity. *Cellular and molecular life sciences* **73**, 3535-3553 (2016).
41 Ishihara, S. & Haga, H. Matrix stiffness contributes to cancer progression by regulating transcription factors. *Cancers* **14**, 1049 (2022).
42 Kaminska, B., Wesolowska, A. & Danilkiewicz, M. TGF beta signalling and its role in tumour pathogenesis. *Acta Biochimica Polonica* **52**, 329-337 (2005).
43 Kumari, A. *et al.* TGFβ signaling networks in ovarian cancer progression and plasticity. *Clinical & Experimental Metastasis* **38**, 139-161 (2021).
44 Liu, S., Ren, J. & Ten Dijke, P. Targeting TGFβ signal transduction for cancer therapy. *Signal transduction and targeted therapy* **6**, 8 (2021).
45 Zhang, M. & Zhang, B. Extracellular matrix stiffness: mechanisms in tumor progression and therapeutic potential in cancer. *Experimental hematology & oncology* **14**, 54 (2025).
46 Prendiville, J., Murphy, D., Renninson, J., Buckley, H. & Crowther, D. Carcinosarcoma of the ovary treated over a 10-year period at the Christie Hospital. *International Journal of Gynecological Cancer* **4**, 200-205 (1994).
47 del Carmen, M. G., Birrer, M. & Schorge, J. O. Carcinosarcoma of the ovary: a review of the literature. *Gynecologic oncology* **125**, 271-277 (2012).

48 Henke, E., Nandigama, R. & Ergün, S. Extracellular matrix in the tumor microenvironment and its impact on cancer therapy. *Frontiers in molecular biosciences* **6**, 160 (2020).
49 Jolly, M. K. & Celià-Terrassa, T. Dynamics of phenotypic heterogeneity associated with EMT and stemness during cancer progression. *Journal of clinical medicine* **8**, 1542 (2019).
50 Brabletz, T. *et al.* Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment. *Proc Natl Acad Sci U S A* **98**, 10356-10361 (2001). https://doi.org/10.1073/pnas.171610498
51 Liu, S. *et al.* Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. *Stem Cell Reports* **2**, 78-91 (2014). https://doi.org/10.1016/j.stemcr.2013.11.009
52 Thangaretnam, K. *et al.* WEE1 inhibition in cancer therapy: Mechanisms, synergies, preclinical insights, and clinical trials. *Critical reviews in oncology/hematology*, 104710 (2025).
53 Xi, Q. *et al.* Cyclin E1 overexpression sensitizes ovarian cancer cells to WEE1 and PLK1 inhibition. *Oncogene*, 1-12 (2025).
54 Agarwala, Y. & Sharma, R. The role of apelin in cancer. *Frontiers in Oncology* **15**, 1683865 (2025).
55 Neelakantan, D. *et al.* Multifunctional APJ pathway promotes ovarian cancer progression and metastasis. *Molecular Cancer Research* **17**, 1378-1390 (2019).
56 Yang, Y. *et al.* Inhibition of neuroactive ligand–receptor interaction pathway can enhance immunotherapy response in colon cancer: an in silico study. *Expert review of anticancer therapy* **23**, 1205-1215 (2023).
57 Gui, T., Sun, Y., Shimokado, A. & Muragaki, Y. The roles of mitogen-activated protein kinase pathways in TGF-β-induced epithelial-mesenchymal transition. *Journal of signal transduction* **2012**, 289243 (2012).
58 Zavadil, J. *et al.* Genetic programs of epithelial cell plasticity directed by transforming growth factor-β. *Proceedings of the National Academy of Sciences* **98**, 6686-6691 (2001).
59 Strippoli, R. *et al.* p38 maintains E-cadherin expression by modulating TAK1–NF-κB during epithelial-to-mesenchymal transition. *Journal of cell science* **123**, 4321-4331 (2010).
60 Song, W.-J., Dong, Y., Luo, C. & Chen, Y.-Y. p38MAPK family isoform p38α and activating transcription factor 2 are associated with the malignant phenotypes and poor prognosis of patients with ovarian adenocarcinoma. *Pathology-Research and Practice* **213**, 1282-1288 (2017).
61 Liu, J. *et al.* Wnt/β-catenin signalling: function, biological mechanisms, and therapeutic opportunities. *Signal transduction and targeted therapy* **7**, 3 (2022).
62 Angonin, D. & Van Raay, T. J. Nkd1 functions as a passive antagonist of Wnt signaling. *PloS one* **8**, e74666 (2013).
63 Ye, G.-D. *et al.* OVOL2, an inhibitor of WNT signaling, reduces invasive activities of human and mouse cancer cells and is down-regulated in human colorectal tumors. *Gastroenterology* **150**, 659-671. e616 (2016).
64 Cieply, B., Farris, J., Denvir, J., Ford, H. L. & Frisch, S. M. Epithelial–mesenchymal transition and tumor suppression are controlled by a reciprocal feedback loop between ZEB1 and Grainyhead-like-2. *Cancer research* **73**, 6299-6309 (2013).

65 Schubert, M. *et al.* Perturbation-response genes reveal signaling footprints in cancer gene expression. *Nature communications* **9**, 20 (2018).
66 Zhang, D. *et al.* VCAM1 promotes tumor cell invasion and metastasis by inducing EMT and transendothelial migration in colorectal cancer. *Frontiers in oncology* **10**, 1066 (2020).
67 Huang, J. *et al.* VCAM1 expression correlated with tumorigenesis and poor prognosis in high grade serous ovarian cancer. *American journal of translational research* **5**, 336 (2013).
68 Tímár, J., Honn, K. V., Hendrix, M. J., Marko-Varga, G. & Jalkanen, S. Newly identified form of phenotypic plasticity of cancer: immunogenic mimicry. *Cancer and Metastasis Reviews* **42**, 323-334 (2023).
69 Zhou, L. *et al.* The role of B7-H4 in ovarian cancer immunotherapy: current status, challenges, and perspectives. *Frontiers in Immunology* **15**, 1426050 (2024).
70 Hu, X. *et al.* DPP4 gene silencing inhibits proliferation and epithelial-mesenchymal transition of papillary thyroid carcinoma cells through suppression of the MAPK pathway. *Journal of Endocrinological Investigation* **44**, 1609-1623 (2021).
71 Yang, F. *et al.* Inhibition of dipeptidyl peptidase-4 accelerates epithelial–mesenchymal transition and breast cancer metastasis via the CXCL12/CXCR4/mTOR axis. *Cancer research* **79**, 735-746 (2019).
72 Yang, X. *et al.* DPPIV promotes endometrial carcinoma cell proliferation, invasion and tumorigenesis. *Oncotarget* **8**, 8679 (2017).
73 Liu, C. *et al.* Down-regulation of DPP4 by TGFβ1/miR29a-3p inhibited proliferation and promoted migration of ovarian cancer cells. *Discover Oncology* **14**, 195 (2023).
74 Wilson, A. L. *et al.* DPP4 inhibitor sitagliptin enhances lymphocyte recruitment and prolongs survival in a syngeneic ovarian cancer mouse model. *Cancers* **13**, 487 (2021).
75 Scott, L. J. Sitagliptin: a review in type 2 diabetes. *Drugs* **77**, 209-224 (2017).
76 Gorrell, M. D. Dipeptidyl peptidase IV and related enzymes in cell biology and liver disorders. *Clinical Science* **108**, 277-292 (2005).
77 Bansal, A. *et al.* Gamma-Glutamyltransferase 1 Promotes Clear Cell Renal Cell Carcinoma Initiation and Progression. *Mol Cancer Res* **17**, 1881-1892 (2019). https://doi.org/10.1158/1541-7786.Mcr-18-1204
78 Wang, Q. *et al.* Tumor and serum gamma-glutamyl transpeptidase, new prognostic and molecular interpretation of an old biomarker in gastric cancer. *Oncotarget* **8**, 36171 (2017).
79 Chen, T. *et al.* MUC 15 promotes osteosarcoma cell proliferation, migration and invasion through livin, MMP-2/MMP-9 and Wnt/β-catenin signal pathway. *Journal of Cancer* **12**, 467 (2021).
80 Wu, S. *et al.* MUC15 loss facilitates epithelial-mesenchymal transition and cancer stemness for prostate cancer metastasis through GSK3β/β-catenin signaling. *Cellular Signalling* **84**, 110015 (2021).
81 Chew, T. *et al.* Phase 1 Pharmacodynamic and Pharmacokinetic Study of the Geranylgeranyltransferase I Inhibitor PTX-100 (GGTI-2418) in Patients with Advanced Malignancies. *Blood* **142**, 1705 (2023).
82 Kinneer, K. *et al.* Design and preclinical evaluation of a novel B7-H4–directed antibody–drug conjugate, AZD8205, alone and in combination with the PARP1-selective inhibitor AZD5305. *Clinical Cancer Research* **29**, 1086-1101 (2023).

**Figures and Figure Legends:**

## Figure 1




B

| PDX model | Tumour pathology |
|---|---|
| #105 | 50% carcinoma; 50% sarcoma |
| #177 | 100% sarcoma |
| #233 | 80% carcinoma; 20% sarcoma |
| #264 | 95% carcinoma; 5% sarcoma |
| #436 | 100% carcinoma |
| #718 | 100% sarcoma |
| PH003mixed | Admixed carcinoma/sarcoma |
| PH003sarcoma | 100% sarcoma |
| PH142 | 40% carcinoma; 60 %sarcoma |
| PH419 | 70% carcinoma; 30% sarcoma |
| PH592sensitive | High grade undifferentiated tumour |
| PH592refractory | High grade undifferentiated tumour |

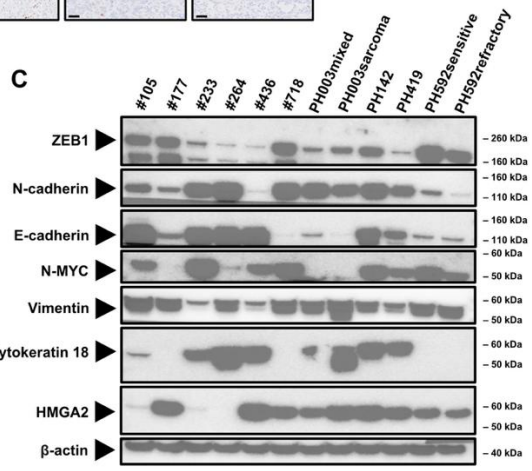
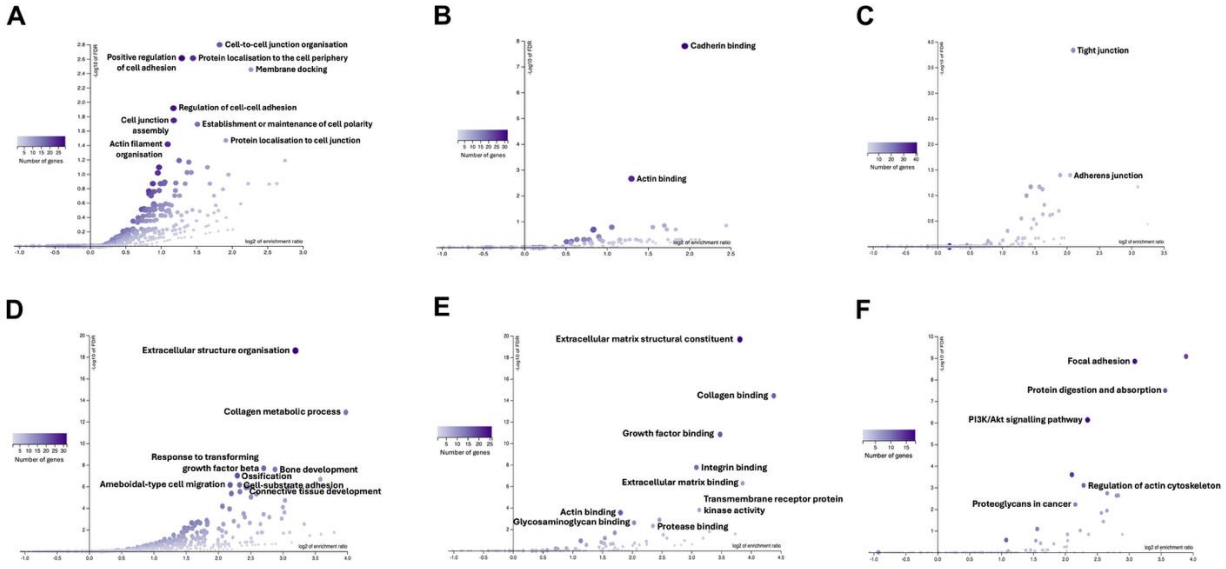


**Figure 1: PDX models selected for gene expression profiling reflected the significant pathological and molecular heterogeneity of OCS**

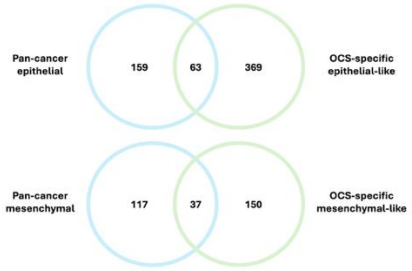
**(A)** Representative images of H&E- and IHC-stained tumour sections from 12 OCS PDX models. Scale bar = 50 μm. **(B)** Proportion of carcinomatous and sarcomatous elements in each tumour based on pathological review using H&E and IHC staining. **(C)** Western blot assessing the expression of common epithelial and mesenchymal proteins across the 12 OCS PDX tumours. β-actin served as the protein loading control.

PDX, patient-derived xenograft; OCS, ovarian carcinosarcoma; H&E, hematoxylin and eosin; IHC, immunohistochemistry

## Figure 2




G




H

| PDX cohort | Tumour pathology |
|---|---|
| #105 | Carcinoma and sarcoma present |
| #177 | Uniformly similar tumour – favour sarcoma |
| #233 | Carcinoma and sarcoma present |
| #264 | Likely all carcinoma |
| #436 | Likely all carcinoma |
| #718 | Entire tumour looks like a carcinoma |
| PH003mixed | Poorly differentiated carcinoma or a sarcoma |
| PH003sarcoma | All of this looks like a sarcoma |
| PH142 | Carcinoma and sarcoma present |
| PH419 | Carcinoma and sarcoma present |
| PH592sensitive | All of this looks like poorly differentiated carcinoma |
| PH592refractory | Uniformly similar tumour – favour sarcoma |

I

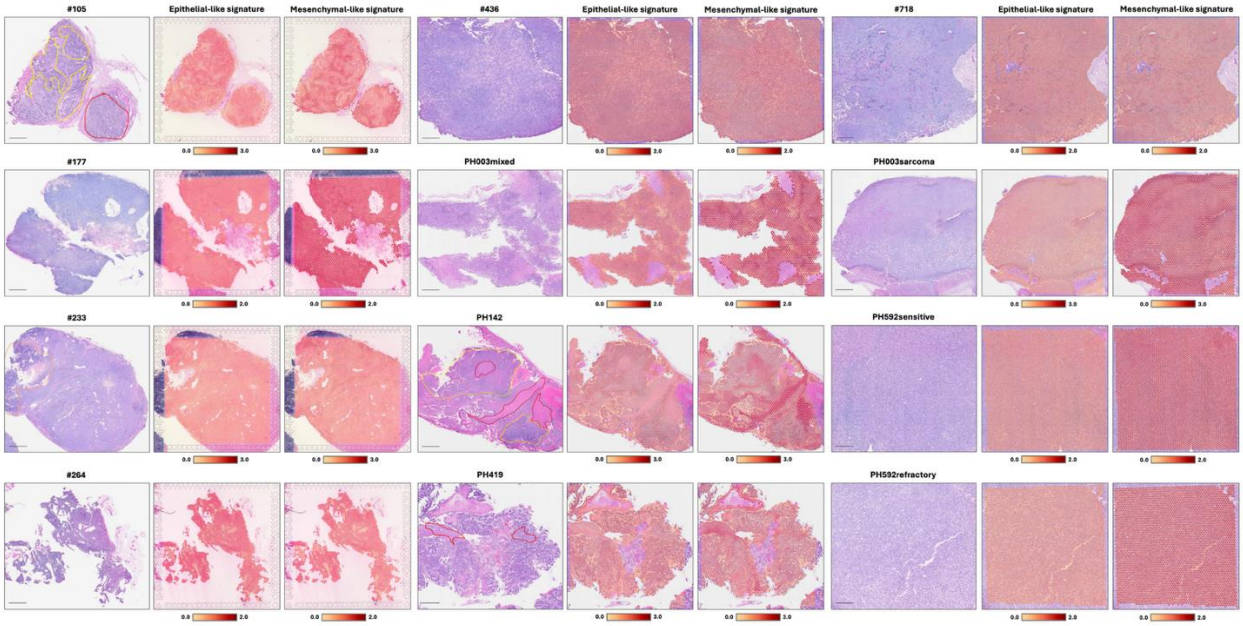


**Figure 2: OCS-specific epithelial- and mesenchymal-like signatures link epithelial and mesenchymal cell biology with tumour pathology**

**(A-C)** Significantly over-represented (FDR < 0.05) GO terms related to **(A)** biological process and **(B)** molecular function, and **(C)** KEGG pathways in the OCS-specific epithelial-like gene signature. **(D-F)** Significantly over-represented (FDR < 0.05) GO terms related to **(D)** biological process and **(E)** molecular function, and **(F)** KEGG pathways in the OCS-specific mesenchymal-like gene signature. **(G)** Venn diagram comparing the genes in the pan-cancer and OCS-specific gene signatures. **(H)** Pathological review of sections from OCS PDX tumours used for ST analysis based on H&E staining. **(I)** Alignment between tumour pathology and mean log-normalised (LogNorm) expression values of OCS-specific epithelial- and mesenchymal-like signature using Loupe Browser (v8.1.1). Scale bar = 1 mm.

OCS, ovarian carcinosarcoma; FDR, false discovery rate; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; ST, spatial transcriptomics; H&E, hematoxylin and eosin

**Figure 3**

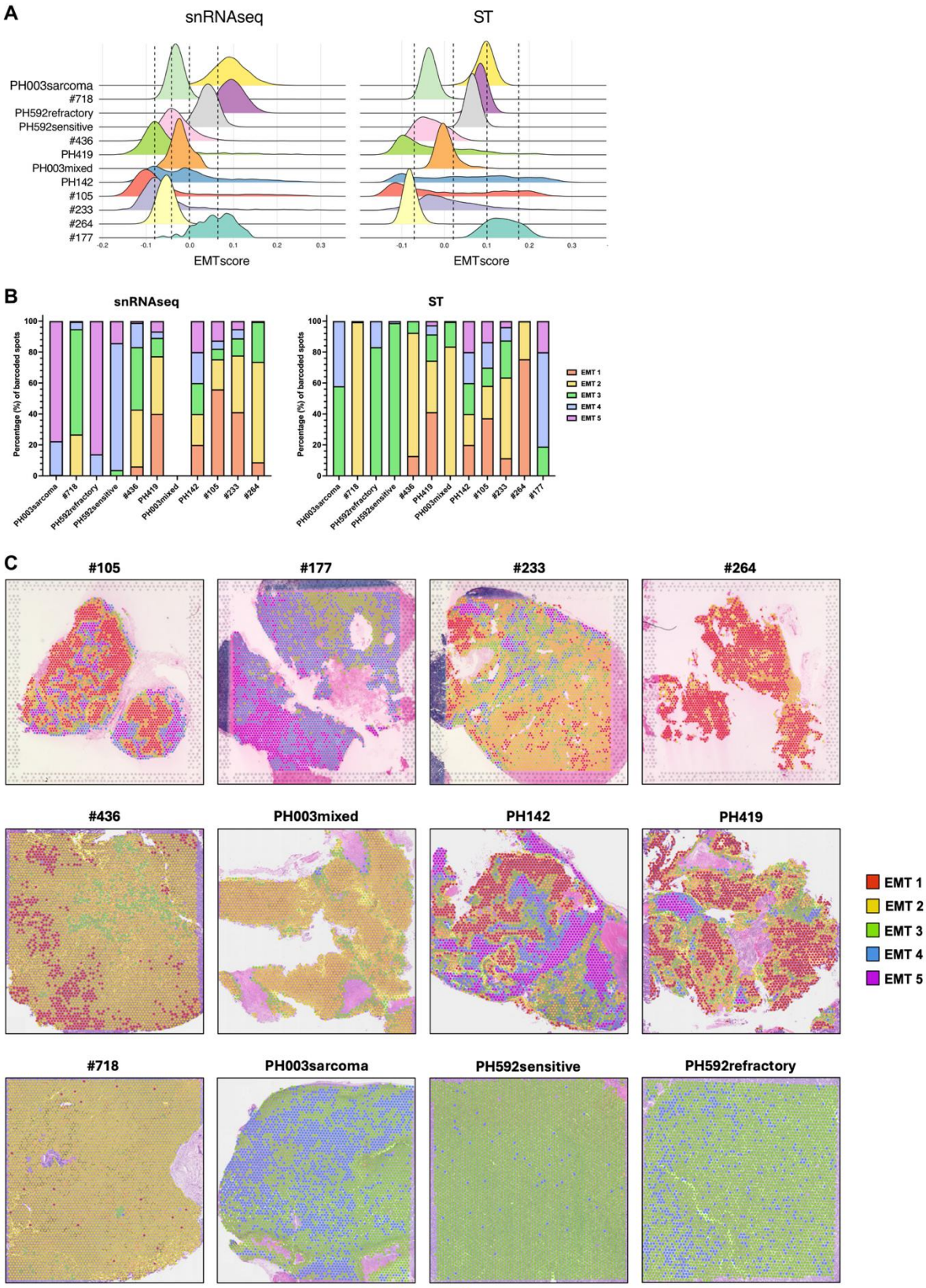

A
snRNAseq
ST
PH003sarcoma
#718
PH592refractory
PH592sensitive
#436
PH419
PH003mixed
PH142
#105
#233
#264
#177
EMTscore
B
snRNAseq
ST
Percentage (%) of barcoded spots
EMT 1
EMT 2
EMT 3
EMT 4
EMT 5
C
#105
#177
#233
#264
#436
PH003mixed
PH142
PH419
#718
PH003sarcoma
PH592sensitive
PH592refractory

**Figure 3: OCS-specific signatures segregated nuclei and barcoded spots into distinct EMT bins which exhibited significant spatial heterogeneity**

**(A)** Distribution of EMT scores calculated using AUCell based on expression of OCS-specific epithelial- and mesenchymal-like gene signatures across all nuclei (snRNA-seq) and barcoded spots (ST) from the OCS PDX tumours. **(B)** Proportion of nuclei and barcoded spots representing each of the five distinct EMT bins in the OCS PDX tumours. **(C)** Spatial distribution of barcoded spots representing distinct EMT bins mapped to the tissue section analysed via ST for each OCS PDX tumour using Loupe Browser (v8.1.1).

OCS, ovarian carcinosarcoma; EMT, epithelial-to-mesenchymal transition; snRNA-seq, single-nucleus

RNA sequencing; ST, spatial transcriptomics; PDX, patient-derived xenograft

## Figure 4

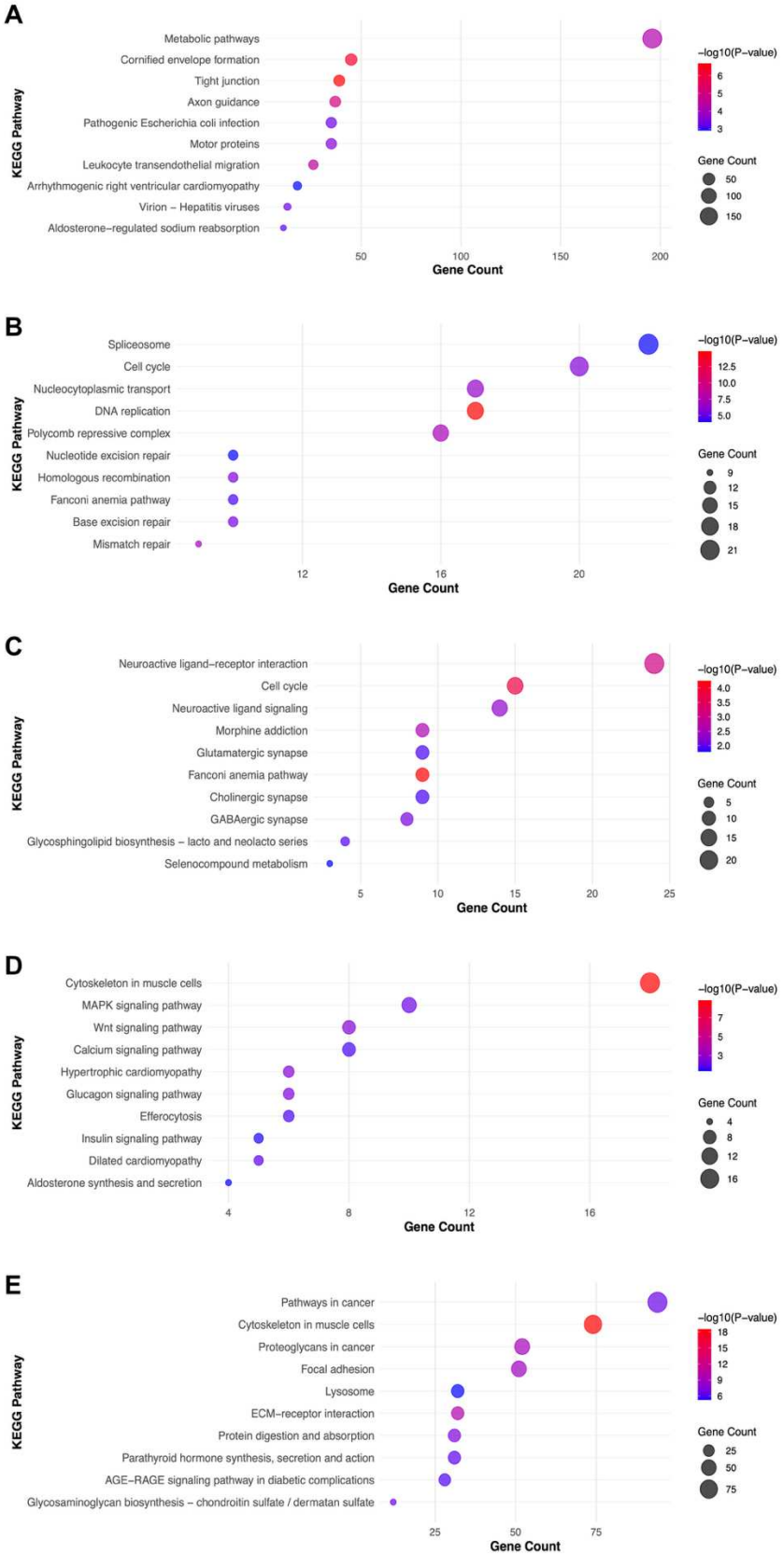
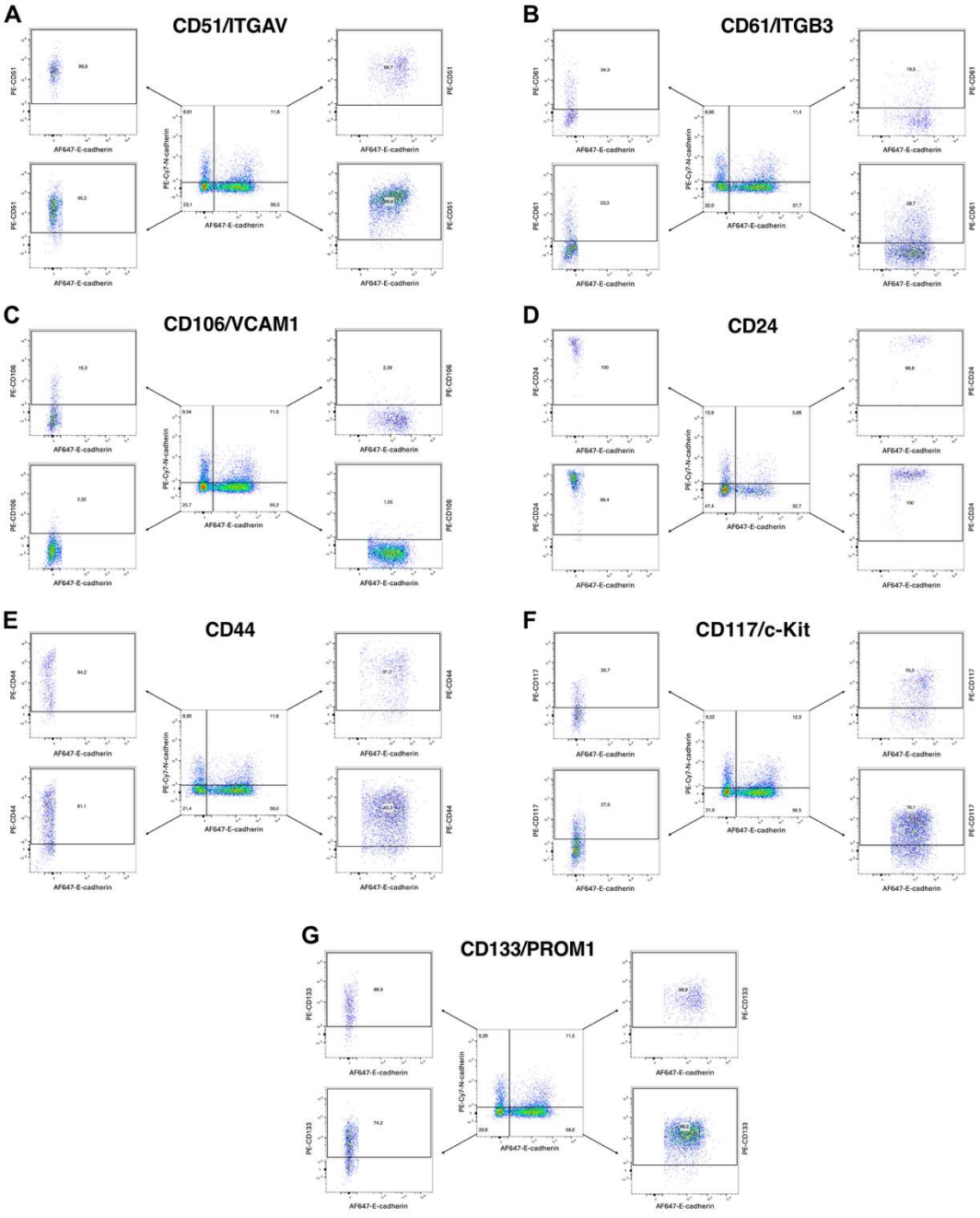

A
Metabolic pathways
Cornified envelope formation
Tight junction
Axon guidance
Pathogenic Escherichia coli infection
Motor proteins
Leukocyte transendothelial migration
Arrhythmogenic right ventricular cardiomyopathy
Virion – Hepatitis viruses
Aldosterone−regulated sodium reabsorption
KEGG Pathway
50
100
150
200
Gene Count
−log10(P−value)
6
5
4
3
Gene Count
50
100
150
B
Spliceosome
Cell cycle
Nucleocytoplasmic transport
DNA replication
Polycomb repressive complex
Nucleotide excision repair
Homologous recombination
Fanconi anemia pathway
Base excision repair
Mismatch repair
KEGG Pathway
12
16
20
Gene Count
−log10(P−value)
12.5
10.0
7.5
5.0
Gene Count
9
12
15
18
21
C
Neuroactive ligand−receptor interaction
Cell cycle
Neuroactive ligand signaling
Morphine addiction
Glutamatergic synapse
Fanconi anemia pathway
Cholinergic synapse
GABAergic synapse
Glycosphingolipid biosynthesis – lacto and neolacto series
Selenocompound metabolism
KEGG Pathway
5
10
15
20
25
Gene Count
−log10(P−value)
4.0
3.5
3.0
2.5
2.0
Gene Count
5
10
15
20
D
Cytoskeleton in muscle cells
MAPK signaling pathway
Wnt signaling pathway
Calcium signaling pathway
Hypertrophic cardiomyopathy
Glucagon signaling pathway
Efferocytosis
Insulin signaling pathway
Dilated cardiomyopathy
Aldosterone synthesis and secretion
KEGG Pathway
4
8
12
16
Gene Count
−log10(P−value)
7
5
3
Gene Count
4
8
12
16
E
Pathways in cancer
Cytoskeleton in muscle cells
Proteoglycans in cancer
Focal adhesion
Lysosome
ECM−receptor interaction
Protein digestion and absorption
Parathyroid hormone synthesis, secretion and action
AGE−RAGE signaling pathway in diabetic complications
Glycosaminoglycan biosynthesis – chondroitin sulfate / dermatan sulfate
KEGG Pathway
25
50
75
Gene Count
−log10(P−value)
18
15
12
9
6
Gene Count
25
50
75

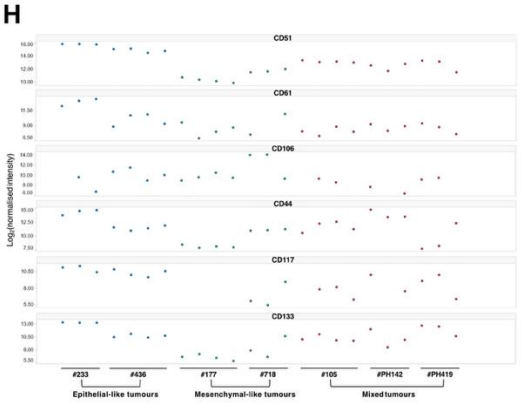
**Figure 4: KEGG pathways or networks significantly over-represented in pseudo-bulked nuclei representing distinct EMT states in OCS**

**(A-E)** Top KEGG pathways or networks significantly over-represented (adjusted P-value < 0.05) by **(A)** EMT 1**, (B)** EMT 2**, (C)** EMT 3**, (D)** EMT 4**, and (E)** EMT 5 marker genes in pseudo-bulked nuclei across the OCS PDX tumours.

KEGG, Kyoto Encyclopedia of Genes and Genomes; EMT, epithelial-to-mesenchymal transition; OCS, ovarian carcinosarcoma; PDX, patient-derived xenograft

## Figure 5


A
CD51/ITGAV
B
CD61/ITGB3
C
CD106/VCAM1
D
CD24
E
CD44
F
CD117/c-Kit
G
CD133/PROM1
PE-CD51
PE-CD61
PE-CD106
PE-CD24
PE-CD44
PE-CD117
PE-CD133
PE-Cy7-N-cadherin
AF647-E-cadherin


H
CD51
CD61
CD106
CD44
CD117
CD133
Log2(normalised intensity)
#233
#436
#177
#718
#105
#PH142
#PH419
Epithelial-like tumours
Mesenchymal-like tumours
Mixed tumours

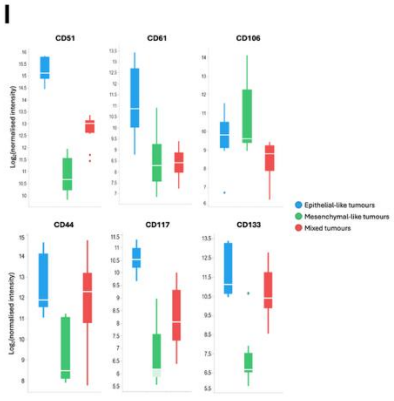

I
CD51
CD61
CD106
CD44
CD117
CD133
Log2(normalised intensity)
Epithelial-like tumours
Mesenchymal-like tumours
Mixed tumours

**Figure 5: Stem cell markers but not previously established murine p-EMT markers were relevant to intermediate EMT states in human OCS**

**(A-G)** Expression of **(A)** CD51/ITGAV, **(B)** CD61/ITGB3, **(C)** CD106/VCAM1, **(D)** CD24, **(E)** CD44, **(F)** CD117/c-Kit, and **(G)** CD133/PROM1 in populations of OCS PDX tumour cells segregated by E- and N-cadherin expression via flow cytometry. **(H-I)** Protein intensity of previously established murine hybrid E/M and human ovarian cancer stem cell markers across **(H)** individual OCS PDX tumours and **(I)** epithelial-like, mesenchymal-like, and mixed OCS PDX tumour groups assessed via cell surface proteomics.

p-EMT, partial epithelial-to-mesenchymal transition; OCS, ovarian carcinosarcoma; PDX, patient-derived xenograft

# Figure 6

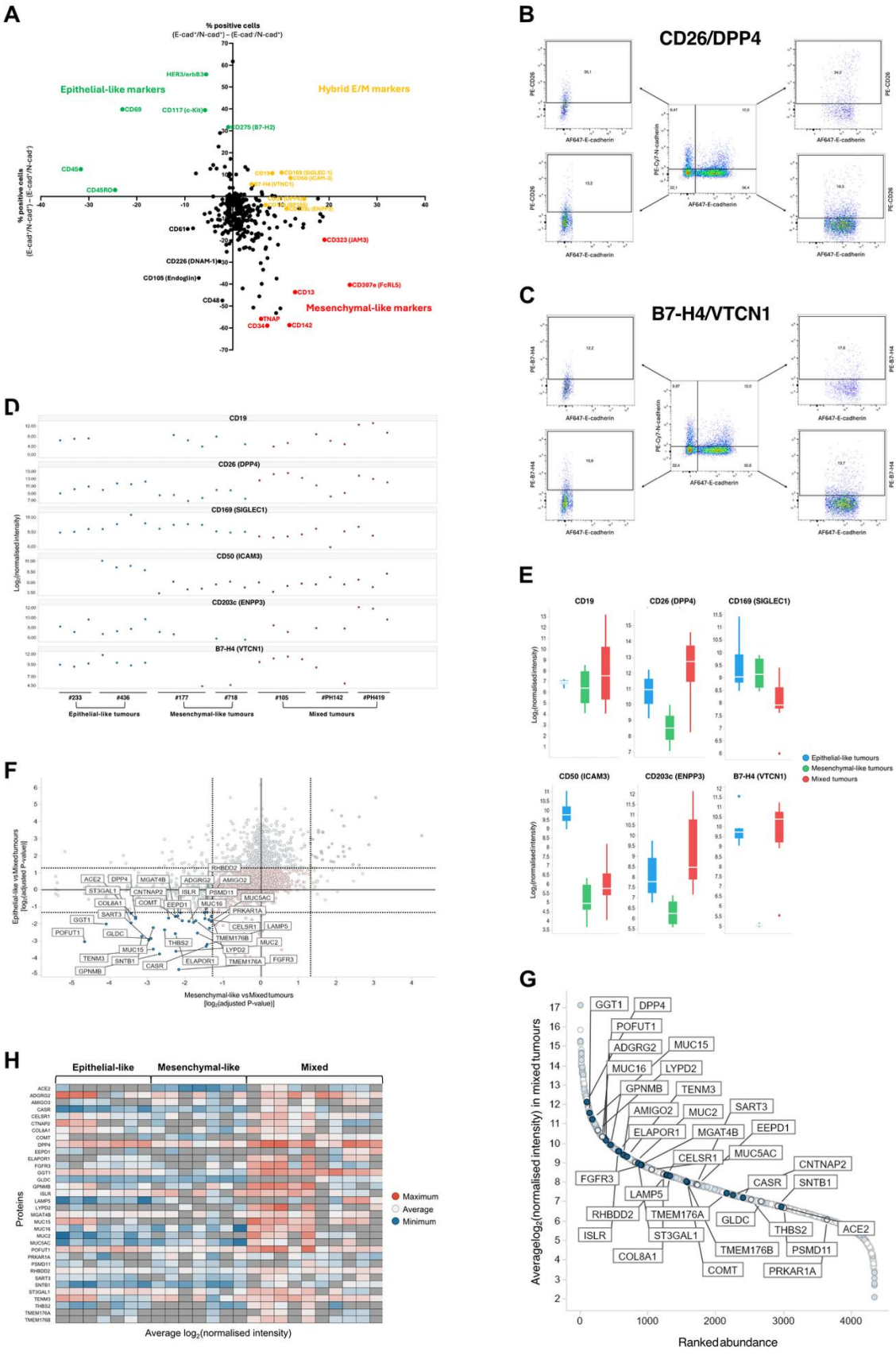

A
% positive cells
(E-cad+/N-cad+) – (E-cad-/N-cad+)
Epithelial-like markers
Hybrid E/M markers
Mesenchymal-like markers
HER3/erbB3
CD69
CD117 (c-Kit)
CD275 (B7-H2)
CD45
CD45RO
CD61
CD226 (DNAM-1)
CD105 (Endoglin)
CD48
CD323 (JAM3)
CD307e (FcRL5)
CD13
TNAP
CD34
CD142
% positive cells (E-cad+/N-cad+) – (E-cad-/N-cad+)
B
CD26/DPP4
PE-CD26
AF647-E-cadherin
PE-Cy7-N-cadherin
C
B7-H4/VTCN1
PE-B7-H4
D
CD19
CD26 (DPP4)
CD169 (SIGLEC1)
CD50 (ICAM3)
CD203c (ENPP3)
B7-H4 (VTCN1)
Log2(normalised intensity)
#233
#436
#177
#718
#105
#PH142
#PH419
Epithelial-like tumours
Mesenchymal-like tumours
Mixed tumours
E
CD19
CD26 (DPP4)
CD169 (SIGLEC1)
CD50 (ICAM3)
CD203c (ENPP3)
B7-H4 (VTCN1)
Log2(normalised intensity)
Epithelial-like tumours
Mesenchymal-like tumours
Mixed tumours
F
Epithelial-like vs Mixed tumours [log2(adjusted P-value)]
Mesenchymal-like vs Mixed tumours [log2(adjusted P-value)]
RHBDD2
ACE2
DPP4
MGAT4B
ADGRG2
AMIGO2
ST3GAL1
CNTNAP2
ISLR
PSMD11
COL8A1
COMT
EEPD1
MUC16
MUC5AC
PRKAR1A
SART3
GGT1
CELSR1
LAMP5
POFUT1
GLDC
TMEM176B
THBS2
MUC2
MUC15
LYPD2
TENM3
SNTB1
ELAPOR1
TMEM176A
FGFR3
CASR
GPNMB
G
Average log2(normalised intensity) in mixed tumours
Ranked abundance
GGT1
DPP4
POFUT1
ADGRG2
MUC15
MUC16
LYPD2
GPNMB
TENM3
AMIGO2
MUC2
SART3
ELAPOR1
MGAT4B
EEPD1
CELSR1
MUC5AC
CNTNAP2
FGFR3
CASR
SNTB1
LAMP5
RHBDD2
TMEM176A
GLDC
THBS2
ACE2
ISLR
ST3GAL1
COL8A1
TMEM176B
PSMD11
COMT
PRKAR1A
H
Epithelial-like
Mesenchymal-like
Mixed
Proteins
Maximum
Average
Minimum
Average log2(normalised intensity)

**Figure 6: Dipeptidyl peptidase 4 (DPP4) was significantly enriched in hybrid E/M cells and mixed OCS PDX tumours**

**(A)** Cell surface markers enriched in OCS PDX cells representing distinct EMT states using flow cytometry. **(B-C)** Expression of **(B)** CD26/DPP4 and **(C)** B7-H4/VTCN1 in populations of OCS PDX tumour cells segregated by E- and N-cadherin expression via flow cytometry. **(D-E)** Protein intensity of surface markers enriched in E-cadherin$^{+}$/N-cadherin$^{+}$ cells across **(D)** individual OCS PDX tumours and **(E)** epithelial-like, mesenchymal-like, and mixed OCS PDX tumour groups assessed via cell surface proteomics. **(F)** Proteins significantly enriched in mixed OCS PDX tumours, relative to epithelial- and mesenchymal-like models assessed via cell surface proteomics. **(G)** Ranked abundance plot of proteins significantly enriched in the mixed OCS PDX tumours. Blue dots represent cell surface proteins. **(H)** Relative expression of proteins significantly enriched in the mixed OCS PDX tumours across all individual tumours assessed via cell surface proteomics. Max = 15.98, Min = 0.37.

E/M, epithelial/mesenchymal; OCS, ovarian carcinosarcoma; PDX, patient-derived xenograft

## Declarations

### Ethics approval and consent to participate

Ethics approval for the WEHI-SFRCP was obtained from the Royal Melbourne Hospital Human Research Ethics Committee (Project number 2015.300) and governance review by the WEHI HREC (Project number G16/02). All animal experiments were conducted according to the National Health and Medical Research Council Australian Code for the Care and Use of Animals for Scientific Purposes 8th Edition, 2013 (updated 2021), and under the approval of the WEHI Animal Ethics Committee (AEC 2019.024, 2022.030, and 2022.038).

### Competing interests

C.L.S. declares Advisory Boards for AstraZeneca, Clovis Oncology, Roche, Eisai Inc, Sierra Oncology, Takeda, MSD and Grant/Research support from Eisai Inc, AstraZeneca, Boehringer Ingelheim and IDEAYA Biosciences. C. Vandenberg and H.E. Barker report research support (paid

to institution) from AstraZeneca Pty Ltd and Boehringer Ingelheim outside the submitted work. Other authors declare no conflicts of interest.

**Funding**

This work was supported by fellowships and grants from the National Health and Medical Research Council (NHMRC Australia; Investigator grant 2009783 (CLS)); the Stafford Fox Medical Research Foundation (AH, AF, CLS and HEB); the Victorian Cancer Agency (Mid-Career Research Fellowship MCRF23021 (HEB)), The Sue Tyerman Scholarship (AH); and the Commonwealth through an Australia Government Research Training Program Scholarship *[DOI: https://doi.org/10.82133/C42F-K220]* (AH).

**Author Contributions**

Conceptualisation, AH, XD, RB, YC, CLS and HEB; Methodology, AH, XD, AF, CJAA, RDM, LL, DAZ, JW, VV, JY, LFG, CJV, RB, YC, CLS and HEB; Investigation, AH, XD, RB, YC, CLS and HEB; Formal analysis, AH, XD, LL, JW, VV, JY, LFG, GR, YC, CLS and HEB; Project administration, DAZ, RB, YC, CLS and HEB; Resources, DAZ, LFG, RB, YC, CLS and HEB; Software, XD, JW, JY, LFG, RB, YC, CLS and HEB; Supervision, AF, DAZ, LFG, CJV, RB, YC, CLS and HEB; Visualisation, AH, ZD, JY, LFG, CJV and HEB; Funding acquisition, DAZ, RB, YC, CLS and HEB; Writing – original draft preparation, AH, XD and HEB; Writing – review & editing, AH, XD, VV, JY, LFG, CJV, YC, CLS and HEB. All authors read and approved the final manuscript.

**Acknowledgments**

We thank John Weroha for the original PH PDX models. We thank Rachel Hancock, Daniel Fayle, Stephanie Bound, Kathy Barber and Silvia Stoev and WEHI Bioservices for their technical assistance. We thank Alex Garnham for technical assistance with transcriptomics data. This work was made possible through the Australian Cancer Research Foundation, the Victorian State Government Operational Infrastructure Support and Australian Government NHMRC IRIISS. The WEHI- Stafford Fox Rare Cancer Program acknowledges support from the Australian and New Zealand Gynaecological Oncology Group and Australian Rare Cancer Portal for facilitating patient accrual. WEHI-Stafford Fox Rare Cancer Program would also like to acknowledge the

contributions of all clinical collaborators, including study nurses, at participating centres throughout Australia for their involvement in this study.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- HadlaetalSupplementaryTables.xlsx
- HadlaetalSuppFigs.docx
- rsflat.pdf