A multi-omic analysis of the heterogeneity of ovarian carcinoma

Judy Sobh

Thesis submitted to the University of Ottawa in partial fulfillment of the requirement for the Master's degree in Cellular and Molecular Medicine

Department of Cellular and Molecular Medicine

Faculty of Medicine

University of Ottawa

Abstract

As part of the OvCan Initiative by Ovarian Cancer Canada, we generated a resource of multi-omic data from 31 human ovarian cancer cell models spanning six histological subtypes (HGSC, CCC, MC, EC, SCCOHT, MMMT). Using RNA sequencing, TMT proteomics, and whole-exome sequencing, we systematically characterized molecular features across models to assess inter-subtype and intra-subtype variability. We demonstrate that these data capture known features of distinct ovarian cancer subtypes, including subtype-specific patterns of gene expression, pathway activation and genomic alterations. We also identify novel molecular patterns, including a partial convergence between EC and SCCOHT models possibly linked to shared SWI/SNF disruption, and concordant RNA-protein signatures highlighting candidate biomarkers and therapeutic targets. Furthermore, we highlight substantial diversity among models within individual subtypes, underscoring the importance of careful model selection in experimental design. Together, this resource will support the ovarian cancer research community by enabling subtype-specific biomarker discovery, therapeutic target identification, and informed model selection, ultimately contributing to the development of new diagnostic and therapeutic approaches.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Dr. David Cook, for his guidance, expertise and constructive feedback throughout this project. I am especially grateful for his patience and dedication in mentoring me in computational and bioinformatic methods. Entering this Master's program with limited prior computational experience was daunting, however his guidance was instrumental in building my technical skills and confidence, enabling me to independently design, implement and interpret the analyses presented in this thesis. His mentorship has provided me with a strong foundation to support my future professional and academic endeavours.

I would also like to thank my Thesis Advisory Committee members, Dr. Barbara Vanderhyden and Dr. Arvind Mer. Thank you for your insightful guidance and valuable feedback regarding my project. I truly enjoyed our TAC meetings and sharing my project with you.

Finally, I would like to thank my family for their constant encouragement and support throughout my academic journey. Their presence was a crucial source of motivation for me during the completion of this project.

I am sincerely grateful to everyone who contributed, directly or indirectly, to the completion of this thesis.

Table of Contents

List of Abbreviations

ADGRL2 – Adhesion G protein–coupled receptor L2

ANGPT1/2 – Angiopoietin-1/2

ARID1A – AT-rich interactive domain-containing protein 1A

ATM – Ataxia-telangiectasia mutated

ATR – ATM- and Rad3-related protein

B3GALT2 – Beta-1,3-galactosyltransferase 2

BRAF – B-Raf proto-oncogene serine/threonine kinase

BRCA1/2 – Breast cancer susceptibility genes 1 and 2

CA125 (MUC16) – Cancer antigen 125

CA19-9 – Carbohydrate antigen 19-9

C3 – Complement component 3

C3 convertase – Central complement activation enzyme complex

C4A/C4B – Complement component 4A/B

C4BPB – C4b-binding protein beta chain

C4BPA – C4b-binding protein alpha chain

CCC – Clear cell carcinoma

CCLE – Cancer Cell Line Encyclopedia

CDKN2A – Cyclin-dependent kinase inhibitor 2A

CEA – Carcinoembryonic antigen

CNV – Copy number variation

COL11A1 – Collagen type XI alpha 1 chain

COL1A2 – Collagen type I alpha 2 chain

COL3A1 – Collagen type III alpha 1 chain

COL5A2 – Collagen type V alpha 2 chain

COLGALT2 – Collagen beta-galactosyltransferase 2

CPQ – Carboxypeptidase Q

CTNNB1 – Catenin beta 1

CXCL10/CXCL11 – C-X-C motif chemokines 10/11

CXCR3 – C-X-C motif chemokine receptor 3

CYB5R3 – Cytochrome b5 reductase 3

DACT1 – Dishevelled binding antagonist of beta-catenin 1

DDR – DNA damage repair

DKK2 – Dickkopf WNT signaling pathway inhibitor 2

DLK1 – Delta like non-canonical Notch ligand 1

DNA – Deoxyribonucleic acid

DSB/DSBs – Double-strand break(s)

D11S860 – Chromosome 11 microsatellite marker

EC – Endometrioid carcinoma

ECM – Extracellular matrix

EGFR – Epidermal growth factor receptor

EGLN3 – Egl-9 family hypoxia inducible factor 3

ENO1 – Enolase 1

EOC – Epithelial ovarian cancer

E2F – E2F transcription factor family

ERBB2 (HER2) – Erb-B2 receptor tyrosine kinase 2

EVI1 – Ecotropic Viral Integration site 1

FDR – False discovery rate

FIGO – International Federation of Gynecology and Obstetrics

FOLR1 – Folate receptor alpha

FUT3 – Fucosyltransferase 3

GEMMs – Genetically engineered mouse models

GI – Gastrointestinal

GO – Gene Ontology

GGT1 – Gamma-glutamyltransferase 1

HDI – Human Development Index

HE4 (WFDC2) – Human epididymis protein 4

HGSC – High-grade serous carcinoma

HGF – Hepatocyte growth factor

HIFs – Hypoxia-inducible factors

HILPDA – Hypoxia inducible lipid droplet associated

HMGA2 – High mobility group AT-hook 2

HOX – Homeobox genes

HR – Homologous recombination

HRD – Homologous recombination deficiency

IL-6 – Interleukin 6

JAK-STAT – Janus kinase–signal transducer and activator of transcription

KRAS – Kirsten rat sarcoma viral oncogene

LAMA1 – Laminin subunit alpha 1

LAMC1 – Laminin subunit gamma 1

LGSC – Low-grade serous carcinoma

logFC – Log fold change

MAPK – Mitogen-activated protein kinase

MC – Mucinous carcinoma

MCUR1 – Mitochondrial calcium uniporter regulator 1

MCM2 – Minichromosome maintenance complex component 2

MCM complex – DNA replication licensing complex

MEK – Mitogen-activated protein kinase kinase

MEKi – MEK inhibitor

MET – MET proto-oncogene receptor tyrosine kinase

MID1 – Midline 1

miRNA – MicroRNA

MIOX – Myo-inositol oxygenase

MRPL2 – Mitochondrial ribosomal protein L2

MSLN – Mesothelin

mRNA – Messenger RNA

MUC1 – Mucin 1

MUC16 – Mucin 16

NFκB – Nuclear factor kappa B

NHEJ – Non-homologous end joining

NRAS – Neuroblastoma RAS viral oncogene

OvCAN – Ovarian Cancer Research Initiative

PARP – Poly(ADP-ribose) polymerase

PARP1 – Poly(ADP-ribose) polymerase 1

PCA – Principal component analysis

PCNA – Proliferating cell nuclear antigen

PDK3 – Pyruvate dehydrogenase kinase 3

PDS – Primary debulking surgery

PI3K – Phosphoinositide 3-kinase

PIK3CA – Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha

PI3K-AKT – PI3K–AKT signaling pathway

PI3K/mTOR – PI3K–mTOR signaling pathway

PROGENy – Pathway RespOnsive GENes analysis

PTPN1 – Protein tyrosine phosphatase non-receptor type 1

QC – Quality control

RAD1 – RAD1 checkpoint DNA exonuclease

RNA – Ribonucleic acid

ROS – Reactive oxygen species

RTK – Receptor tyrosine kinase

SLC34A2 – Sodium-dependent phosphate transporter 2B

SLPI – Secretory leukocyte protease inhibitor

SMARCA4 – SWI/SNF related matrix associated actin dependent regulator of chromatin A4

SOX11 – SRY-box transcription factor 11

ST8SIA4 – ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 4

STAT3 – Signal transducer and activator of transcription 3

STC2 – Stanniocalcin 2

STR – Short tandem repeat

SWI/SNF – SWItch/Sucrose Non-Fermentable chromatin remodeling complex

TACSTD2 (TROP-2) – Tumor associated calcium signal transducer 2

TCGA – The Cancer Genome Atlas

TF – Transferrin

THBS3 – Thrombospondin 3

TNFα – Tumor necrosis factor alpha

TNKS – Tankyrase

TP53 – Tumor protein p53

TREH – Trehalase

t-SNE – t-distributed stochastic neighbor embedding

WES – Whole-exome sequencing

WHO – World Health Organization

WNT – Wingless/integrated signaling pathway

WT1 – Wilms tumor 1

XRCC2 – X-ray repair cross-complementing protein 2

YAP1 – Yes-associated protein 1

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Ovarian Carcinoma

According to the World Health Organization (WHO), each year globally, an estimated total of 225,500 cases of ovarian cancer will be diagnosed, and 140,200 patients will succumb to the disease[1]. The global burden of this disease is set to rise, and despite higher incidence found in countries with higher human developmental index (HDI), mortality rates are similar across HDI levels[2]. Ovarian carcinoma is often mislabeled as a single disease entity, despite comprising various heterogeneous subtypes, each with distinct complex molecular landscapes[3]. Most ovarian cancers are epithelial in origin, accounting for about 95% of cases, while the remaining 5% are represented by non-epithelial ovarian cancers[4]. The WHO has identified five primary subtypes of epithelial ovarian cancers: high-grade serous carcinoma (HGSC), low-grade serous carcinoma (LGSC), mucinous carcinoma (MC), endometrioid carcinoma (EC) and clear cell carcinoma (CCC)[5]. Non-epithelial ovarian cancers include germ cell and sex-cord stromal cancers[4]. Epithelial ovarian carcinomas typically arise from epithelial tissues within the adnexa (specifically surface epithelium of the ovaries or the fallopian tube epithelium for serous ovarian cancers), whereas non-epithelial ovarian carcinomas arise from non-surface tissue such as primordial germ cells, sex cords or stromal cells[6].

Subtypes of epithelial ovarian cancers differ in morphology, biomarkers, mutations, responses to therapy and prognoses. Molecular landscapes also evolve throughout disease progression or therapy, amplifying their complexity. Understanding the distinct features of ovarian cancer subtypes is critical to develop personalized strategies for clinical management.

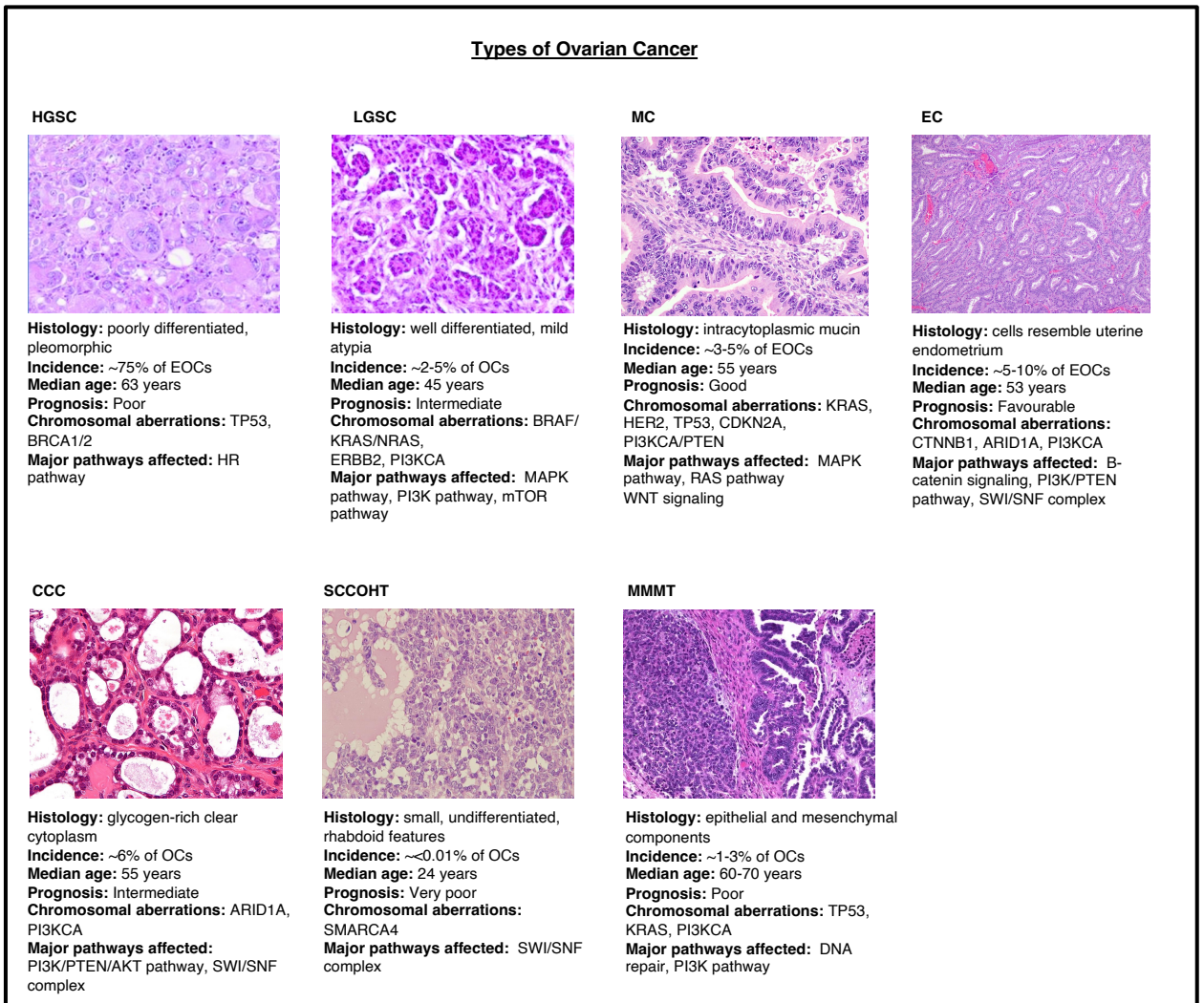

Figure 1. Overview of ovarian cancer subtypes. Overview includes histology, epidemiology, and key molecular alterations.

1.1.1 High grade serous ovarian carcinoma

HGSC, the most prevalent subtype accounting for approximately 75% of epithelial ovarian cancers, is highly aggressive and characterized by poorly differentiated, rapidly dividing cells that spread sporadically[7]. The median age of diagnosis for HGSC is 63 years[8]. HGSC is defined by nearly ubiquitous TP53 mutations (~96%)[9] and defects in homologous recombination repair (HRD; ~50%)[10,11], such as BRCA1/2 mutations (~25%)[12], leading to its genomic instability[13]. Although historically thought to arise solely from ovarian surface epithelium, accumulating molecular, pathological and genetic evidence now supports that HGSC may arise from multiple tissues of origin, including both the ovarian surface epithelium and the fallopian tube. Furthermore, it is now better supported that the distal fallopian tube secretory epithelium is a primary site of origin for many HGSCs. More specifically, serous tubal intraepithelial carcinoma (STIC) lesions are a possible precursor for HGSC. STIC lesions arise through the transformation of fallopian tube secretory epithelium, often initiated by early TP53 mutations, leading to localized epithelial proliferation and further progression to HGSC. However, not all HGSC cases exhibit identifiable STIC lesions, suggesting biological heterogeneity of this subtype.[14] In addition, HGSC transcriptomic signatures more closely resemble fallopian tube secretory epithelium than ovarian surface epithelium[15]. HGSC is frequently diagnosed at advanced stages (about 70% of cases at FIGO stages III or IV)[16], and treatment typically involves a combination of cytoreductive surgery and platinum-based chemotherapy[17]. Maintenance therapy using PARP inhibitors has significantly improved outcomes for BRCA1/2 and HRD-positive patients, however novel treatment options are still needed, as approximately 70% of patients develop resistance and the prognosis for this subtype remains poor[18,19].

1.1.2 Low-grade serous ovarian carcinoma

LGSC is the rarer serous subtype of ovarian cancer, accounting for approximately 2-5% of cases of all ovarian cancers, and 5-10% of serous cases. The median age of diagnosis for this subtype is around 45 years, which is younger than HGSC patients.[20] Histologically, LGSC is characterized by well-differentiated epithelial cells with low nuclear atypia, reflecting its slower proliferative rate in comparison to HGSC. Due to this, LGSC exhibits lower genomic instability and fewer copy number alterations across the chromosomes.[21] Molecularly, LGSC is distinct from HGSC as it often lacks mutations in TP53 (~8-20%)[22], and instead is typically driven by mutations in the mitogen-activated protein kinase (MAPK) pathway; specifically in KRAS (~16-44%), BRAF (2-20%), NRAS (up to 26%)[23,24]. Patients typically have a better prognosis than HGSC, yet often experience recurrent disease due to intrinsic chemoresistance. Nonetheless, this disease remains deadly. Despite being chemo-resistant, standard platinum-based chemotherapy remains the main treatment option for patients, as few effective alternative treatment options are currently available. Furthermore, most LGSC tumours express estrogen and progesterone receptors, keeping them hormonally driven. Therefore, maintenance endocrine therapy post-chemotherapy has been shown to improve patient outcomes, however, most patients ultimately experience disease progression and require additional systemic therapies.[20]

1.1.3 Clear cell carcinoma of the ovary

CCC is also a rare ovarian carcinoma subtype, accounting for approximately 6% of cases[25]. CCC is named after its histological appearance, where the cytoplasm appears clear due to a high intracytoplasmic accumulation of glycogen and lipids that are dissolved during standard tissue processing, leaving behind an empty or vacuolated cytoplasm[26,27]. Patients commonly have loss-

of-function mutations in ARID1A (~46-58%)[28], and activating mutations in PIK3CA (~32-50%)[29], indicating disruption of chromatin remodeling and PI3K/AKT signalling as an oncogenic driver. Endometriosis is a main risk factor and a recognized precursor lesion for CCC, with endometriosis identified in over 50% of patients[30]. Studies have shown that the key driver mutations found in CCC, specifically the loss-of-function mutations in ARID1A, can be detected in adjacent endometriosis and atypical endometriosis lesions, which supports the theory that endometriosis can progress to carcinoma[26,31]. The median age of diagnosis for CCC is approximately 55 years old[32]. This subtype exhibits intrinsic chemoresistance, resulting in a lower rate of response to chemotherapy (~11%)[33]. However, standard treatment regimens typically mirror HGSC - aggressive cytoreductive surgery and platinum-based chemotherapy, highlighting the need for improved subtype-specific treatment strategies for this disease.[30]

1.1.4 Mucinous carcinoma of the ovary

MC is a rare epithelial ovarian cancer subtype, representing about 3-5% of epithelial ovarian cancer cases, with a median age of diagnosis of approximately 55 years. MC is biologically and clinically distinct from other epithelial ovarian cancer subtypes, particularly HGSC.[34,35] The common risk factors of HGSC, such as nulliparity, late menopause, early menarche, BRCA mutations and lack of breastfeeding are not associated with MC[36]. Furthermore, it has a distinct molecular landscape, which includes frequent activating MAPK mutations such as KRAS (~40-50%) and ERBB2 amplifications (~19%)[37]. Other mutations commonly found include CDKN2A (~33%), and TP53 (~64%)[38,39]. This subtype is also named after its histological appearance, as tumour cells contain a large amount of intracellular mucin (~50%) in at least 90% of cells[40]. Serum markers include elevated carcinoembryonic antigen (CEA) (~80%)[40], CA19-9 (~47%)[41] and CA125 (~42%)[41], which can support diagnostic evaluation for this subtype. Previously, MC

cases were misdiagnosed as metastases from gastrointestinal (GI) sites such as the stomach, colon or pancreas[42]. This is because certain primary MC cases, specifically those which arise from teratomas, exhibit morphological and immunohistochemical characteristics that resemble intestinal and upper GI tract tumours[43]. These characteristics include expression of GI markers such as CDX2 and CK20, closely resembling colorectal adenocarcinoma, and KRAS-driven intestinal differentiation[44,45]. MC is frequently diagnosed at an early stage, with approximately 80% of cases being diagnosed as FIGO stage I, where prognosis is generally favourable. However, patient outcomes decline in advanced disease due to limited responsiveness to platinum-based chemotherapy.[40]

1.1.5 Endometrioid carcinoma of the ovary

EC is an endometriosis-associated epithelial ovarian cancer subtype and represents approximately 5-10% of EOCs, with patients being diagnosed at a median age of approximately 53 years[46]. EC is characterized by malignant glands that resemble endometrial epithelium, often explaining its classification as an endometrioid carcinoma. EC also displays well-formed squamous differentiation, consistent with a more differentiated phenotype than HGSC.[47,48,49] Molecularly, EC frequently harbours CTNNB1 (~40-43%), ARID1A (~40-43%) and PIK3CA (~36%) mutations, implicating dysregulation of WNT/β-catenin signalling, chromatin remodelling, and PI3K/AKT pathways in tumour development[46,50]. An estimated 15-50% of EC cases are linked to endometriosis[51], and it is thought that the oxidative stress associated with iron overload in endometriosis is responsible for the DNA mutations that could lead to EC[49]. Therefore, endometriosis is a non-obligatory precursor lesion[47].[52,53] EC is frequently detected at an earlier stage, contributing to a favourable prognosis (~95% in early stages)[54]. As in HGSC, standard management of EC consists of cytoreductive surgery and platinum-based

chemotherapy; however, activation of the PI3K pathway, including the PIK3CA amplification observed in this subtype, has been associated with reduced sensitivity to chemotherapy.[55,56]

1.1.6 Small-cell carcinoma of the ovary - hypercalcemic type

Small-cell carcinoma of the ovary - hypercalcemic type (SCCOHT) is a rare and highly aggressive ovarian carcinoma that accounts for less than 0.01% of ovarian neoplasms. SCCOHT typically affects young children, adolescents and young adults (median age ~ 24 years)[57,58]. The designation "hypercalcemic type" arises from the formation of elevated serum calcium levels due to abnormal bone formation and resorption processes from the cancer, which is observed in ~60% of patients[59]. This occurs through tumour secretion of parathyroid hormone-related protein (PTHrP), which mimics parathyroid hormone activity and disrupts normal calcium homeostasis.[60–62] SCCOHT is defined by biallelic inactivation of SMARCA4 (~95% of cases), a core catalytic subunit of the SWI/SNF chromatin remodeling complex. Loss of SMARCA4 results in epigenetic dysregulation and the loss of cellular differentiation.[59,63,64] Clinically, SCCOHT is lethal, and 5-year survival rates are around 10% [58]. Current management typically involves intensive surgical resection and chemotherapy/radiation; however, outcomes remain poor. These features collectively make SCCOHT one of the most aggressive ovarian malignancies.

1.1.7 Malignant mixed Müllerian tumour

Malignant mixed Mullerian ovarian carcinoma (MMMT) is an extremely rare ovarian neoplasm characterized by the presence of malignant epithelial (carcinoma) and stromal components (sarcoma). Due to this, this subtype is often referred to as carcinosarcoma. Although historically considered a biphasic tumour, accumulating studies indicate that both the carcinomatous and

sarcomatous components could arise from a single malignant epithelial precursor, with the sarcomatous elements developing through differential orientation of the carcinoma cells.[65,66] The median age of diagnosis is between 60-70 years, and MMMT most often develops in the ovaries, uterus and fallopian tubes[67,68]. Common mutations occur in TP53 (>50%), PIK3CA (~40%) and KRAS (~15%)[67,69]. MMMT represents only 1-3% of ovarian malignancies, and patients have an extremely low survival rate (~less than 2 years), even with chemotherapy, surgery and adjuvant protocols[70]. Overall, the aggressive nature and limited treatment success emphasizes the critical need for continued investigation into its biology.

1.2 Treatment Challenges

Rarer subtypes such as LGSC, MC, EC, CCC, SCCOHT and MMMT remain underrepresented in literature and clinical research compared to HGSC. This underrepresentation is typically due to low incidence and limited availability in clinical datasets, resulting in a disproportionate focus on HGSC. As a result of this, the current treatment options for patients with ovarian cancer are largely shaped by the biology and response of HGSC, and these options, primary debulking surgery (PDS) and platinum-based chemotherapy, tend to be far less effective in other subtypes of ovarian carcinoma. The initial chemotherapy response rate for HGSC is about 70%[71], whereas those of rarer subtypes are significantly lower (LGSC ; 4.9-23%[72,73], CCC ; 10-25%[74], MC ; 12.5-38.5%[75,76], EC/MMMT ; no clear estimate reported). Blanc-Durand *et al* treated SCCOHT patients with cytoreductive surgery and multiagent chemotherapy and achieved an 89% response rate[77]. Despite these encouraging responses, this disease remains highly aggressive, with most patients experiencing early relapse and poor long-term survival[58].

These differences in response rates highlight how the current HGSC-focused treatment models fail to address the biology of rarer subtypes. Furthermore, this mismatch between therapy and understanding of disease biology for each subtype has contributed to a gap in the understanding of less common subtypes and their therapeutic vulnerabilities.

1.3 Novel Targeted Therapies

Current investigations for targeted treatments are limited, but several approaches are emerging based on recurrent molecular alterations and biological characteristics unique to each subtype. These strategies aim to improve patient outcomes beyond conventional chemotherapies by using subtype-specific oncogenic drivers and targets.

Mitogen-activated protein kinase Extracellular signal-regulated kinase (MEK) inhibitors have been identified as targeted anti-cancer drugs that block MEK1/2, kinases involved in the MAPK/ERK signalling pathway. LGSC tumours are highly dependent on MAPK signalling due to frequent activating mutations in KRAS, BRAF and NRAS, making MEK a critical target for inhibiting downstream tumour growth and survival.[78] The MEKi trametinib demonstrated improved clinical outcomes relative to other treatment options (chemotherapy or hormonal therapy) in phase 3 trials[79]. However, objective response rates remain modest (~12-26%), and progression-free survival has been inconsistent across trials, indicating that only a subset of patients are fully benefiting from this therapy.[78] This further highlights the need for predictive biomarkers beyond MAPK to identify MEKi responsive tumours[80].

ARID1A loss-of-function mutations are present in CCC and EC, therefore targeting this vulnerability with anti-cancer drugs, such as EZH2 and ATR inhibitors are currently being explored[81,82]. Loss of ARID1A shifts chromatin regulation towards EZH2-mediated gene

repression, making ARID1A-mutant tumours dependent on EZH2, and vulnerable to EZH2 inhibition. In parallel, ARID1A normally plays a role in repairing DNA damage, therefore its loss causes cells to depend on ATR survival, further making them sensitive to ATR inhibition.[81,82,83,84]

Furthermore, MC and gastrointestinal tumours share similar pathological and molecular features, where MC tumour cells resemble intestinal goblet cells or gastric foveolar epithelium[45]. Consistent with this biology, studies have found that MC patients benefited from GI chemotherapy regimens and had improved survival.[85,86] The Gynecologic Oncology Group conducted the only randomized trial as of yet trying to compare capecitabine and oxaliplatin (a gastrointestinal regimen) with standard carboplatin and paclitaxel for MC, but the trial was unfortunately terminated because of poor accrual.[34,87]

Overall, due to the high degree of inter-patient, intra-patient and intra-tumour heterogeneity, patients with rarer subtypes of ovarian carcinoma suffer from a lack of tailored therapies and have fewer evidence-based treatment options despite being biologically distinct. This highlights the urgent need for more research on all subtypes of ovarian cancer in order to develop novel and efficient targeted therapies.

1.4 The Importance of Preclinical Models

Preclinical models are experimental systems derived from human or animal tumours used to study cancer biology and evaluate therapies prior to clinical testing[88]. These can include *in vitro* models (2D cell lines, 3D spheroids/organoids), *ex vivo* models (fresh tumour explants/slices) and *in vivo* models (tumors generated by genetic engineering or transplantation). These models allow researchers to study tumour-intrinsic pathways in controlled environments and mechanisms of therapeutic response. Models are generated either by deriving cultures from tumours (human or murine samples into cell lines and organoids) or creating tumours in mice via genetic engineering or transplantation (ex. patient-derived xenografts).[89–91]

Genetically engineered mouse models (GEMMs) are created by introducing defined oncogenic alterations into the mouse genome to push tumour development in a tissue of interest. These models can be used to study tumour initiation and progression, validate cancer genes and drug targets, assess therapy efficacy and evaluate mechanisms of drug resistance. However, it is important to note that murine models differ in their mutational burden and immune context, therefore limiting their ability to fully recapitulate human disease biology.[92,93]

Syngeneic mouse models are generated by implanting mouse tumour cell lines into immunocompetent mice of the same genetic background, allowing tumours to grow in the presence of a functioning immune system. These models are ideal for studying tumour immunology and possible immunotherapy treatments, and they can be particularly valuable for checkpoint blockade studies and evaluating host-tumour interactions.[94] We have previously systematically characterized syngeneic mouse models of HGSC and identified phenotypic variations among models that influence tumour characteristics and progression. For example,

among the various findings, this work identified inherent differences in the expression of immunoregulatory factors between different cell models. These differences contributed to distinct immune landscapes in the tumours they produced, suggesting that different models may be variably appropriate for testing immunotherapies. Despite these models being effective for understanding the mechanisms of tumor development, they don't faithfully recapitulate human disease. The murine genome and the tumour microenvironment differ substantially from those of humans, and there are fundamental differences in tumour evolution, genetic diversity, and immune response.[95,96]

Human-derived models are often more reliable in mimicking the complexity and heterogeneity of malignant cells, providing insight into disease biology. They can be derived from human tumours and used for in vitro testing or as xenografts in immunocompromised mice. However, since these mice lack components of their immune system, it limits the ability to study the immune response in the presence of the cancer. Despite this limitation, human models are valuable for therapeutic testing and are needed for offering a more personalized approach to ovarian cancer treatment. Human models also include established cell lines and patient-derived organoids, which serve as efficient, reproducible platforms for mechanistic studies and drug response studies while maintaining molecular and phenotypic features of the original tumours.

1.5 The Challenges of Current Preclinical Models

Despite significant advancements and diverse applications of both mouse and human models in ovarian cancer research, several critical challenges remain. Models are often generated by many independent labs and are supported by varying degrees of validation and characterization. Also, models are rarely collected and compared systematically or in a controlled way. In particular,

they are subject to variability in passage numbers, media composition, culture conditions, and different degrees of molecular validation.

Misidentification and cross-contamination of cell lines also remains a major issue in cancer research[97]. It is estimated that 15-20% of commonly used human cancer cell lines are mislabeled and not properly authenticated using techniques such as Short Tandem Repeat (STR) DNA genotyping[98]. It is crucial to understand the properties and limitations of the models and how they compare to one another. Together, these limitations underscore the urgent need for systematically characterized preclinical models to advance biological understanding and enable the discovery of effective diagnostic and therapeutic strategies in ovarian cancer.

1.6 The Importance of Multi-Omic Characterization

Multi-omic analyses integrate various layers of biological information, including transcriptomic profiles, proteomics, genomic sequencing, epigenomics and metabolomics to provide a holistic view of biological samples[99].

RNA-sequencing measures mRNA abundance by sequencing cDNA libraries derived from cellular RNA, providing a quantitative readout of gene expression across the entire transcriptome. RNA-seq provides the transcriptomic information that can identify gene expression profiles, differentially expressed genes, biomarker expression/abundance and relevant biological pathways.[100]

Proteomics uses mass spectrometry to quantify protein abundance. In this process, proteins are enzymatically digested into peptides, separated by liquid chromatography and analyzed by tandem mass spectrometry, allowing peptide spectra to be matched to protein sequences for quantitative analyses. Proteomics specifically measures protein abundance and post-translational

modifications and provides insight into signalling pathway activity and metabolic processes relevant to cancer cell models.[101,102]

Whole-exome sequencing (WES) profiles the DNA sequence of all protein-coding regions of the genome (~1-2% of the genome) by selectively capturing and sequencing exonic DNA[103]. In WES, genomic DNA is fragmented, and the protein-coding regions are selectively captured using hybridization probes prior to high-throughput sequencing. Unlike whole-genome sequencing, which analyzes the entire genome, WES focuses on coding regions, where the majority of known cancer-driving mutations occur, allowing the detection of somatic mutations, small insertions and deletions, and copy-number alterations. Therefore, WES characterizes the mutational landscape and structural variations driving tumour behaviour.[104–106]

Each of these 3 molecular modalities capture a distinct aspect of tumour biology, and their integration provides a more comprehensive and accurate representation of biology and malignancy.

Prior multi-omic efforts, such as the Broad Institute's Cancer Cell Line Encyclopedia (CCLE) and Pan-Cancer initiatives, have contributed genomic and transcriptomic datasets across many tumour types and human cancer cell lines. These datasets include somatic mutation profiles, copy-number alterations, RNA-seq-based gene expression data, DNA methylation and miRNA profiles, structural variation data and associated clinical and technical metadata across diverse tumour types. These initiatives enable systematic tumour classification, identification of oncogenic drivers, and provide insights for clinical decision-making.[107] However, ovarian cancer remains comparatively underrepresented in these resources.[108] In addition, the majority of the characterized models are HGSC, while the remaining rarer subtypes are represented by fewer

validated models. This highlights the need for comprehensive characterization across ovarian cell models to improve therapeutic discovery and model selection in preclinical research.

1.7 Rationale and Objectives

Given the limitations of existing multi-omic resources and the underrepresentation of rarer ovarian cancer subtypes, I sought to systemically characterize human ovarian carcinoma cell models as part of the OVCAN Initiative. Our objective is to use multi-omic profiling to evaluate and characterize a collection of 31 human cell models of diverse ovarian cancer subtypes from 3 sources across Canada. To address this, I propose the following 3 aims.

1. Collect and preprocess multi-omic data comprising RNA-sequencing, whole-exome sequencing, and proteomics from human cell models of ovarian cancer.
2. Identify molecular and phenotypic features that distinguish ovarian cancer subtypes.
3. Evaluate the similarity of models of each subtype to determine the extent of biological and technical variation.

Hypothesis

Integrating these three molecular modalities will provide a comprehensive understanding of these models and enable us to further understand complex molecular relationships that define different ovarian cancer subtypes. I hypothesized that models of the same subtype would be more similar to one another than to models of different subtypes.

Chapter 2: Materials and Methods

2.1 RNA collection and library preparation

Cell lines were derived and cultured as previously described[109–113]. Total RNA was extracted according to the manufacturer's instructions with the RNeasy Plus Mini Kit (Qiagen). RNA-seq libraries were generated (Genome Quebec) from 250ng of total RNA as following: mRNA enrichment was performed using the NEBNext Poly(A) Magnetic Isolation Module (New England BioLabs). cDNA synthesis was achieved with the NEBNext RNA First Strand Synthesis and NEBNext Ultra Directional RNA Second Strand Synthesis Modules (New England BioLabs). The remaining steps of library preparation were done using and the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England BioLabs). Adapters and PCR primers were purchased from New England BioLabs.

The libraries were normalized and pooled and then denatured in 0.02N NaOH and neutralized using HT1 buffer. The pool was loaded at 200pM on a Illumina NovaSeq S4 lane using Xp protocol as per the manufacturer's recommendations. The run was performed for 2x100 cycles (paired-end mode). A phiX library was used as a control and mixed with libraries at 1% level. Base calling was performed with RTA v3. Program bcl2fastq2 v2.20 was then used to demultiplex samples and generate fastq reads.

2.2 RNA-seq transcript quantification and processing

Kallisto was used to pseudoalign fastq files for each sample to the GRCh38 build of the human transcriptome. The R package tximport (v1.36.1) was used to load transcript quantifications, converting to gene-level transcript estimates. Principal component analysis (PCA) was performed on log-transformed transcript-per-million (TPM) expression values obtained via tximport from Kallisto quantification. Genes with zero expression across all samples were

removed, and for each subtype-specific PCA, only samples belonging to the corresponding subtype were included. Pairwise Euclidean distances were calculated using the first 10 principal components, which accounted for approximately 80% of the total variance.

2.3 Differential gene expression

Differential gene expression analysis was performed using the DESeq2 package (v1.48.2) where subtype-specific differential expression was evaluated by comparing samples of that subtype to all others grouped as "Other". A Wald test was used to calculate *p*-values. Signatures for each subtype were created by identifying differentially expressed genes with an adjusted *p*-value < 0.05 and log2FoldChange > 1 from the DESeq2 output. For each subtype, the resulting genes were sorted by significance (padj) and the top 100 genes were retained as the molecular signature. This cutoff was chosen to balance biological interpretability with sufficient pathway representation and is consistent with prior transcriptomic signature analyses.

2.4 GO Term over-representation analysis

Gene ontology (GO) over-representation analysis was conducted using the topGOtable() function from the pcaExplorer R package (v3.2.0), serving as a wrapper for the topGO package (v2.60.1), and uses org.Hs.eg.db gene annotation. The analysis was conducted for each subtype, using the corresponding molecular signatures for each. The elim algorithm was applied to reduce redundancy among enriched terms.

2.5 Gene set scoring and inference of signalling

Gene set activity scores for individual samples were computed using the R package singscore (v1.28.1). TPM-normalized gene expression values were log-transformed, and scores were

standardized across samples using Z-score formation. Scores reflect a rank-based statistic from genes comprising each set analogous to the Wilcoxon rank sum test. Signalling activity for cell lines were calculated using the R package PROGENy (v1.30.0) based on pretrained regression models of gene activity linked with 14 varying signalling pathways. The top 500 genes of each model were used to calculate the scores. The relative activity was visualized by reshaping scaled scores into matrices and plotted using the ComplexHeatmap package (v2.24.1), with clustering applied to columns using Ward's method.

2.6 Proteomics sample preparation and TMT labeling

Quantitative proteomics was performed by the Proteomics Platform at Canada's Michael Smith Genome Sciences Centre (BC Cancer, Vancouver, BC). Extracted protein lysates were processed using the single-pot, solid-phase-enhanced sample preparation (SP3) protocol on paramagnetic beads. Briefly, proteins were reduced, alkylated, and bound to SP3 beads, followed by on-bead enzymatic digestion with trypsin. Resulting peptides were labeled with isobaric tandem tag (TMT) reagents (Thermo Fisher Scientific) and combined at equal ratios. The multiplexed samples were fractionated by offline high-pH reversed-phase chromatography to reduce sample complexity and increase proteome coverage. Fractions were analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) on a Thermo Orbitrap mass spectrometer using synchronous precursor selection (SPS)-MS3 acquisition to minimize ratio compression from co-isolated precursor interference[114]. Raw files were searched against the UniProt human reference proteome, and protein quantification was derived from TMT reporter ion intensities.

2.7 Whole-exome sequencing

Genomic DNA from cell lines was extracted and exome sequencing libraries were prepared using the Nimblegen SeqCap EZ Exome 3.0 capture kit. Libraries were sequenced to an approximate depth of 50M reads per sample with 100bp paired-ends reads on a HiSeq 2500. Sequencing data was processed using the Sarek NextFlow pipeline bundling read process best practices and the CNVkit, Mutect2, and VEP for the identification and annotation of structural rearrangements and somatic variants. Reference normal tissue samples from Lai et al. (GEO Accession: GSE85671) were provided as a reference for CNV inference.[115]

Chapter 3: Collecting and preprocessing multi-omic data

3.1 Overview of the data set

To characterize human cell models of ovarian cancer, we generated a resource of multi-omic data from the OvCAN Collection of gold standard ovarian cancer cell models assembled by Ovarian Cancer Canada **(Table 1)**. I systematically generated RNA-seq, tandem-mass-tag (TMT) multiplexed proteomics, and whole-exome sequencing (WES) data from 31 human cell models. These models comprised diverse subtypes, including high-grade serous (HGSC; n=15), clear cell (n=7), mucinous (n=3), endometrioid (n=2), small cell carcinoma of the ovary, hypercalcemic type (SCCOHT; n=2), and malignant mixed-Mullerian tumour (MMMT; n=2). Patient-derived cell lines were provided by the Ottawa Hospital Research Institute (OHRI), British Columbia Cancer Research Institute (BCCRC) and the Montreal University Hospital Research Centre (CRCHUM). Subtypes were determined based on the clinical diagnosis of the patient. All data was processed using standard analysis pipelines **(see Methods)**.

Table 1. Summary table of cell lines, their histological subtype, and data availability across RNA-seq, proteomics and whole-exome sequencing. Subtype representation includes high-grade serous carcinoma (n = 15), clear cell carcinoma (n = 7), mucinous carcinoma (n = 3), endometrioid carcinoma (n = 2), small-cell carcinoma of the ovary, hypercalcemic type (SCCOHT; n = 2), and malignant mixed-Müllerian tumour (MMMT; n = 2).

Cell Line	Subtype	Stage	Chemo Status	RNA-seq	Proteomics	WES
BIN67	SCCOHT	N/A	N/A	✓	✓	x
COV434	SCCOHT	N/A	N/A	✓	✓	x
OV1369-R2	High-grade serous	IIIC	post	✓	✓	✓
OV1946	High-grade serous	IIIC	pre	✓	✓	x
OV2085	High-grade serous	IIIC	post	✓	✓	✓
OV2295	High-grade serous	IIIC	pre	✓	✓	✓
OV2295-R2	High-grade serous	IIIC	post	✓	✓	✓
OV3133-R	High-grade serous	IIIC	post	✓	x	✓
OV3331	High-grade serous	IIIC	post	✓	✓	✓
OV4453	High-grade serous	IIIC	pre	✓	✓	x
OV4485	High-grade serous	IIIC	post	✓	✓	x
OV866-2	High-grade serous	IIIC	post	✓	✓	x
OV90	High-grade serous	IIIC	pre	✓	✓	✓
TOV1369	High-grade serous	IIIC	pre	✓	✓	✓
TOV3041G	High-grade serous	IVA	post	✓	✓	x
TOV3133G	High-grade serous	IIIC	pre	✓	✓	✓
TOV3291G	High-grade serous	IIIC	pre	✓	x	x
TOV21G	Clear cell	III	pre	✓	✓	✓
TOV3392D	Clear cell	IIIC	pre	✓	x	✓
VOA10816	Clear cell	IA	N/A	✓	✓	x
VOA12539	Clear cell	IC1	N/A	✓	✓	x
VOA295	Clear cell	IIIC	pre	✓	✓	x
VOA4841	Clear cell	IC	pre	✓	✓	x
VOA6861	Clear cell	IC	N/A	✓	x	x
TOV112D	Endometrioid	IIIC	pre	✓	✓	✓
VOA4395	Endometrioid	IIIB	N/A	✓	✓	x
VOA8762	Mucinous	IV	pre	✓	✓	x
VOA8771	Mucinous	IV	pre	✓	✓	x
TOV2414	Mucinous	IIIC	post	✓	✓	✓
VOA5217	MMMT	N/A	post	✓	✓	x
VOA5436	MMMT	N/A	post	✓	x	x

3.2 Exclusion of LGSC due to batch effect

Prior to beginning my planned analyses and addressing my aims, I performed exploratory analyses to assess sample behaviour and to ensure reliability of downstream analyses. I began with dimensionality reduction analyses such as Principal Component Analysis (PCA), and a t-distributed Stochastic Neighbour Embedding (t-SNE) analysis to visualize the global structure of our models based on their gene expression profiles. This allowed me to assess whether samples exhibit subtype-specific clustering. From this analysis, I observed that all the LGSC models formed a tight cluster completely separating from the rest of the samples and subtypes **(Fig. 2a/b)**. While subtype-specific grouping can reflect real biology, the degree of separation was unusually high. I then evaluated the relative activity of relevant cancer biology signalling pathways using gene set scoring with PROGENy. Normally, different samples show distinct patterns of pathway activity that reflect their oncogenic drivers. However, almost all LGSC samples showed strong and consistent upregulation across all signaling pathways **(Fig. 2c)**. This behaviour does not align with known patterns of LGSC and is unlikely to represent real biology. For example, LGSC tumours are predominantly characterized by elevated MAPK signalling, but reduced inflammatory signalling compared to HGSC. The LGSC samples were also the only samples whose RNA-seq data was not generated together with the rest of the cohort. This pattern is consistent with a batch effect arising due to this independent processing. Since I had no reliable way to separate technical and biological, or to correct for the batch effect, the inclusion of LGSC would have made it difficult to distinguish genuine subtype-specific biology. Therefore, the LGSC samples were removed due to batch effects and were no longer included in the downstream analyses.

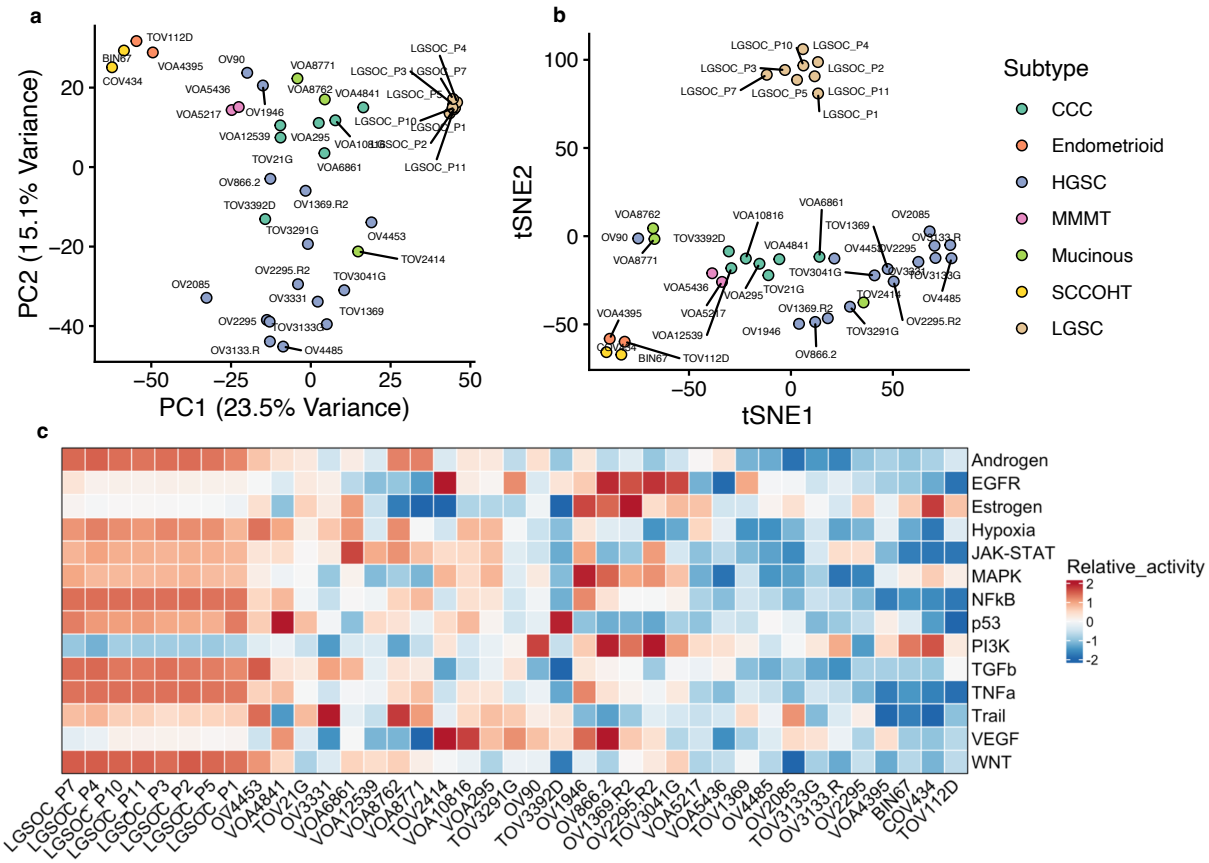

Figure 2. Initial exploratory analyses of LGSC highlighting its distinct molecular profile. **a)** PCA based on transcriptional molecular profiles. **b)** t-SNE embedding based on transcriptomic molecular profiles. **c)** Heatmap showing relative activity of cancer-relevant signaling pathways across cell models.

3.3 Quality Control Metrics

After establishing the final dataset, I evaluated various quality control (QC) metrics for each molecular modality to ensure the remaining samples were reliable. For the transcriptomic data, I examined the metrics provided by the processing pipeline which included the number of processed reads and the pseudoalignment rate. The number of processed reads reflects the total sequencing depth and is essential for accurate gene expression quantification. RNA-seq libraries were sequenced to a depth ranging from 45-97 million processed reads (median: 64.3M) **(Fig. 3a, left)**. According to ENCODE consortium guidelines, a sequencing depth above 30 million reads per sample is sufficient for robust quantification of gene-level expression in bulk RNA-seq[116,117]. Our sequencing depth median exceeded these commonly recommended thresholds, indicating a consistently high sequencing depth, and confirmation that all samples were sequenced at a depth sufficient for reliable downstream analyses. The pseudoalignment rate represents the percentage of reads that were assigned to a transcript during quantification. All samples exhibited high pseudoalignment rates of >85% (median=91.1%), which indicates a good library quality and proper transcript representation **(Fig. 3a, right)**.[118]

To assess the consistency between the transcriptomic and proteomic data, I compared the RNA-seq expression levels with the protein abundance values for each cell line. I computed the Pearson correlation coefficients for each comparison, and the values all ranged between 0.34-0.46. This aligns with literature where ranges should fall between 0.3-0.5, reflecting the expected level of concordance and supporting the suitability of models for downstream analyses[119,120] **(Fig. 3b)**. I next evaluated the whole-exome data to ensure the characteristic copy number variations (CNV) and mutations were identified. It should be noted that as matched samples were unavailable for these cell lines, reference normal tissue from Lai et al. was used as a surrogate

for CNV inference and somatic variant calling. As expected, the HGSC models exhibited a high level of copy-number alterations, with widespread amplifications and deletions across the genome. This highlights genomic instability, which is a canonical feature of HGSC, as it is related to disrupted DNA repair pathways[121] **(Fig. 3c)**. To further evaluate the genomic profile of the cell line models, I selected relevant driver genes and assessed their mutations across samples. When analyzing the sequenced whole-exome lines, all HGSC samples had TP53 mutations, and 5 showed BRCA2 mutations. BRCA1 mutations were also found. These findings are validating, aligning with the canonical mutational landscape of HGSC[122]**(Fig. 3d)**. The Non-HGSC lines, including clear cell lines (TOV21G, TOV3392D), a Mucinous line (TOV2414) and an Endometrioid line (TOV112D) all exhibited KRAS mutations, linking to the MAPK pathway, however this mutation is only known to be frequently found in Mucinous lines. Both clear cell samples exhibited ARID1A mutations, which are frequently found in this subtype[123] **(Fig. 3e)**. Together, these findings confirm the transcriptomic, proteomic and genomic profiles are consistent with known biology, supporting their suitability as ideal preclinical models for subsequent analyses.

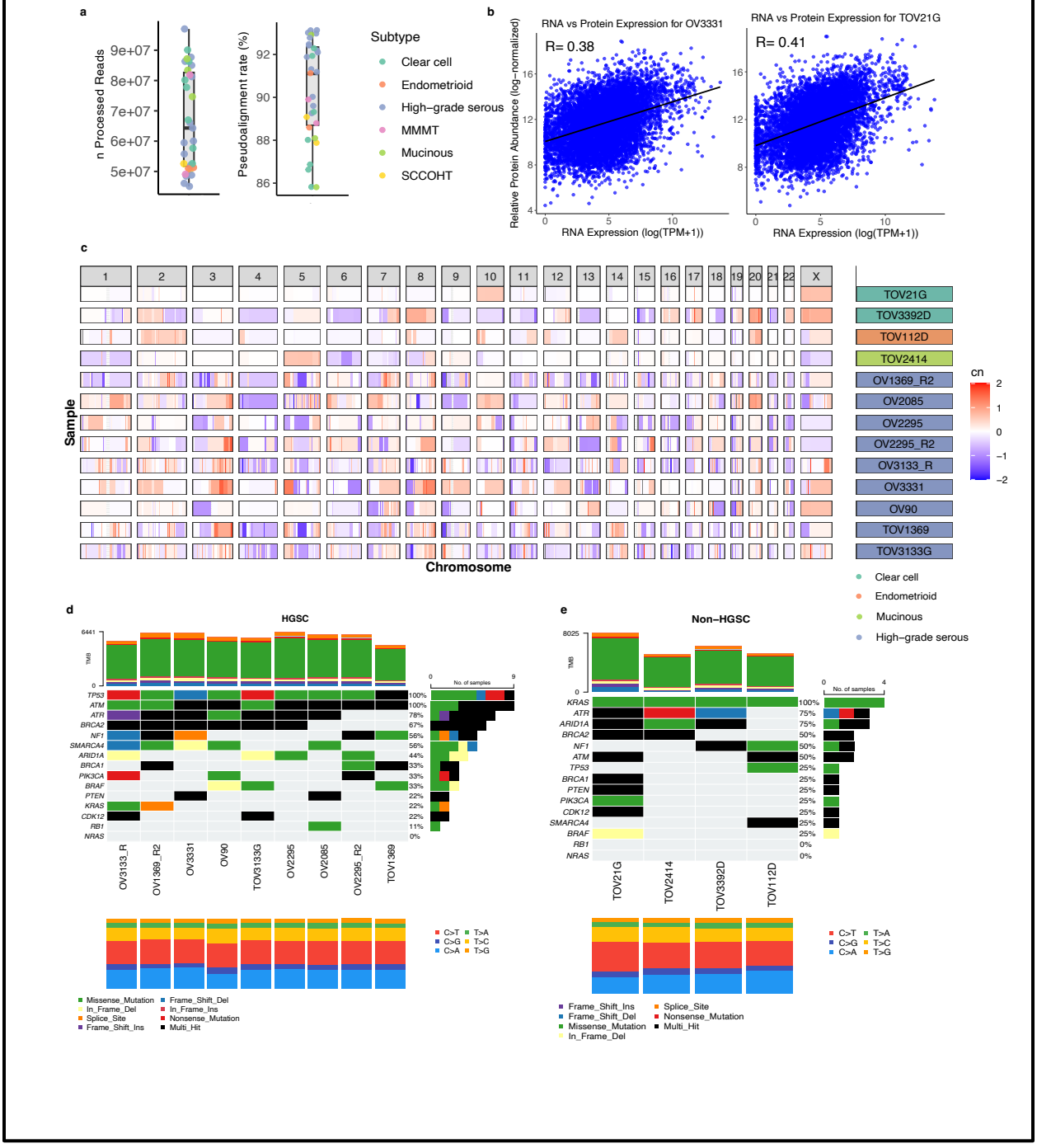
a
n Processed Reads
9e+07
8e+07
7e+07
6e+07
5e+07
Pseudoalignment rate (%)
92
90
88
86
Subtype
Clear cell
Endometrioid
High−grade serous
MMMT
Mucinous
SCCOHT
b
RNA vs Protein Expression for OV3331
R= 0.38
RNA vs Protein Expression for TOV21G
R= 0.41
Relative Protein Abundance (log−normalized)
RNA Expression (log(TPM+1))
RNA Expression (log(TPM+1))
c
Sample
Chromosome
TOV21G
TOV3392D
TOV112D
TOV2414
OV1369_R2
OV2085
OV2295
OV2295_R2
OV3133_R
OV3331
OV90
TOV1369
TOV3133G
cn
Clear cell
Endometrioid
Mucinous
High−grade serous
d
HGSC
TMB
No. of samples
TP53
ATM
ATR
BRCA2
NF1
SMARCA4
ARID1A
BRCA1
PIK3CA
BRAF
PTEN
KRAS
CDK12
RB1
NRAS
OV3133_R
OV1369_R2
OV3331
OV90
TOV3133G
OV2295
OV2085
OV2295_R2
TOV1369
C>T
C>G
C>A
T>A
T>C
T>G
Missense_Mutation
In_Frame_Del
Splice_Site
Frame_Shift_Ins
Frame_Shift_Del
In_Frame_Ins
Nonsense_Mutation
Multi_Hit
e
Non−HGSC
TMB
No. of samples
KRAS
ATR
ARID1A
BRCA2
NF1
ATM
TP53
BRCA1
PTEN
PIK3CA
CDK12
SMARCA4
BRAF
RB1
NRAS
TOV21G
TOV2414
TOV3392D
TOV112D
C>T
C>G
C>A
T>A
T>C
T>G
Frame_Shift_Ins
Frame_Shift_Del
Missense_Mutation
In_Frame_Del
Splice_Site
Nonsense_Mutation
Multi_Hit

Figure 3. Multi-omic characterization and quality control of ovarian cancer cell lines. a) Boxplots display the number of processed reads (n_processed, left), and the percentage of pseudoaligned reads (p_pseudoaligned, right) for each sample, coloured by subtype. Each point represents an individual sample. **b)** Comparison of log-normalized RNA expression (log TPM + 1) and proteomic relative protein abundance for HGSC line OV3331 and the clear cell line TOV21G. **c)** Copy-number alterations displaying amplifications or deletions across all chromosomes for 2 clear cell lines, 1 Endometrioid, 1 Mucinous and 9 HGSC lines. **d-e)** Mutational landscape of cancer-relevant driver genes across HGSC models and Non-HGSC models, respectively.

Chapter 4: Identifying distinct expression profiles associated with ovarian cancer subtypes

This chapter examines whether ovarian cancer models recapitulate known histological subtype distinctions at the molecular level. Using integrated transcriptomic, proteomic and whole-exome profiling, subtype-specific expression patterns are assessed.

4.1 Spearman Correlation Analysis

I next investigated if distinct expression patterns can effectively distinguish ovarian cancer subtypes, and whether those distinctions either align with, or unveil novel disease biology. I started by evaluating the similarity of the models' expression profiles to confirm that our data captures subtype-specific features through a Spearman rank correlation analysis. I calculated the correlation for all pairwise samples and performed hierarchical clustering to visualize subtype-specific patterns. Samples of each subtype tend to cluster together. HGSC lines comprise most of our samples and are generally well correlated. However, three lines classified as HGSC (OV866-2, OV1369-R2, OV1946) clustered with CCC lines, and one line (OV90) clustered with Mucinous lines (VOA8762/VOA8771). The mucinous line TOV2414 clustered with CCC lines, and Endometrioid and SCCOHT lines formed a cluster together, which might indicate molecular similarity between these samples and/or subtypes **(Fig. 4)**. The intra-subtype variations will be discussed further in Chapter 5, but overall, I observed subtype-specific clustering among our samples, generally supporting the classification of each line and consistent with the expectation that histological subtype is the largest source of variation among the models.

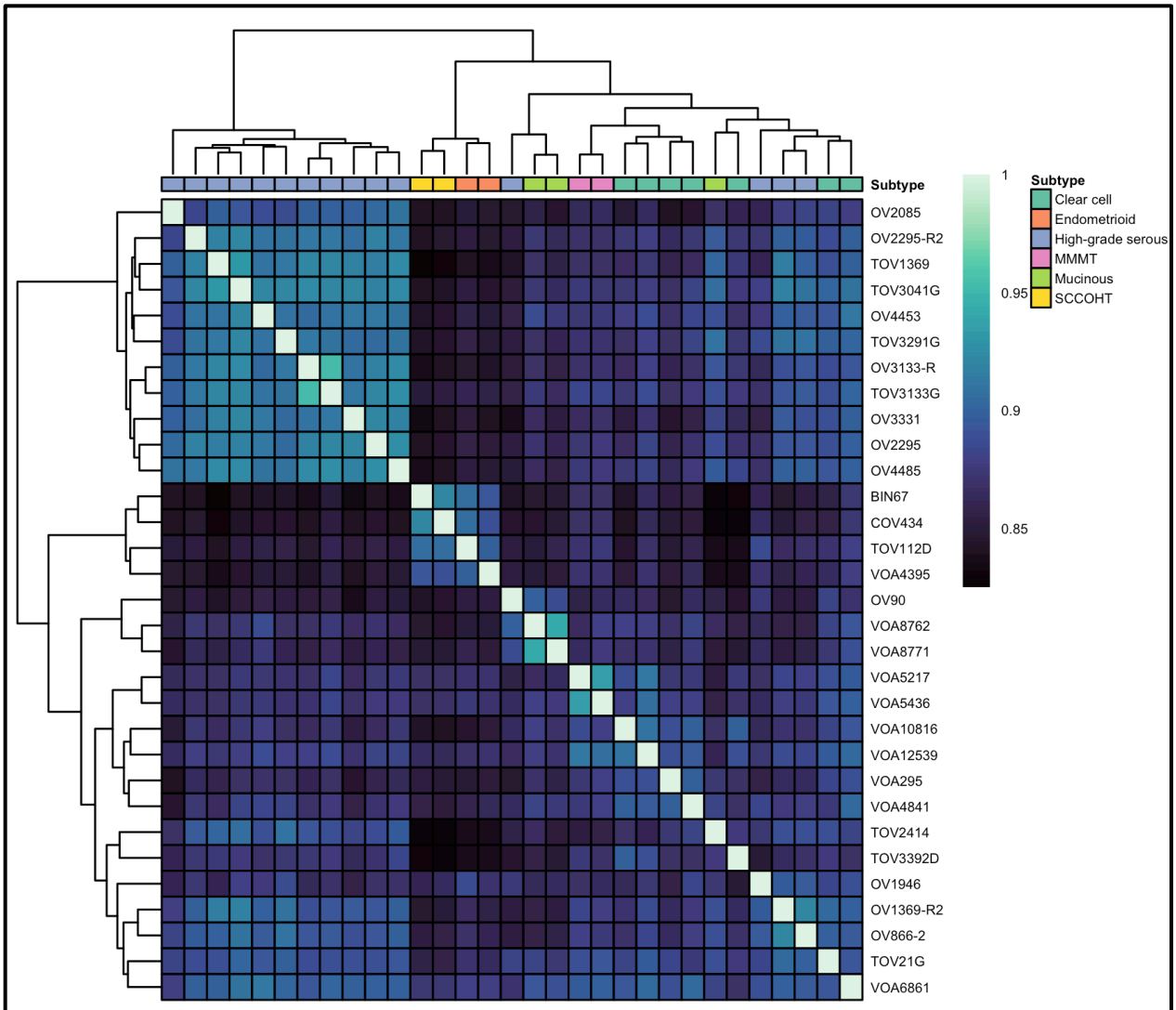

Figure 4. Hierarchical clustering heatmap based on the Spearman correlation of RNA expression profiles across ovarian cancer cell lines. Clustering reflects transcriptomic similarity between models and highlights transcriptomic similarity between samples.

4.2 Dimensionality Reduction

Dimensionality reduction techniques were performed again, and subtype-specific grouping was also observed. Through a t-SNE, I saw clear separation of cell lines based on their subtype, with samples of the same subtype aggregating together. The PCA plot further confirmed subtype-specific aggregation, as samples of each subtype formed groups across the first and second principal components, reflecting distinct gene expression profiles for each **(Fig. 5a/b)**. This clustering suggests that gene expression patterns are subtype-specific. These analyses were repeated using protein abundance values and similar subtype-specific grouping was observed **(Fig. 5d/e)**. These results highlight transcriptional and proteomic similarity among subtypes of ovarian cancer and distinct gene expression patterns between subtypes.

In both transcriptomic and proteomic profiles, I observed Endometrioid samples and SCCOHT samples clustering near each other, consistent with the Spearman correlation analysis. Despite the overall subtype-specific clustering present among the samples, some few samples clustered outside of their annotated subtype. For example, the Mucinous sample TOV2414 consistently grouped with HGSC models in both the transcriptomic and proteomic embeddings. This could indicate molecular similarity or misclassification of the Mucinous sample, but would need further investigation to confirm. To further quantify the transcriptional separation between subtypes, I calculated the average silhouette widths for each subtype group based on the PCA embeddings. A silhouette width quantifies how similar a sample is to other samples within its assigned group compared to samples in other neighbouring groups. Silhouette width values range from -1 to 1, where values close to 1 indicate that samples are matching to their own subtype and clearly separating from others. Values near 0 indicate overlap between groups, and negative values suggest potential misclassification where samples are assigned to subtypes other than their

own.[124,125] MMMT, SCCOHT, HGSC, and Endometrioid samples exhibited positive silhouette scores, ranging from 0.3-0.8, consistent with their similarity. Notably, Mucinous and CCC models both had silhouette widths near 0, suggesting a more elevated intra-subtype variability **(Fig. 5c)**. Similar results were observed with the proteomic profiles, however Endometrioid and SCCOHT samples had negative silhouette width scores, indicating that they group to other subtypes more than their own **(Fig. 5f)**.

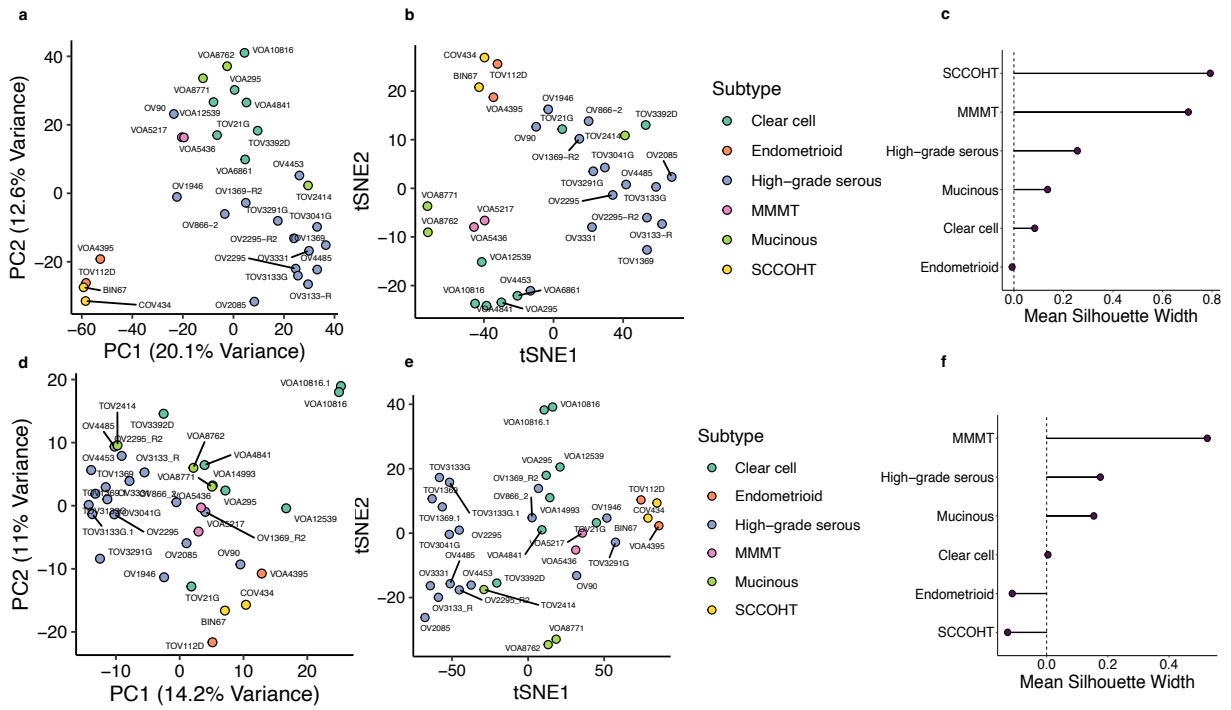

Figure 5. Dimensionality reduction and clustering for RNA-seq and proteomic datasets.

a) Principal component analysis (PCA)**, b)** t-distributed stochastic neighbor embedding (t-SNE)**, c)** silhouette width analyses based on RNA expression. Mean silhouette widths of samples by subtype based on the first 10 principal component embeddings. **d–f)** Corresponding analyses based on proteomic profiles.

4.3 Euclidean Distances

To further assess the global similarity of the samples, I generated a Euclidean distance heatmap using the first 10 principal components which capture approximately 80% of the variance within the samples. While the Spearman correlation heatmap **(Fig. 4)** captures concordance in the relative ranking of gene expression patterns between samples regardless of magnitude, the Euclidean distance heatmap measures the magnitude of expression differences between samples in the reduced PCA space, providing a complementary view of sample similarity that is sensitive to actual expression levels. The hierarchical clustering of the distances revealed evident grouping of the models according to their annotated subtypes, consistent with the PCA and t-SNE analyses. However, there are still some samples that cluster outside of their assigned subtype. Consistent with our previous analyses, OV90 does not cluster with the rest of the HGSC samples, indicating that there is some variability, and instead clusters with the same MC samples from the Spearman Correlation analysis. The MC sample (TOV2414) clusters with HGSC and CCC samples. The EC and SCCOHT samples also cluster together, consistent with the Spearman correlation and dimensionality reduction analyses **(Fig. 6)**.

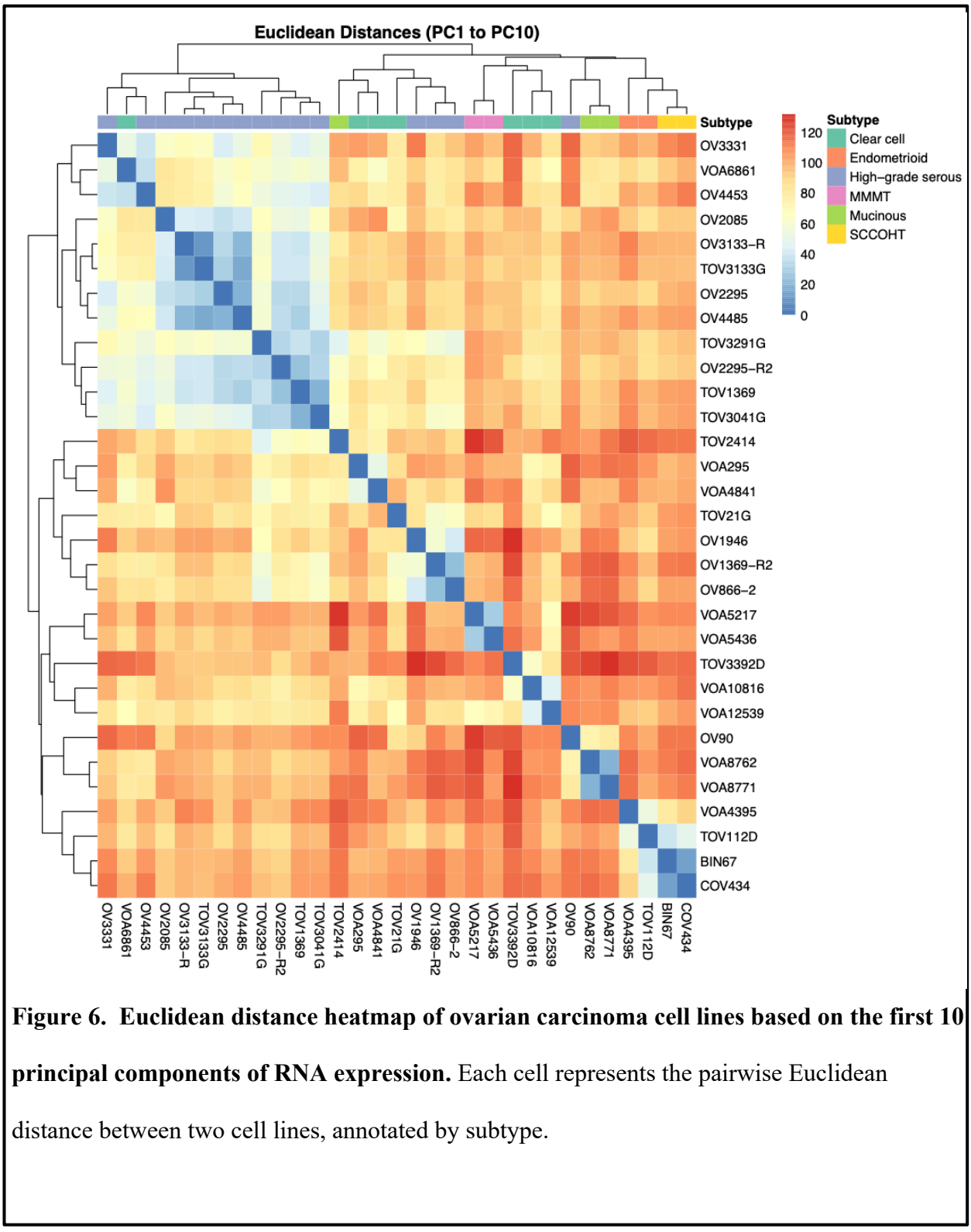

Figure 6. Euclidean distance heatmap of ovarian carcinoma cell lines based on the first 10 principal components of RNA expression. Each cell represents the pairwise Euclidean distance between two cell lines, annotated by subtype.

4.4 Identifying molecular signatures

To define molecular signatures for each subtype, I performed differential expression analyses to compare the expression profiles of each subtype against all others combined. This one-versus-all design was chosen to identify genes that distinguish each subtype from the broader heterogeneous landscape of ovarian cancer, which is most reflective of a real-world classification context. However, given that HGSC comprises the largest proportion of our dataset (n=15), the "other" group is inherently HGSC-dominated. To mitigate this, signatures were restricted to positively upregulated genes within each subtype of interest, ensuring that the resulting signatures reflect what is molecularly distinct about each subtype, rather than what is merely different from an HGSC-enriched background. Our differential expression analysis produced signatures of hundreds of genes significantly associated with each of the subtypes (FDR-adjusted $p<0.05$, logFC > 1). Genes were ranked by adjusted *p*-value, and the top 100 genes were selected for each subtype to form the respective subtype-specific signatures. The expression patterns of each signature are evident, supporting the high specificity and sensitivity of the identified gene signatures **(Fig. 7)**.

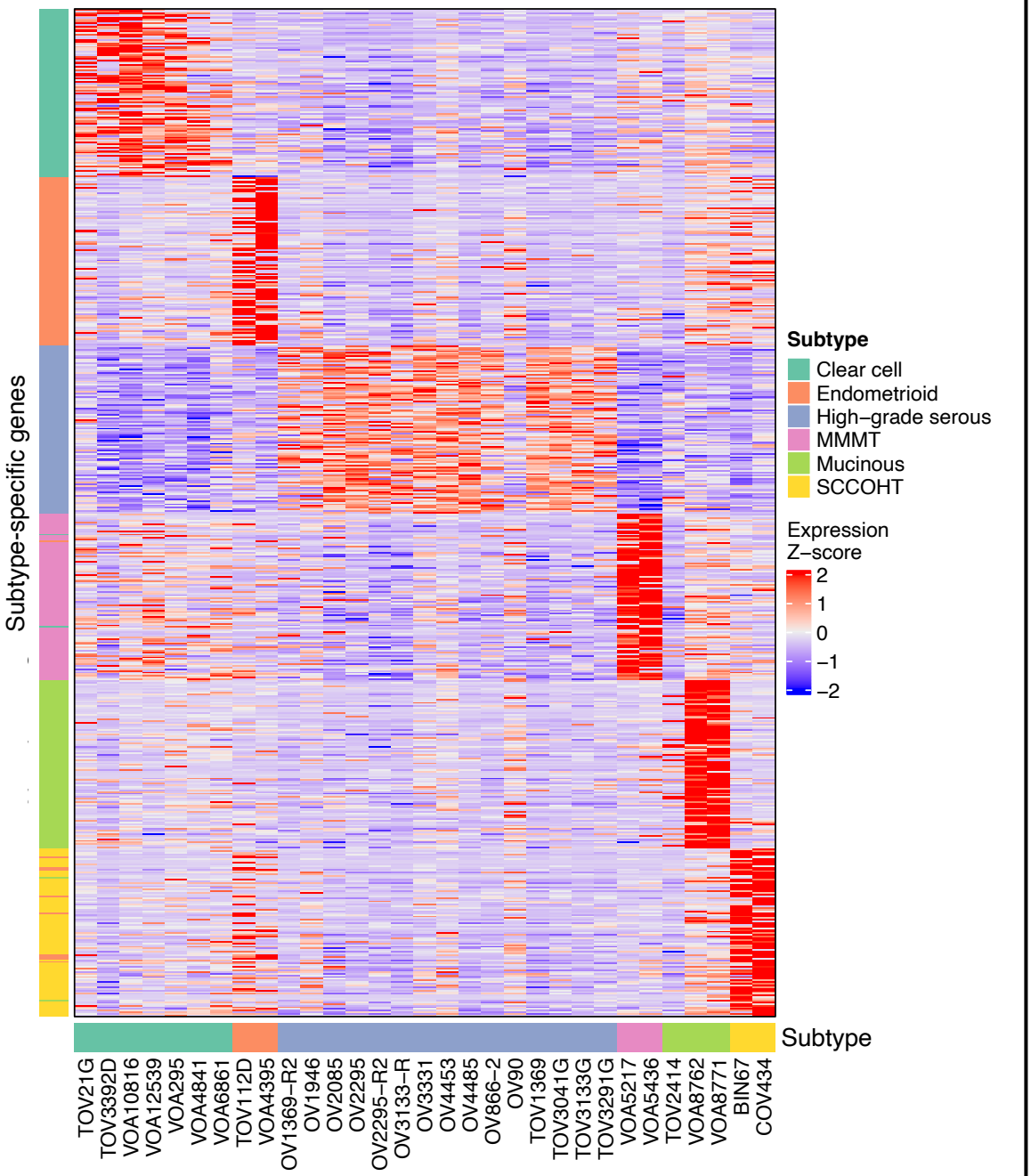

Figure 7. Transcriptomic molecular signature. Differential gene expression analysis was performed using DESeq2 and genes with an adjusted *p*-value < 0.05 and $|\log_2 FC| > 1$ were considered significant. The heatmap shows the Z-score scaled expression of the top 100 genes per subtype across all cell lines, ranked by adjusted *p*-value and arranged by subtype

We further generated Volcano Plots with the top 30 upregulated genes to visually highlight the most significantly upregulated genes within each subtype for key subtype-specific markers **(Fig. 8)**.

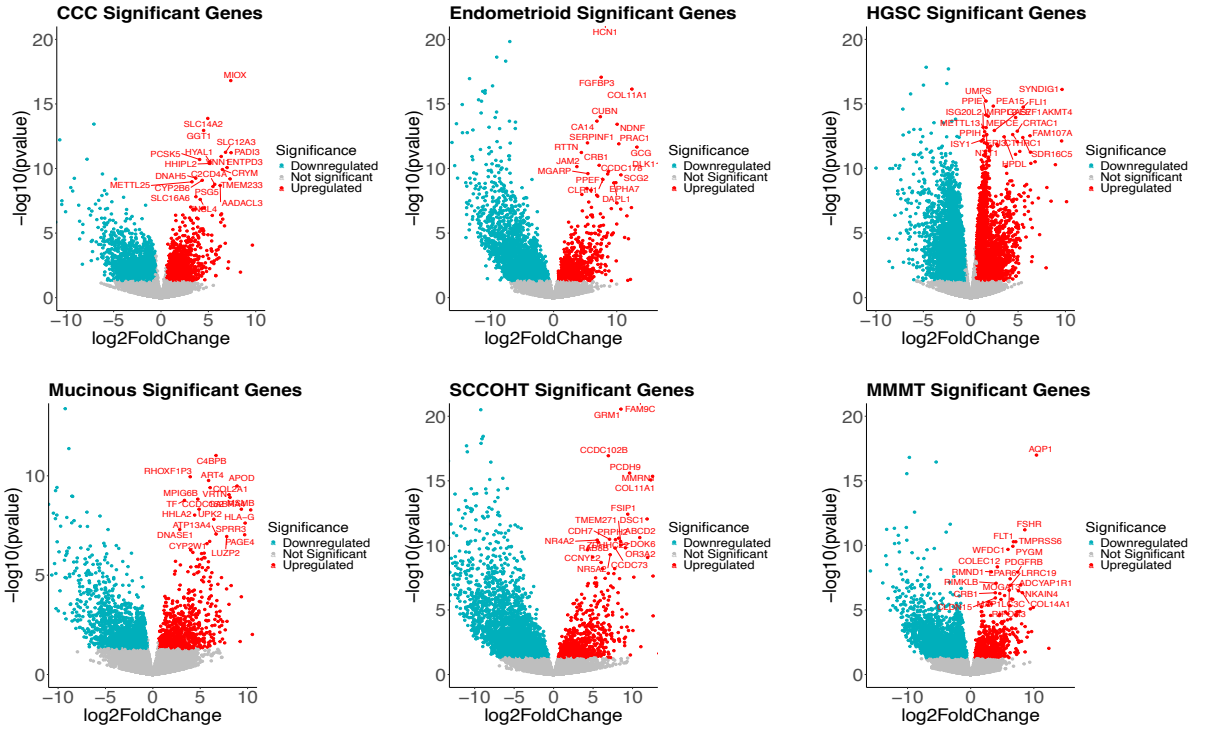

Figure 8. Most significant upregulated genes for each subtype from differential gene expression analysis. Volcano plots showing differential gene expression results for each ovarian cancer subtype, highlighting the top 30 significantly upregulated genes. Genes are plotted according to log_2FC (x-axis) and -log10 adjusted *p*-value (y-axis) derived from DESeq2 analysis.

4.5 Gene Ontology Terms

Then, to begin understanding the differences between subtypes, I identified the biological pathways associated with the most strongly upregulated genes from the differential expression analysis. Gene Ontology (GO) overrepresentation analysis was used to determine whether a greater number of signature genes mapped to specific GO terms than expected by chance **(Fig. 9)**.

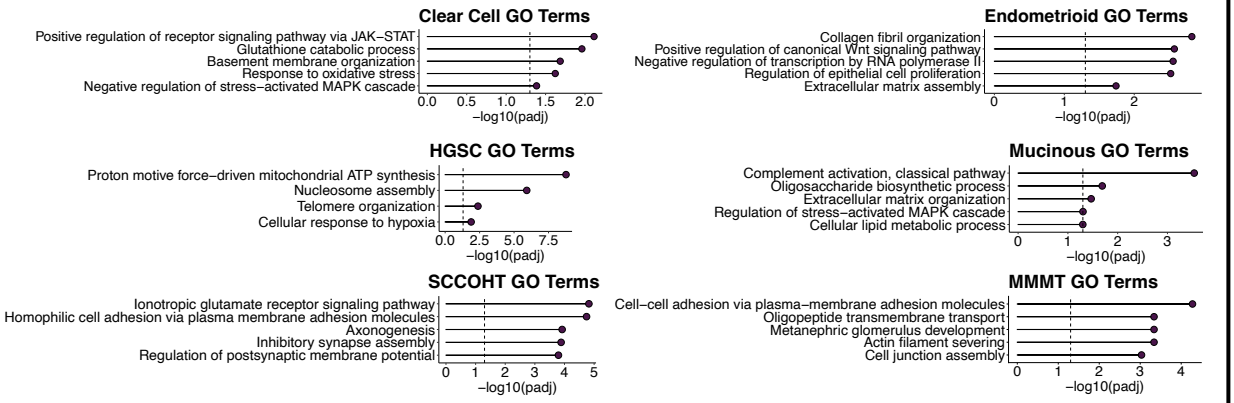

Figure 9. Gene Ontology (GO) term overrepresentation analysis of each ovarian cancer subtype. Plots showing significantly enriched GO biology process terms for each ovarian cancer subtype based on subtype-specific differentially expressed genes. Enrichment was performed using the top upregulated genes identified from the RNA-seq differential expression analysis. The x-axis represents -log10 adjusted *p* values (elim adjusted), with the dashed vertical line indicating the significance threshold.

4.5.1 HGSC results

MRPL2 and MCUR1 were identified as significantly upregulated genes in our HGSC samples compared to other subtypes. Previous studies show MCUR1 to be overexpressed in breast cancers, with highest expression in triple-negative breast cancer, and to be associated with poor overall survival in patients.[126] Furthermore, MCUR1 was found to be overexpressed in hepatocellular carcinoma, and promoted cancer cell proliferation[127,128]. However, the complete role of MCUR1 has not yet been investigated in ovarian carcinomas. One of the most expressed GO terms identified was nucleosome assembly, which is likely associated with high levels of proliferation in HGSC **(Fig. 9)**. Numerous histone proteins were identified through this GO term, such as H4C15, H4C2, H4C3, H4C4, H4C6, H4C8, and H4C9. This is consistent with the highly proliferative nature of HGSC, which requires extensive DNA replication and chromatin packaging.[129] This is also consistent with the genomic instability that characterizes this subtype, as high proliferation reflects rapidly dividing and poorly differentiated cells. The HGSC signature was also enriched with genes related to double-strand break (DSB) repair. HGSC is often characterized by homologous recombination (HR) deficiency, especially in cases where BRCA1/2 mutations are present[130]. However, recent studies suggest that tumours with HR deficiencies may compensate for this by using alternative DNA repair mechanisms, such as NHEJ or replication stress response pathways[131]. The genes associated with this GO term included several DSB-associated genes, such as RAD1, XRCC2, PARP1, TP53, and multiple MCM complex components. GO terms for these samples also revealed enrichment in genes associated with cellular response to hypoxia **(Fig. 9)**. Hypoxia-inducible factors (HIFs) are involved in the adaptation of cancer cells to low oxygen levels[132,133], and I observed an upregulation of canonical HIF target genes including EGLN3, ENO1, PDK3 and HILPDA, and

stress response regulators such as TP53 and STC2[134]. These genes participate in metabolic reprogramming, resistance to apoptosis and enhanced survival in low oxygen environments, all features associated with tumour aggressiveness and therapy resistance.[135–137] Despite being cultured under normoxic conditions, these cell lines retained enrichment of hypoxia-associated genes, suggesting tumour cell-intrinsic activation of HIF signalling pathways, consistent with prior reports of HIF dysregulation in HGSC[138].

4.5.2 CCC results

MIOX (Myo-Inositol Oxygenase) was identified as a significantly upregulated gene among the CCC samples **(Fig. 8)**. MIOX has previously been identified in clear-cell renal carcinoma[139], and various studies identify molecular similarities between clear cell ovarian carcinoma and clear-cell renal carcinoma. These similarities include the dysregulation of the PI3K/mTOR proliferative signalling pathway, a hypoxia-like mRNA profile and frequent disruptions in the SWI-SNF chromatin remodeling complex[140].[141,142] GO terms for the CCC samples revealed positive regulation of the receptor signaling pathway via JAK-STAT suggesting a pro-inflammatory and cytokine-driven phenotype of the samples **(Fig. 9)**. IL6 was identified from this GO term, and due to it being a potent activator of STAT3, it is presumed to be associated with immune evasion and tumour proliferation[143]. Basement membrane organization was also among the enriched GO terms **(Fig. 9)**, consistent with its histological hallmark of ECM remodeling and stromal hyalinization[144]. LAMA1 and LAMC1 were the genes associated with this GO term, and their upregulation suggests laminin-driven invasion and apoptosis resistance and were previously reported in CCC lines.[145]

4.5.3 MC results

Among the MC samples, C4BPB, a gene encoding the regulation of the complement activation pathway[146], was significantly upregulated in our differential expression analysis **(Fig. 8)**. Consistent with this finding, GO term enrichment analysis of the MC signature revealed significant enrichment of the complement activation, classical pathway **(Fig. 9)**. Other genes identified with this GO term include C4A, C4B, and C4BPA. This suggests the activation of complement related immune processes in the MC samples. C4A and C4B are effector genes that are required for C3 convertase formation and downstream complement activation[147]. C4BPA and C4BPB are complement regulatory proteins that prevent the pathway from being destructive and instead are tuned towards a tumour-adaptive immune modulation[148]. Together, the expression of complement effector and regulator genes suggests tumour cell-intrinsic expression of classical complement pathway components. Confirmation of functional complement activation would require validation in systems containing immune cells. MC samples also had significant enrichment of genes associated with the oligosaccharide biosynthetic process **(Fig. 9)**. This enrichment is consistent with the mucinous phenotype of this subtype, as extensive oligosaccharide biosynthesis is required for mucin glycosylation and secretion[149]. The genes B3GALT2, FUT3, ST8SIA4 and TREH were upregulated, suggesting increased glycan remodeling activity. Alterations in glycosylation are a hallmark of mucinous tumours, contributing to the production of mucin within the cytoplasm, immune evasion and altered cell-cell interactions.[150–152]

4.5.4 EC results

The differential expression for EC samples confirmed significant upregulation of DLK1, a gene that plays a role in the development and progression of endometriosis[153,154] **(Fig. 8)**. Specifically, DLK1 participates in inhibiting the Notch signalling pathway, which is essential for cell maintenance and tissue homeostasis. Dysregulation of DLK1-mediated Notch signalling may contribute to abnormal cellular differentiation, endometriosis lesions, resistance to apoptosis and immune evasion.[155] This aligns with the molecular similarities between endometriosis and endometrioid ovarian carcinoma, and therefore may represent a molecular link between the two.[156] COL11A1 (alpha-1 chain of type XI collagen) was one of the most significantly upregulated genes in the differential expression analysis **(Fig. 8)**, along with other fibrillar collagen genes (COL11A1, COL1A2, COL3A1, COL5A2, and COLGALT2) that were flagged in the significant enrichment of collagen fibril organization **(Fig. 9)**. Enhanced expression of fibrillar collagen genes suggests active ECM remodeling, which is linked to invasion, resistance and stromal reprogramming.[157,158] It may also further reflect the fibrotic environment common to both endometriosis and endometrioid ovarian carcinoma[156]. GO terms also revealed significant enrichment of positive regulation of the canonical WNT pathway among the EC samples **(Fig. 9)**. Genes involved in both activation (TNKS[159],YAP1[160]) and inhibition (DACT1[161], DKK2[162]) were flagged for this GO term. This aligns with previous studies showing that dysregulated WNT/β-catenin signaling plays a role in endometrioid ovarian tumorigenesis.

4.6 Developing a concordant RNA-protein signature

To strengthen confidence in the specific markers identified from our signatures, I selected the top genes that were consistently upregulated in both the transcriptomic and proteomic data to form a

concordant RNA-protein signature. From each RNA-derived signature, the top genes were selected (ranked by adjusted *p*-value). The method for differential expression implemented in DESeq2 is not applicable to proteomic data, therefore a t-test with the same significance thresholds (FDR-adjusted $p<0.05$, logFC > 1) was applied to identify the significantly upregulated genes for the protein data. By intersecting these proteins with those significantly upregulated in the RNA-seq data, I identified a conserved molecular signature associated with each subtype. This intersection defines a consensus RNA-protein signature that increases confidence in subtype-specific markers by retaining features present at both molecular levels, including known and potentially novel candidates **(Fig. 10)**.

4.7 Signature genes are linked to relevant features of subtype biology

GGT1 (type I gamma-glutamyltransferase) was identified as part of the concordant CCC subtype signature **(Fig. 10)**. GGT1, involved in glutathione metabolism[163], has been linked to chemoresistance and tumour aggressiveness when overexpressed. Previous work has shown that GGT1 inhibition might reduce cell migration and improve response to chemotherapy agents such as cisplatin.[164] C4BPB was identified in the concordant MC signature, further supporting the link between this subtype and the complement activation pathway, as discussed previously **(Fig. 10)**. COL3A1 (Collagen type III alpha 1 chain) was identified as part of the Endometrioid signature **(Fig. 10)**, and the Kaplan-Meier survival analysis revealed that COL3A1 expression in epithelial cancers, including Endometrioid, is associated with shorter patient survival and tumour aggressiveness[165]. WFDC2 (WAP four-disulfide core domain 2), an FDA approved biomarker for ovarian cancer was identified as part of the HGSC subtype signature **(Fig. 10)**. WFDC2, also called HE4, is frequently expressed in ovarian cancer, specifically HGSC, and is implicated in tumour progression, extracellular matrix remodeling and pro-survival. Elevated WFDC2

expression is also associated with lower survival in ovarian cancer patients.[166–168] Several of the MMMT gene signatures (CPQ, CYB5R3, PTPN1, THBS3) have been involved in tumour progression[169], cancer cell colonization[170,171] and poor prognosis[172,173] in varying cancers, however none have been reported in relation to MMMT. Therefore, their enrichment in our signature suggests potentially novel therapeutic targets that can warrant future investigation. HGF (Hepatocyte Growth Factor) was flagged as a signature for the SCCOHT subtype **(Fig. 10)**. HGF is a ligand for the MET receptor, and previous work has shown that SCCOHT samples are sensitive to c-MET inhibitors, as they can attenuate tumour growth in SCCOHT lines.[174] Additional genes such as MID1, MTM1, ADGRL2 may represent novel gene targets for this subtype.

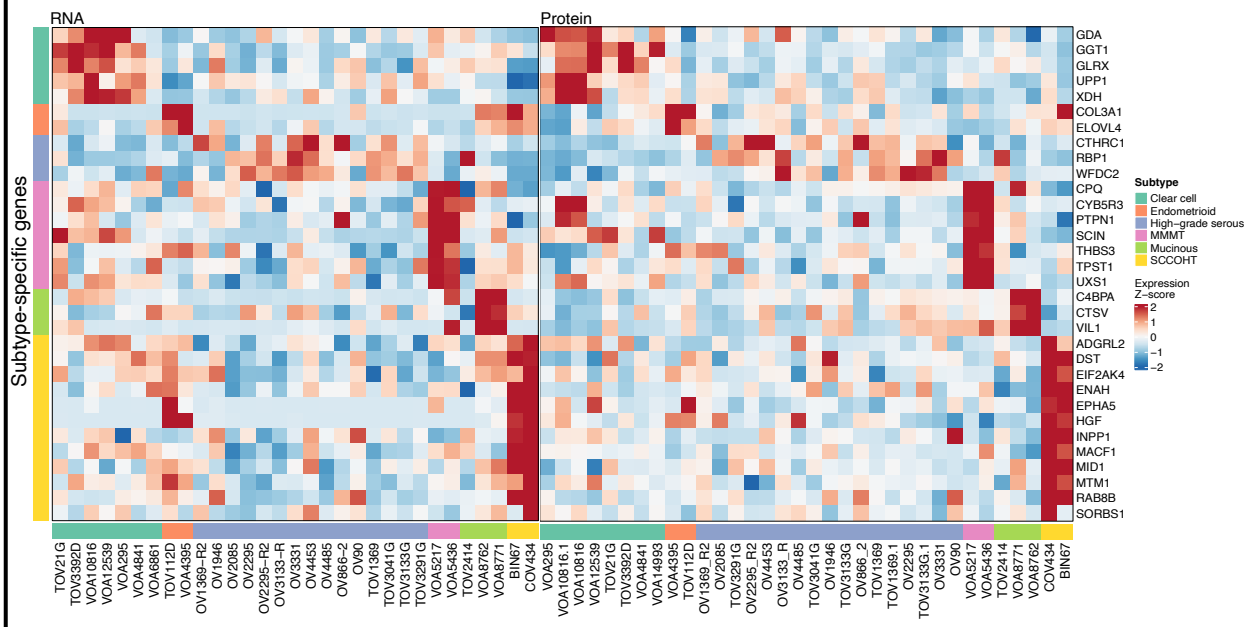

Figure 10. Identification of concordant RNA/protein molecular signatures and expression patterns across ovarian cancer subtypes. Integrated RNA-protein signature derived from concordant genes and proteins meeting the same significance thresholds (*padj* < 0.05, $|\log_2 FC| > 1$ by DESeq2 for RNA and t-test for proteomics).

4.8 ADC Targets

Having identified subtype-specific molecular signatures and pathway differences across the ovarian cancer models, I next examined whether established therapeutic targets also exhibited subtype-associated expression patterns. I quantified the protein abundance of antibody drug conjugate (ADC) targets currently in clinical trials across the ovarian cancer cell lines to identify potential subtype-specific therapeutic targets **(Fig. 11)**. The HGSC samples show the highest abundance of TACSTD2, which encodes the protein Trop-2. This cell-surface glycoprotein is involved in metastatic and proliferative signalling. Its overexpression and high tumour-specificity in many cancers makes it a successful therapeutic target for ADCs.[175,176] The MC lines have high expression of FOLR1 (Folate Receptor Alpha), excluding TOV2414. This ADC target is a cell surface receptor that mediates folate uptake and is found to be overexpressed in epithelial ovarian carcinomas. Due to its low expression in most normal tissue, it has become a valuable therapeutic target for ADCs for ovarian cancer.[177,178] This analysis can support future experimental model selection. For example, there is some variability in the ADC target expression among the CCC lines, and some lines might be better fits for certain therapeutic experiments than others. In particular, TOV3392D has a very high abundance of ERBB2 (HER2), a tyrosine kinase receptor that amplifies growth signalling through the PI3K-AKT pathway[179]. When overexpressed, HER2 drives uncontrolled proliferation in epithelial cancers and provides a therapeutic vulnerability, as HER2-targeting ADCs have shown antitumour activity in HER2-high tumours[180,181]. Furthermore, the CCC sample VOA12539 has high expression of the target SLC34A2, a sodium dependant phosphate transporter that is frequently overexpressed on the surface of epithelial tumours, making it a strong ADC target[182,183].

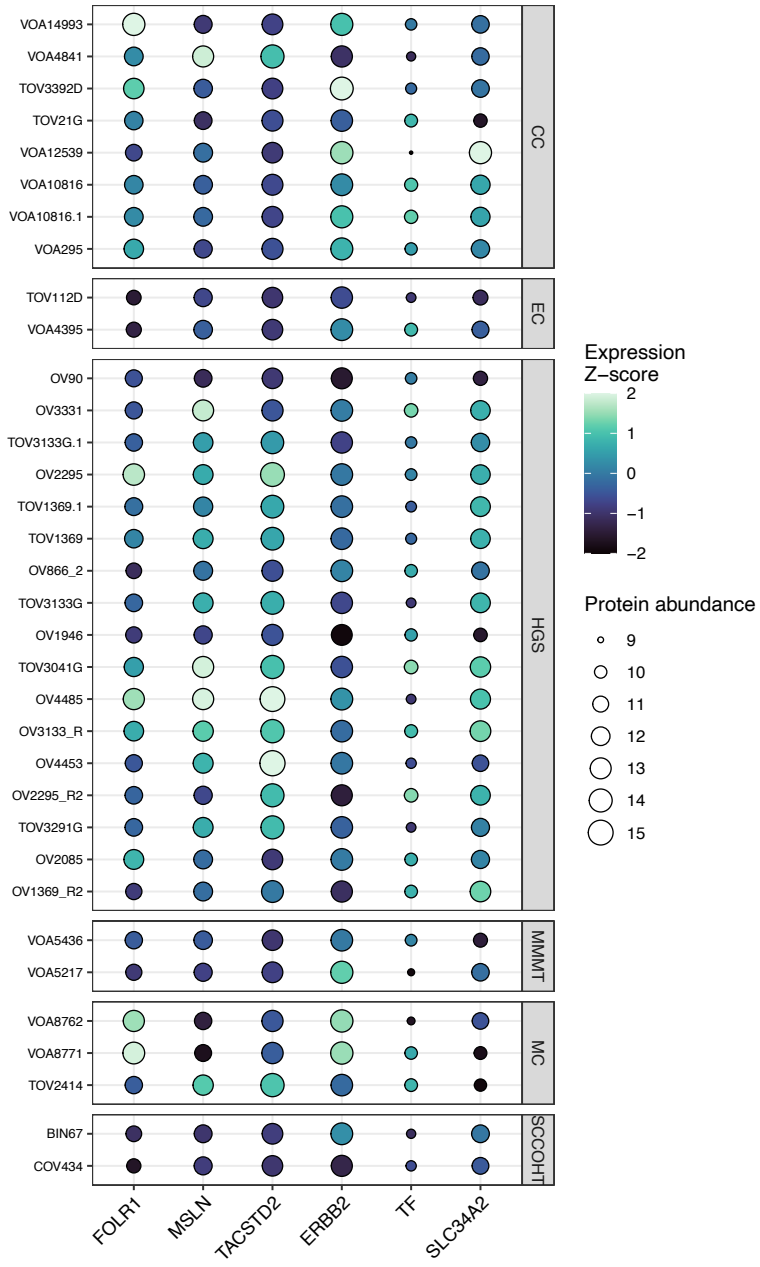

Figure 11. Antigen expression heatmap based on proteomics data showing subtype expression patterns of Antibody Drug Conjugate (ADC) targets. Each column represents an ADC target, and each row represents an ovarian cancer sample. Protein expression values were z-score normalized across samples for each target and displayed by colour intensity. Size of each point corresponds to the absolute protein abundance. Targets are currently in clinical trials.

In this chapter, I demonstrated that ovarian cancer subtypes are characterized by distinct molecular expression profiles at both the transcriptomic and proteomic levels. Unsupervised analyses such as Spearman correlation distance, PCA and t-SNE revealed subtype-specific clustering and overall concordance between the RNA and protein data. Most samples clustered according to their annotated subtype, however some deviated, reflecting intra-subtype heterogeneity. Quantitative analyses such as silhouette width and Euclidean distances confirmed separation between subtypes across both molecular layers. Differential expression analysis revealed subtype-specific molecular signatures concordant across transcriptomic and proteomic profiles. These signatures generated the top upregulated genes for each subtype which can be useful for novel therapeutic targets or as diagnostic markers. Lastly, antigen target profiling further highlighted molecular distinctions between subtypes, also supporting future therapeutic target discovery and development.

The observed deviations within subtypes motivated further investigation into intra-subtype heterogeneity, which is explored in Chapter 5 to better understand molecular heterogeneity among models of the same histological classification.

Chapter 5: Evaluating variation among models of matched subtypes

Building on the subtype-level analyses in Chapter 4, this chapter focuses on variability within individual subtypes to identify heterogeneity that may influence experimental systems and therapeutic responses.

5.1 Evaluating variation among HGSC models

I next wanted to further explore the diversity within subtypes to gain a deeper understanding of the models and the subtypes they are representing. While effectively all HGSC models shared traits consistent with the subtype, I noted distinct clustering patterns among HGSC that could reflect relevant disease heterogeneity. I sought to demonstrate the utility of this resource to understand this heterogeneity and aid in the selection of ideal cell models for specific experimental aims. To explore tangible biological properties of these models, I collected transcriptional signatures of distinct signalling pathways, biological functions, and metabolic processes. I computed gene set scores based on the ranked expression of signature genes within the sample and then compared relative (*Z*-score transformed) gene set scores among the HGSC models.[184] These gene set scores represent relative biological pathway level activity across our models, allowing us to compare the cancer relevant pathways and processes between all models, which could aid in future diagnostic and therapeutic endeavours.

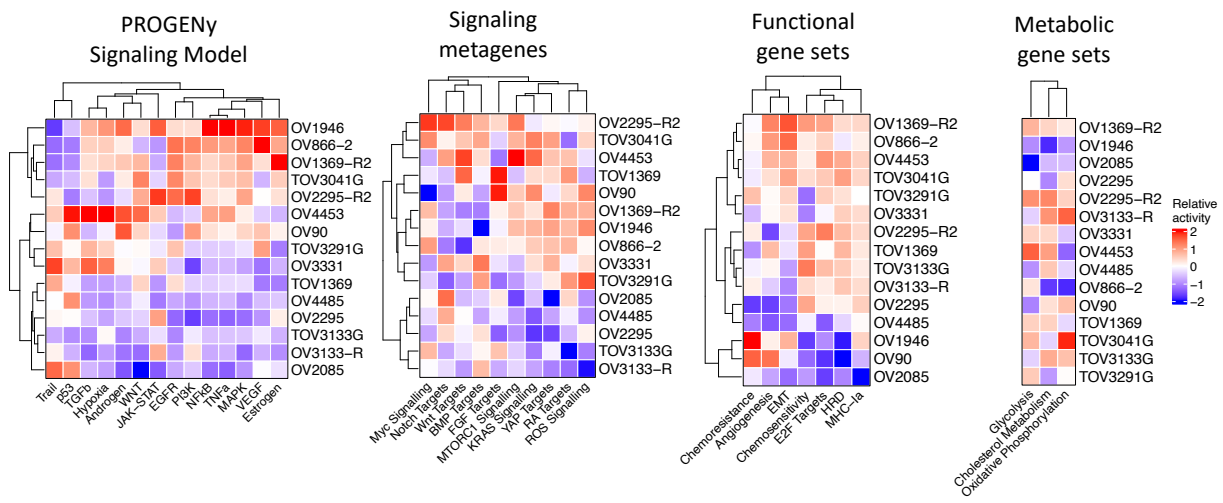

Figure 12. Heatmap of the relative activity of various gene sets across HGSC ovarian cancer samples. The gene sets are organized into groups representing the PROGENy signaling model, signaling metagenes, functional gene sets, and metabolic gene sets, highlighting distinct pathway activity patterns across samples.

I observed a cluster of models with elevated growth factor signalling through EGFR, PI3K and other receptor tyrosine kinases (RTKs). These models include OV1946, OV866, OV1369.R2, OV2295.R2, and TOV3041G **(Fig. 12)**. EGFR is most commonly activated in cancer from amplifications and point mutations at the genomic locus, as well as transcriptional upregulation or ligand overexpression[185,186]. Activation of the EGFR pathway and related RTKs can activate downstream PI3K-AKT signalling, which can promote cell survival from inhibited apoptotic pathways, uncontrolled proliferation and rapid cell-cycle progression[187,188]. The specific lines identified seem to adopt a proliferative state due to the elevated expression of growth-factor / related receptors. These models may be relevant for studies that focus on proliferative signalling, adaptive resistance mechanisms, or therapies targeting RTKs, PI3K or other relevant pathways.

The model OV4453 was unique in its particularly elevated activation of inflammatory signalling linked to NF-κB, JAK-STAT, and WNT signalling. These pathways are characteristic of an inflammatory tumour, suggesting this line is associated with an immunoregulatory signalling and may represent a useful model for studying rare, inflamed "hot" tumours and immunoregulatory signalling. These signalling pathways were only moderately active in the models associated with growth factor signalling, and a subset of models (OV2085, OV3133, TOV3133G, OV2295, OV4485, and TOV1369) had extremely low activity of both sets of signalling pathways. Chemoresistance signatures were notably high in OV1946 and OV90, which were also associated with low proliferation rates (E2F Targets). Lower proliferation has been associated with lower sensitivity to chemotherapy, as it preferably targets cells that are rapidly dividing and have higher proliferation, which might contribute to chemoresistance in these models[189,190].

To assess similarity between samples in terms of shared biological patterns, I relied on Spearman correlation, which captures concordance in relative gene expression patterns. Euclidean distances can further identify variation present within subtypes by quantifying absolute differences in expression based on the first 10 principal components. As stated earlier, HGSC samples tend to group together in various analyses such as PCA, t-SNE, Spearman correlation analyses and Euclidean distance analyses. However, we do observe intra-subtype variation among the HGSC samples. The HGSC samples OV1946, OV1369-R2 and OV866-2 formed a separate cluster from the main HGSC cluster in the Spearman correlation analyses **(Fig. 4)**. OV90 also clustered away from the main HGSC cluster **(Fig. 4)**. Evidently, OV90 displays consistent deviation across our analyses. I observed OV90 clustering with Mucinous samples (VOA8762 and VOA8771) in both the Spearman correlation and Euclidean distances analyses **(Fig. 4/6)**. OV90 was also the only HGSC sample to deviate from its transcriptomic molecular signature **(Fig. 7)**. While OV90 did

contain TP53 mutations, it also expressed a mutation in PIK3CA as well **(Fig. 3d)**, which is not typically common in HGSC[191,192]. The PI3K/AKT pathway can get dysregulated indirectly through copy-number alterations or PTEN losses[193], however specific mutations in PIK3CA are not common in HGSC.

5.1.1 Genomic pathway concordance

To complement these findings, I examined whole-exome sequencing data to relate mutational profiles with the observed transcriptional activity. OV2295_R2 has a PI3K mutation **(Fig. 3d)** and also demonstrated high relative activity of PI3K signalling in the pathway activity analyses **(Fig. 12)**, suggesting concordance between genomic and transcriptional profiles. From the genomic profile, I observed ATM (Ataxia-Telangiectasia Mutated) mutations in 100% of the HGSC lines, and ATR (ATM- and Rad3-Related) in 75% of these lines **(Fig. 3d)**. ATM and ATR are classified as the most upstream DNA-Damage Repair (DDR) kinases. ATM is primarily activated from Double-Strand Breaks (DSBs), while ATR responds to a broad spectrum of DNA damage, including DSBs, replication stress and single-strand breaks.[189,194] HGSC is characterized by extensive genomic instability and the presence of DNA damage, therefore the mutations in these DDR genes reflects a defining feature of HGSC biology. Furthermore, the mutations in the ATM and ATR signalling pathways can increase the reliance on compensatory DNA repair pathways. Since ATM and ATR are central regulators in DNA repair, their disruption can compromise canonical repair and cell-cycle control, forcing tumour cells to depend on alternative repair mechanisms. Alongside this, compensatory pathways can further contribute to genomic instability.[195,196] Recent studies have been utilizing ATR inhibitors to enhance chemotherapy and radiation response in ovarian, endometrial and cervical cancer cells. Increased expression of ATM and ATR have been previously reported to contribute to

chemoresistance as they bind competitively to DNA against platinum-based regiments.[197] This led to the discovery of combining small molecular pharmacological ATM/ATR inhibitors with platinum drugs or ionizing radiation. Results demonstrated a significant enhancement of the platinum drug response in all the gynecological lines tested when ATR was inhibited. In addition, all lines displayed an enhanced response to ionizing radiation when inhibiting either ATM or ATR, and further enhancement with co-inhibition.[194,198] Another study reported the presence of ATM mutations specifically in HGSC, and also suggested its inhibition would support the treatment of this cancer[199]. As matched normal samples were not available, these variants may include germline alterations, potentially contributing to the higher observed frequencies relative to published somatic ATM/ATR mutation rates in HGSC, that are lower than what was observed in our results (~2% of cases)[200].

When further analyzing the copy number alterations across our models, I wanted to look at specific arm level alterations and identify patterns within subtypes. I observed consistent amplifications on the 3q chromosome arm across all HGSC samples, which have been previously reported in about 50% of HGSC cases[201,202] **(Fig. 3c)**. However, OV90 is excluded from this observation, distinguishing it from other HGSC models. Amplifications on the 3q arm are implicated in tumour-driver mutations because it harbours numerous oncogenic drivers, such as PIK3CA[201] and EVI1[203]. These oncogenes support proliferation and survival, and therefore their amplification might contribute to the aggressive biology of HGSC.[202,204]

I also observed recurrent deletions on all chromosome 11p arms in the HGSC lines **(Fig. 3c)**. The 11p region harbours tumour suppressor genes including WT1, and microsatellite markers such as D11S860, that are commonly used to assess loss of heterozygosity in ovarian cancer.[205–207] Loss on this chromosomal arm may result in lack of tumour suppression and uncontrolled tumour

growth, along with increased genomic instability. Deletions on the 11p arms have been previously identified as frequent events in ovarian cancers, however their specific biological relevance for HGSC has not yet been characterized[208]. Overall, the presence of consistent 11p losses warrants future investigation on its therapeutic and diagnostic implications.

5.1.2 Mapping HGSC models to established HGSC molecular subtypes

To further characterize the heterogeneity of HGSC, I next assessed whether the models recapitulate the established molecular subtypes of HGSC defined by The Cancer Genome Atlas (TCGA)[209]. The TCGA has identified four transcriptional subtypes of HGSC: immunoreactive, differentiated, proliferative and mesenchymal. Each subtype is associated with distinct biological properties. The immunoreactive subtype is characterized by high expression of immune-related genes and reflects immune cell infiltration, the differentiated subtype exhibits tumour markers such as MUC1 and MUC16, the proliferative subtype is characterized by proliferation markers such as MCM2 and PCNA, and lastly the mesenchymal subtype has transcriptional programs linked to stromal components and has expression of HOX genes, and ANGPTL1/2.[210] To assign our HGSC samples to these subtypes, I applied the ConsensusOV framework to the HGSC subset of our data, assigning each model to a specific HGSC molecular subtype[211] **(Table 2)**. Given that cell lines inherently lack stromal and immune components, ConsensusOV classifications reflect tumour cell-intrinsic transcriptional programs, and assignment to immunoreactive or mesenchymal subtypes should be interpreted in that context.

Table 2. HGSC subtype-assignments based on ConsensusOV's pipeline. Each HGSC sample in our cohort and the assigned established TCGA subtype.

Cell Line	Subtype
OV1369-R2	Differentiated
OV1946	Mesenchymal
OV2085	Immunoreactive
OV2295	Differentiated
OV2295-R2	Immunoreactive
OV3133-R	Differentiated
OV3331	Proliferative
OV4453	Proliferative
OV4485	Differentiated
OV866-2	Mesenchymal
OV90	Immunoreactive
TOV1369	Differentiated
TOV3041G	Mesenchymal
TOV3133G	Differentiated
TOV3291G	Mesenchymal

I wanted to further contextualize the ConsensusOV subtype assignments by examining the expression patterns of the established TCGA subtype marker genes across our cell lines. While ConsensusOV classifies samples based on relative gene-pair comparisons based on a predefined signature, rather than absolute expression of individual markers, evaluating the expression of these markers still provides additional biological insight on the variation present in the HGSC samples.

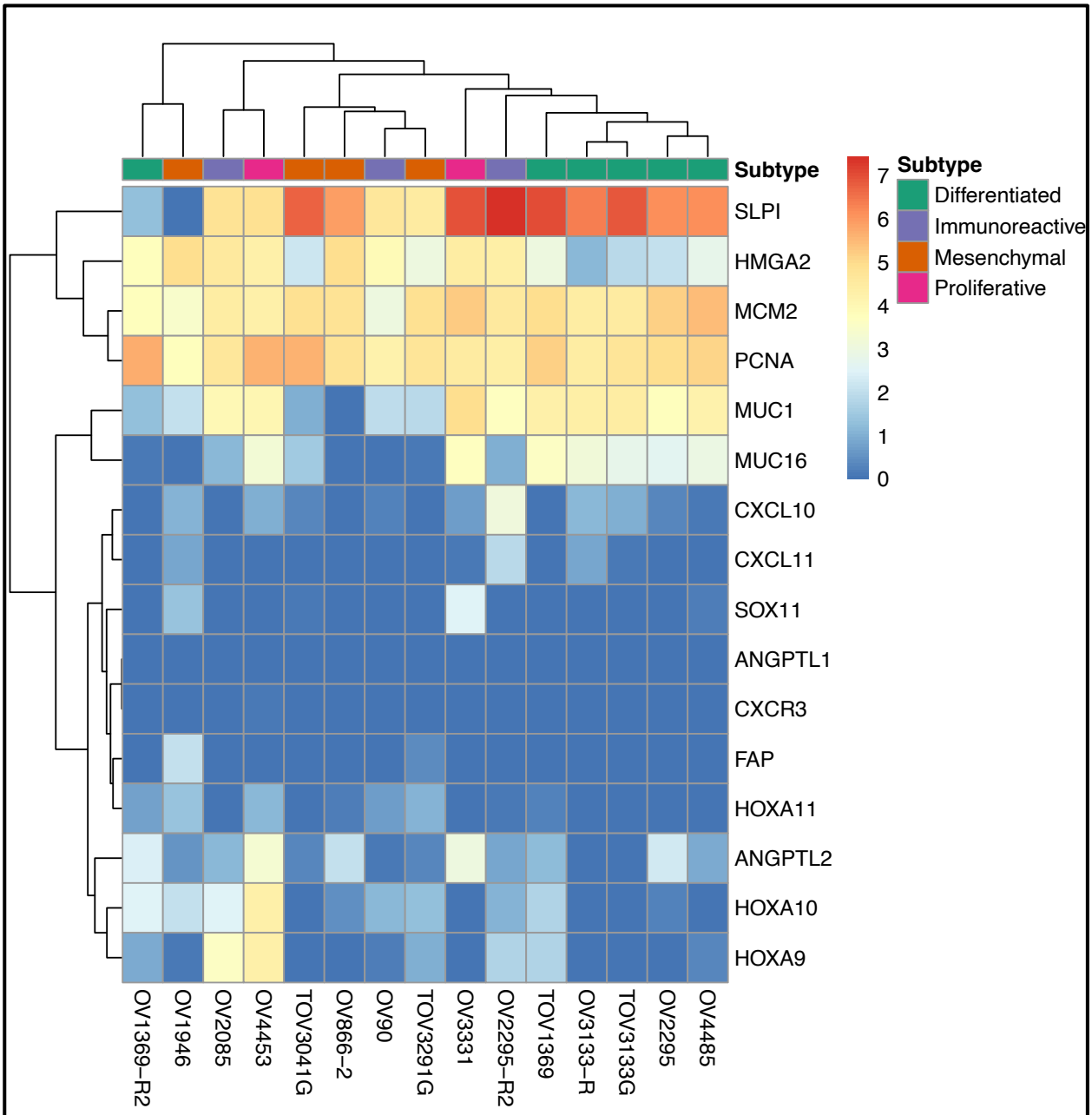

Figure 13. Transcriptomic expression of molecular subtype-specific marker genes in HGSC models. Marker genes previously defined by TCGA were assessed to visualize concordance between the ConsensusOV classification and the known TCGA subtype signatures.

5.1.2.1 Immunoreactive subtype (IMR)

The following lines were classified as immunoreactive: OV2085, OV2295-R2, OV90. OV2295-R2 showed the highest expression of the immunoreactive markers CXCL10/CXCL11 in comparison to the other HGSC lines. Although none of the lines showed expression of the other immunoreactive marker CXCR3 **(Fig. 13)**. Referring back to the relative activity gene set scores, OV2295-R2 displayed elevated relative activity of immune-associated signalling pathways such as JAK-STAT, NF-κB and TNF-α signalling **(Fig. 12)**.

5.1.2.2 Differentiated subtype (DIF)

The HGSC lines classified as differentiated include OV1369-R2, OV2295, OV3133-R, OV4485, TOV1369 and TOV3133G. All these lines consistently exhibited high expression of the markers MUC1, MUC16 and SLPI relative to other HGSC models, excluding OV1369-R2 **(Fig. 13)**. These genes are established markers of the TCGA differentiated subtype and represent mature ovarian epithelial tissue differentiation[212].

5.1.2.3 Proliferative subtype (PRO)

The proliferative lines include: OV3331 and OV4453. Both of these lines demonstrated expression of HMGA2, a transcription factor associated with proliferation and chromatin organization[213]. These lines also expressed other proliferation markers including MCM2 and PCNA. However, these markers were also expressed in other lines of HGSC, reflecting the generally high proliferative nature of HGSC. In contrast, SOX11, another proliferative marker identified by TCGA, was not highly expressed in these lines **(Fig. 13)**.

5.1.2.4 Mesenchymal subtype (MES)

The lines classified as mesenchymal include OV1946, OV866-2, TOV3041G and TOV3291G. The mesenchymal subtype is characterized by the expression of HOX genes and markers associated with increased stromal components, such as microvascular pericytes (markers ANGPTL1/ANGPTL2). OV866-2 demonstrated modestly elevated expression of ANGPTL2, however other mesenchymal markers were not consistently elevated across the HGSC lines **(Fig. 13)**.

5.2 Evaluating variation among Mucinous carcinoma models

The MC sample TOV2414 clustered with the CCC sample TOV3392D and further clustered with other HGSC and CCC samples in the Spearman correlation **(Fig. 4)**. This was further observed in all PCA and t-SNE plots **(Fig. 5)**. TOV2414 also deviated from its molecular signature **(Fig. 7)**. TOV2414 has the most elevated p53 signalling compared to the other MC samples and expresses elevated growth factor related signalling processes such as PI3K, VEGF, MAPK and EGFR, which are very characteristic of HGSC. TOV2414 also expressed a mutation in ARID1A, commonly observed in CCC lines. TOV2414 also displayed modestly elevated HRD transcriptional activity, which suggests an involvement of DNA damage and replication stress. TOV2414 also had the highest activity of MHC-1a signalling in comparison to the other MC samples **(Fig. 14)**.

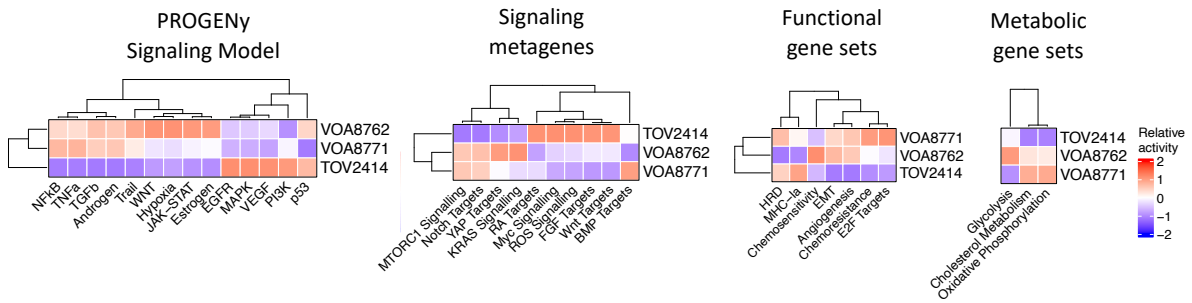

Figure 14. Heatmap of the relative activity of various gene sets across MC ovarian cancer samples. The gene sets are organized into groups representing the PROGENy signaling model, signaling metagenes, functional gene sets, and metabolic gene sets, highlighting distinct pathway activity patterns across samples.

These characteristics are more frequently observed in HGSC and CCC samples, rather than MC samples that tend to be more genomically stable and less HRD associated[214]. Lastly, TOV2414 has a mutation in ARID1A **(Fig. 3e)**, which is strongly associated and considered a hallmark in CCC. Overall, these findings suggest that TOV2414 represents a biologically distinct sample that, although annotated as MC, displays transcriptional signatures more consistent with HGSC and CCC. Among the remaining MC samples, VOA8762 displays the most elevated relative activity of inflammatory and cytokine-associated pathways including NF-κB, TNF-α, TGFb and JAK-STAT compared to the other MC samples, indicating it has the most inflammatory phenotype among the samples. Furthermore, VOA8762 along with VOA8771 display the lowest relative activity of growth factor driven pathways such as EGFR, MAPK, VEGF and PI3K, which is more expected of MC **(Fig. 14)**.

5.3 Evaluating variation among clear cell carcinoma models

I performed a similar gene set scoring strategy to clear cell carcinoma models.

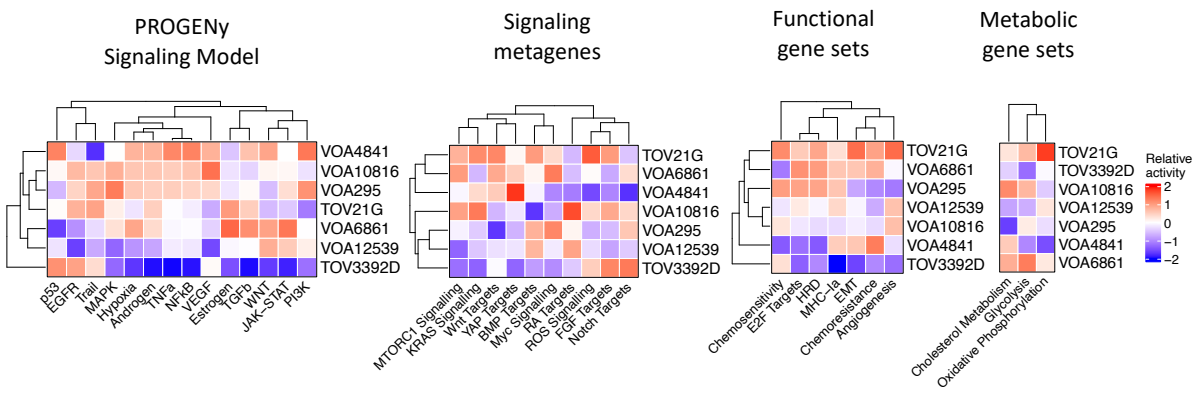

Figure 15. Heatmap of the relative activity of various gene sets across CCC ovarian cancer samples. The gene sets are organized into groups representing the PROGENy signaling model, signaling metagenes, functional gene sets, and metabolic gene sets, highlighting distinct pathway activity patterns across samples.

VOA10816 and VOA295 share very similar expression patterns, with high relative activity of RTKs such as EGFR, MAPK and VEGF, indicating these lines have more growth-factor signalling compared to the other CCC lines **(Fig. 15)**. Increased VEGF signalling has been identified as a hallmark of CCC, and reflects their adaptation to hypoxia, which I also see elevated relative activity in the majority of our cell lines[215]. Simultaneously, these lines, in addition to VOA4841, exhibited high relative activity of stress-related and inflammatory signalling through TNF-α and NF-κB. This is characteristic of CCC tumours as they are known to be pro-inflammatory. This is because they frequently arise from inflamed endometriosis lesions, which undergo recurrent cyclic hemorrhage that results in the formation of reactive oxygen species (ROS). These ROS activate and induce inflammatory transcription factors like NF-κB, and inflammatory cytokines such as TNF-α and IL-6, as well as other stress response

genes.[215–217] CCC tumours are also sensitive to immunotherapy treatments[218,219]. I also observed consistent patterns in hormone signalling across CCC samples. VOA4841, VOA10816 and VOA295 all exhibited high relative activity of androgen, and low relative activity of estrogen. TOV21G and VOA6861 had elevated relative activity of androgen, and elevated activity of estrogen simultaneously. However, VOA12539 and TOV3392D have low expression of both these hormone signalling pathways. TOV3392D also has consistently low relative activity of the majority of the cancer-relevant signalling pathways used in our analyses, however it had the highest relative activity of Notch targets compared to all the other CCC samples **(Fig. 15)**. Notch signalling has highly context-dependent roles in cancer and can act as both an oncogene and regulator of cancer cell fate. Notch activation has previously been identified to support the maintenance of stem-like or progenitor cell states and to regulate differentiation programs, contributing to long-term cell survival[220]. Considering that TOV3392D has low relative activity of canonical growth signals like PI3K, MAPK or Hypoxia, the elevation of Notch targets activity could suggest a stem-like transcriptional state rather than a highly proliferative one.[221]

In summary, models within each subtype share a common molecular structure, yet are not identical and exhibit varying degrees of biological and technical heterogeneity. The variability should be considered when choosing certain models for different experimental aims.

Chapter 6: Discussion

Ovarian cancer is a highly heterogeneous disease comprising various histological subtypes, each with distinct molecular drivers, biomarkers, prognoses and therapeutic vulnerabilities. Despite this diversity, there is a lack of molecular characterization across all subtypes due to limited availability of high-quality preclinical models and uneven representation across all subtypes. These gaps further interfere with the development of successful subtype-specific therapeutic strategies. I sought to address this challenge by performing integrated transcriptomic, proteomic and genomic profiling of 31 human ovarian cancer cell models to systematically characterize these models and develop a resource for the ovarian cancer community as part of the OvCAN Initiative. To achieve this, I first collected and preprocessed all the multi-omic data, I then identified molecular and phenotypic features that distinguished ovarian cancer subtypes, and finally I evaluated the similarity of models for specific subtypes to determine the extent of biological and technical variation.

Our multi-omic analyses revealed that ovarian cancer subtypes are characterized by distinct molecular expression profiles at both the RNA and protein levels. Dimensionality reduction analyses such as PCA and t-SNE, along with hierarchical clustering methods generally demonstrated subtype-specific clustering across the models. This aligns with known biological distinctions among subtypes, including distinct cells of origin, driver mutations and molecular phenotypes that define each ovarian cancer subtype.

Differential expression and molecular signature analyses identified subtype-specific genes and biological processes consistent with known biological features of each subtype. They also identified novel processes and biomarkers that represent valuable resources for future biomarker discovery, experimental endeavours and therapeutic strategies, including ADCs.

6.1 Relationship between Endometrioid and SCCOHT

One notable observation from this study was the modest overlap in Endometrioid and SCCOHT models. Across numerous transcriptomic and proteomic analyses, EC and SCCOHT samples clustered closely together and reflected overlapping molecular signatures. First, I observed this overlap in our PCAs where in both transcriptomic and proteomic profiles, the EC and the SCCOHT samples clustered closely together in principal component space. I also saw an overlap between EC and SCCOHT in the molecular signatures, where samples of each subtype flagged for each other's signatures. Furthermore, from our mutation analysis using WES, the Endometrioid line TOV112D displayed a SMARCA4 mutation, which is typically found in 95% of SCCOHT cases[222]. Although EC and SCCOHT are distinct histological subtypes, they shared features relating to the chromatin-remodeling dysregulation, which may provide a mechanistic explanation for this similarity. EC frequently harbours alterations, specifically deletions, in ARID1A, which is a key subunit of the SWI/SNF chromatin remodeling complex[223]. And SCCOHT frequently harbours inactivating mutations of SMARCA4, which is another essential SWI/SNF ATPase subunit. This results in the disruption of the SWI/SNF pathway.[224] Both ARID1A loss in EC tumours and SMARCA4 mutations in SCCOHT can impair normal chromatin remodeling and transcriptional regulation[225]. Together, these shared alterations in the SWI/SNF complex could be the mechanistic explanation for the convergence of expression profiles between EC and SCCOHT in our analyses, however, given the small sample sizes (n=2 per subtype), this finding should be considered hypothesis-generating.

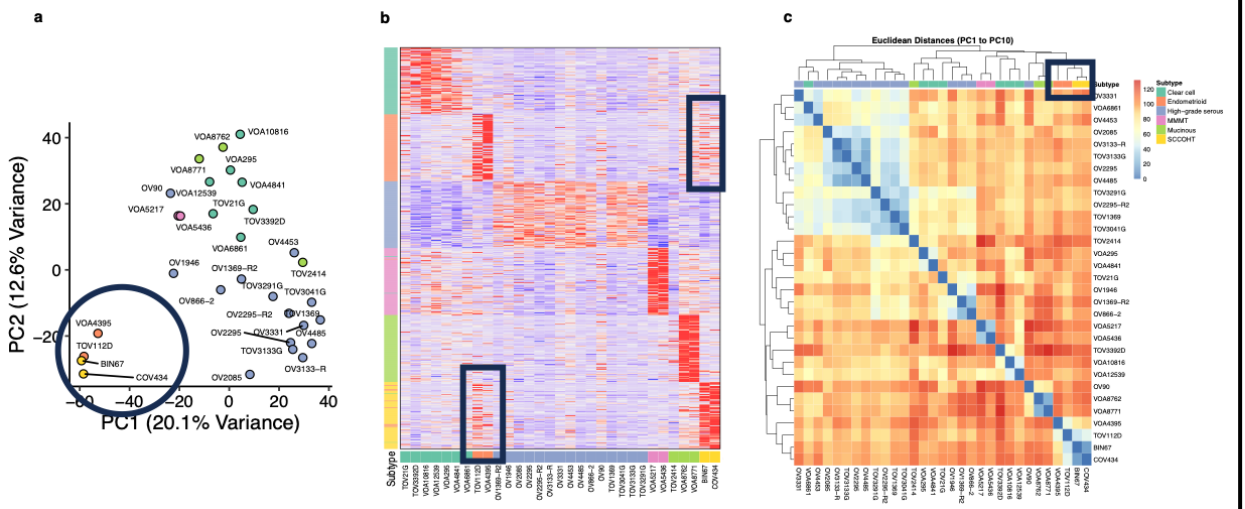

Figure 16. Integrated transcriptomic analyses reveal partial molecular convergence between EC and SCCOHT models. a) Principal component analysis of RNA expression profiles showing EC and SCCOHT samples clustering closely along the first two principal components. **b)** Molecular signatures overlap demonstrating reciprocal enrichment of EC and SCCOHT subtype signatures. **c)** Hierarchical clustering based on Euclidean distances across the top principal components, further highlighting similarity between EC and SCCOHT models relative to other subtypes.

6.2 Intra-subtype variability

Despite the subtype-specific features we observed, intra-subtype variability was still present, where certain models of specific subtypes deviated from their assigned subtype. This was observed with the HGSC model OV90, and the MC model TOV2414, which both deviated from the molecular signatures of their annotated subtypes, and consistently clustered with other subtypes in various analyses. Although OV90 harboured a hallmark TP53 mutation, it also lacked hallmark genomic instability and instead harbored an activating PIK3CA mutation,

suggesting altered pathway signaling relative to typical HGSC. Similarly, the MC model TOV2414 clustered with HGSC and CCC lines and exhibited molecular features associated with other subtypes, including ARID1A, commonly observed in CCC. Several factors may contribute to these deviating patterns, including genuine biological heterogeneity, atypical clinical representation, misclassification of the lines, or molecular evolution *in vitro* leading to divergence from the original tumour state. Technical factors such as contamination could also be responsible for these outliers. Overall, these observations highlight that individual models may not faithfully recapitulate their presumed histological subtype, which underscores the importance of systematic molecular characterization when selecting experimental systems.

I further characterized intra-subtype variability by examining relative activity scores of relevant cancer biology pathways using PROGENy, allowing the assessment of signalling heterogeneity across specific ovarian cancer subtypes. This is useful for the resource because it captures functional signalling differences between models that are not apparent from subtype analyses alone, overall supporting more informed model selection, pathway targeted studies, and overall strengthening knowledge on ovarian cancer cell models.

Application of the ConsensusOV framework enabled classification of HGSC models into TCGA-defined molecular subtypes. This provides a structured approach to group HGSC cell lines based on shared gene expression patterns. In several cases, cell lines showed expression profiles consistent with their specific HGSC subtypes, supporting the use of this framework to guide model selection for HGSC-focused studies. However, TCGA subtypes were originally defined using bulk tumour tissue containing stromal and immune cells. Since 2D cell lines lack these components, subtype assignments reflect gene expression within the cancer cells rather

than the full tumour context, therefore these classifications should be viewed as approximations rather than direct equivalents of primary tumours.

This work complements large-scale cancer cell lines resources such as CCLE and Pan-Cancer Initiatives, which have limited representation of ovarian cancer subtypes, especially the rarer ones, in comparison to other cancers. This study extends existing datasets through the integration of proteomic and whole-exome profiling and focused characterization of specific ovarian cancer biology. Furthermore, the inclusion of the rarer subtypes addresses key gaps in existing resources and enhances the relevance of this data for the ovarian cancer research community.

Overall, this study partially supports the hypothesis that models of the same histological subtype would exhibit greater molecular similarity. While we observed subtype specific clustering and grouping, we also observed intra-subtype heterogeneity where some models of different histological subtypes were exhibiting greater molecular similarity. These findings indicate that histological classification alone does not fully capture molecular diversity and reinforces the importance of systematic molecular profiling for model selection.

6.3 Limitations

Despite the strengths of this study, a few limitations should be addressed. First, while cell models offer reproducible controlled systems for experimental investigations, they do not always fully recapitulate the complexity of human tumours and the tumour microenvironment. Therefore, future studies should focus on the utilization of 3D models, such as organoids, to obtain more physiologically relevant information on subtype-specific biology. Second, whole-exome sequencing was performed without matched normal samples, limiting the ability to distinguish

somatic mutations from germline variations, and potentially inflating observed mutation frequencies. Third, the limited availability of models for rarer subtypes, such as the two models for EC, SCCOHT and MMMT, constrained the study and resulted in unequal sample sizes across subtypes, which may have influenced the statistical power and robustness of subtype-level comparisons. This also includes the unfortunate elimination of LGSC samples from our data set, which hindered our ability to provide more information on this rare subtype to the ovarian cancer community. Ultimately, the issue of limited sample size reflects the inherent rarity and underrepresentation of ovarian cancer subtypes, rather than shortcomings of the study design. Finally, the differential expression analyses were conducted using a one-versus-all framework, where each subtype was compared against a heterogeneous "other" group combining all other subtypes. This "other" group was dominated by the HGSC subtype since it had the most models (n=15), which may bias differential expression results and reduce sensitivity.

6.4 Future Directions

A major strength of this study is the integration of transcriptomic, proteomic and genomic data to provide a multi-layered view of ovarian cancer subtype biology. By profiling models across three molecular modalities, we were able to capture profound aspects of tumour biology- including gene expression, protein abundance and genomic alterations, enabling more robust characterization of subtype specific features than a single data type alone. Future work applying formal multi-omic integration frameworks may further refine these relationships. These frameworks include Multi-Omics Factor Analysis and Multiple Co-Inertia Analysis, which join transcriptomic, proteomic and genomic data to identify shared patterns of variation across datasets[226,227].

Another future direction involves comparing our cell line profiles with primary tumour data from TCGA to help identify which models most closely resemble patient tumours. This would support better model selection and increase translational relevance. However, it is important to note that TCGA cohort only comprises HGSC samples. Further studies prioritizing comprehensive molecular profiling of rarer ovarian cancer subtypes would provide valuable reference datasets.

Furthermore, it would be ideal to complete experimental validation of candidate subtype-specific markers and pathways identified in this study to establish functional significance. This includes validation of differentially expressed genes, pathway activity signatures and therapeutic targets, including ADCs. Immunohistochemical validation of signature genes in patient tumour samples would also provide important spatial and histopathological context, helping confirm subtype-specific expression patterns. Functional assays assessing pathway dependency and therapeutic sensitivity across models would also further support clinical applicability.

As stated earlier, future expansion of this resource to include additional models, particularly for underrepresented subtypes, would further enhance its translational relevance. Incorporating three-dimensional models, such as organoids, would also provide a more physiologically relevant approach and would further strengthen our biological conclusions. In parallel, deeper investigation on the molecular convergence between Endometrioid and SCCOHT models would clarify the mechanistic role of SWI/SNF chromatin remodelling and determine whether the shared alterations of this pathway drive their molecular profile similarity.

Overall, this study represents a valuable multi-omic resource for the ovarian cancer research community. By systematically characterizing diverse ovarian cancer cell models across transcriptomic, proteomic and genomic layers, this work provides a foundation for rational

model selection, and the future discovery of subtype and model-specific therapeutic strategies. Collectively, these efforts support the advancement of ovarian cancer research to improve the outcomes of patients across Canada, and the rest of the world.

References

1. Matulonis, U. A. *et al.* Ovarian cancer. *Nat. Rev. Dis. Primer* **2**, 16061 (2016).
2. Cabasag, C. J. *et al.* Ovarian cancer today and tomorrow: A global assessment by world region and Human Development Index using GLOBOCAN 2020. *Int. J. Cancer* **151**, 1535–1541 (2022).
3. Kurman, R. J. & Shih, I.-M. The Origin and Pathogenesis of Epithelial Ovarian Cancer- a Proposed Unifying Theory. *Am. J. Surg. Pathol.* **34**, 433–443 (2010).
4. Arora, T., Mullangi, S., Vadakekut, E. S. & Lekkala, M. R. Epithelial Ovarian Cancer. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2025).
5. De Leo, A. *et al.* What Is New on Ovarian Carcinoma: Integrated Morphologic and Molecular Analysis Following the New 2020 World Health Organization Classification of Female Genital Tumors. *Diagnostics* **11**, 697 (2021).
6. Cheung, A. *et al.* Non-Epithelial Ovarian Cancers: How Much Do We Really Know? *Int. J. Environ. Res. Public. Health* **19**, 1106 (2022).
7. Kurman, R. J. Origin and molecular pathogenesis of ovarian high-grade serous carcinoma. *Ann. Oncol.* **24**, x16–x21 (2013).
8. Filippova, O. T. *et al.* Molecular Characterization of High-Grade Serous Ovarian Cancers Occurring in Younger and Older Women. *Gynecol. Oncol.* **161**, 545–552 (2021).
9. Widjaja, A. O. *et al.* Abstract 7287: High-grade serous ovarian carcinoma detected with TP53 mutation panel and SiMSen-Seq. *Cancer Res.* **84**, 7287 (2024).
10. Bouberhan, S., Philp, L., Hill, S., Al-Alem, L. F. & Rueda, B. Exploiting the Prevalence of Homologous Recombination Deficiencies in High-Grade Serous Ovarian Cancer. *Cancers* **12**, 1206 (2020).

11. Integrated Genomic Analyses of Ovarian Carcinoma. *Nature* **474**, 609–615 (2011).

12. Turashvili, G. *et al.* Tumor BRCA Testing in High Grade Serous Carcinoma: Mutation Rates and Optimal Tissue Requirements. *Cancers* **12**, 3468 (2020).

13. Chui, M. H., Momeni Boroujeni, A., Mandelker, D., Ladanyi, M. & Soslow, R. A. Characterization of TP53-wildtype tubo-ovarian high-grade serous carcinomas: rare exceptions to the binary classification of ovarian serous carcinoma. *Mod. Pathol.* **34**, 490–501 (2021).

14. Kim, J. *et al.* Cell Origins of High-Grade Serous Ovarian Cancer. *Cancers* **10**, 433 (2018).

15. Ducie, J. *et al.* Molecular analysis of high-grade serous ovarian carcinoma with and without associated serous tubal intra-epithelial carcinoma. *Nat. Commun.* **8**, 990 (2017).

16. Liberto, J. M. *et al.* Current and Emerging Methods for Ovarian Cancer Screening and Diagnostics: A Comprehensive Review. *Cancers* **14**, 2885 (2022).

17. Mahmood, R. D., Morgan, R. D., Edmondson, R. J., Clamp, A. R. & Jayson, G. C. First-Line Management of Advanced High-Grade Serous Ovarian Cancer. *Curr. Oncol. Rep.* **22**, 64 (2020).

18. Dao, F. *et al.* Characteristics of 10-year survivors of high-grade serous ovarian carcinoma. *Gynecol. Oncol.* **141**, 260–263 (2016).

19. Luyckx, M., Squifflet, J.-L., Bruger, A. M. & Baurain, J.-F. Recurrent High Grade Serous Ovarian Cancer Management. in *Ovarian Cancer* (ed. Lele, S.) (Exon Publications, Brisbane (AU), 2022).

20. Babaier, A., Mal, H., Alselwi, W. & Ghatage, P. Low-Grade Serous Carcinoma of the Ovary: The Current Status. *Diagnostics* **12**, 458 (2022).

21. Chui, M. H. *et al.* Clinicopathologic Features, Molecular Landscape, and Prognostic Implications of Ovarian Low-grade Serous Tumors with Histologic Transformation. *Clin. Cancer Res.* **31**, 3084–3095 (2025).

22. Zwimpfer, T. A. *et al.* Low grade serous ovarian cancer – A rare disease with increasing therapeutic options. *Cancer Treat. Rev.* **112**, 102497 (2023).

23. Gershenson, D. M. *et al.* The genomic landscape of low-grade serous ovarian/peritoneal carcinoma and its impact on clinical outcomes. *Gynecol. Oncol.* **165**, 560–567 (2022).

24. Nakayama, K., Razia, S., Ishibashi, T. & Kyo, S. Current concept of low-grade serous ovarian carcinoma. *Transl. Cancer Res.* **13**, 6–10 (2024).

25. Clear Cell Ovarian Cancer: Speed of Spread and Outlook. *Healthline* https://www.healthline.com/health/ovarian-cancer/clear-cell-ovarian-cancer (2023).

26. Sun, M. & Jiang, W. Ovarian clear cell carcinoma with or without endometriosis origin in a single institution cohort. *Discov. Oncol.* **14**, 39 (2023).

27. Clear Cell Carcinoma - an overview | ScienceDirect Topics. https://www.sciencedirect.com/topics/pharmacology-toxicology-and-pharmaceutical-science/clear-cell-carcinoma#.

28. Lyu, C., Zhang, Y., Zhou, X. & Lang, J. ARID1A gene silencing reduces the sensitivity of ovarian clear cell carcinoma to cisplatin. *Exp. Ther. Med.* **12**, 4067–4071 (2016).

29. Wong, O. G. W., Li, J. & Cheung, A. N. Y. Targeting DNA Damage Response Pathway in Ovarian Clear Cell Carcinoma. *Front. Oncol.* **11**, 666815 (2021).

30. Gadducci, A. *et al.* Clear cell carcinoma of the ovary: Epidemiology, pathological and biological features, treatment options and clinical outcomes. *Gynecol. Oncol.* **162**, 741–750 (2021).

31. Ayhan, A. *et al.* Loss of ARID1A Expression Is an Early Molecular Event in Tumor Progression From Ovarian Endometriotic Cyst to Clear Cell and Endometrioid Carcinoma. *Int. J. Gynecol. Cancer* **22**, 1310–1315 (2012).

32. Clear Cell Ovarian Cancer - Ovarian Cancer Foundation NZ. https://ovariancancerfoundation.org.nz/clear-cell-ovarian-cancer/ (2023).

33. Stewart, J., Cunningham, N. & Banerjee, S. New therapies for clear cell ovarian carcinoma. *Int. J. Gynecol. Cancer* **33**, 385–393 (2023).

34. Wang, Y., Liu, L. & Yu, Y. Mucins and mucinous ovarian carcinoma: Development, differential diagnosis, and treatment. *Heliyon* **9**, e19221 (2023).

35. Mucinous carcinoma. https://www.pathologyoutlines.com/topic/ovarytumormucinouscarcinoma.html.

36. Goodman, A. Mucinous Ovarian Malignancy in a Young Woman: A Case Report - Clinical Case Reports Journal (ISSN 2767-0007). *Clin. Case Rep. J.* https://clinicalcasereportsjournal.com/article/1000117/mucinous-ovarian-malignancy-in-a-young-woman-a-case-report# (2021).

37. Perren, T. J. Mucinous epithelial ovarian carcinoma. *Ann. Oncol.* **27**, i53–i57 (2016).

38. Jang, M. *et al.* Identifying and capitalizing on unique molecular alterations of mucinous ovarian carcinoma for the development of novel therapeutic strategies. *Gynecol. Oncol. Rep.* **63**, 102008 (2026).

39. Ryland, G. L. *et al.* Mutational landscape of mucinous ovarian carcinoma and its neoplastic precursors. *Genome Med.* **7**, 87 (2015).

40. Babaier, A. & Ghatage, P. Mucinous Cancer of the Ovary: Overview and Current Status. *Diagnostics* **10**, 52 (2020).

41. Lertkhachonsuk, A. *et al.* Serum CA19-9, CA-125 and CEA as tumor markers for mucinous ovarian tumors. *J. Obstet. Gynaecol. Res.* **46**, 2287–2291 (2020).

42. Cheasley, D. *et al.* The molecular origin and taxonomy of mucinous ovarian carcinoma. *Nat. Commun.* **10**, 3935 (2019).

43. Wang, Y., Peng, L., Ye, W. & Lu, Y. Multimodal diagnostic strategies and precision medicine in mucinous ovarian carcinoma: a comprehensive approach. *Front. Oncol.* **14**, 1391910 (2024).

44. Vang, R. *et al.* Ovarian mucinous tumors associated with mature cystic teratomas: morphologic and immunohistochemical analysis identifies a subset of potential teratomatous origin that shares features of lower gastrointestinal tract mucinous tumors more commonly encountered as secondary tumors in the ovary. *Am. J. Surg. Pathol.* **31**, 854–869 (2007).

45. Chui, M. H. & Ellenson, L. H. Mixed ovarian neoplasms with gastrointestinal-type mucinous and Mullerian epithelial components: a rare group of tumors demonstrating the phenotypic plasticity of the Mullerian epithelial cell. *Am. J. Surg. Pathol.* **47**, 756–765 (2023).

46. de Nonneville, A. *et al.* Endometrioid ovarian carcinoma landscape: pathological and molecular characterization. *Mol. Oncol.* **18**, 2586–2600 (2024).

47. Parra-Herran, C. *et al.* Molecular-based classification algorithm for endometrial carcinoma categorizes ovarian endometrioid carcinoma into prognostically significant groups. *Mod. Pathol.* **30**, 1748–1759 (2017).

48. Sohn, J., Lee, Y. & Kim, H.-S. Endometrioid Carcinomas of the Ovaries and Endometrium Involving Endocervical Polyps: Comprehensive Clinicopathological Analyses. *Diagnostics* **12**, 2339 (2022).

49. Updates on Rare Epithelial Ovarian Carcinoma. in *Translational Advances in Gynecologic Cancers* 181–195 (Academic Press, 2017). doi:10.1016/B978-0-12-803741-6.00010-0.

50. Hollis, R. L. *et al.* Molecular stratification of endometrioid ovarian carcinoma predicts clinical outcome. *Nat. Commun.* **11**, 4995 (2020).

51. Pavone, M. E. & Lyttle, B. M. Endometriosis and ovarian cancer: links, risks, and challenges faced. *Int. J. Womens Health* **7**, 663–672 (2015).

52. Barreta, A. *et al.* Endometriosis-Associated Ovarian Cancer: Population Characteristics and Prognosis. *Int. J. Gynecol. Cancer Off. J. Int. Gynecol. Cancer Soc.* **28**, 1251–1257 (2018).

53. Bulun, S. E., Wan, Y. & Matei, D. Epithelial Mutations in Endometriosis: Link to Ovarian Cancer. *Endocrinology* **160**, 626–638 (2019).

54. Peres, L. C. *et al.* Invasive Epithelial Ovarian Cancer Survival by Histotype and Disease Stage. *J. Natl. Cancer Inst.* **111**, 60–68 (2019).

55. Kolasa, I. K. *et al.* PIK3CA amplification associates with resistance to chemotherapy in ovarian cancer patients. *Cancer Biol. Ther.* **8**, 21–26 (2009).

56. (PDF) A Review of the Clinical Characteristics and Novel Molecular Subtypes of Endometrioid Ovarian Cancer. *ResearchGate* https://www.researchgate.net/publication/352753290_A_Review_of_the_Clinical_Characteristics_and_Novel_Molecular_Subtypes_of_Endometrioid_Ovarian_Cancer (2026).

57. SCOCF | About Small Cell Ovarian Cancer. https://smallcellovarian.org/about-small-cell-ovarian-cancer.html#.

58. Wens, F. S. P. L. *et al.* Small Cell Carcinoma of the Ovary, Hypercalcemic Type (SCCOHT): Patient Characteristics, Treatment, and Outcome—A Systematic Review. *Cancers* **15**, 3794 (2023).

59. Tischkowitz, M. *et al.* Small-Cell Carcinoma of the Ovary, Hypercalcemic Type–Genetics, New Treatment Targets, and Current Management Guidelines. *Clin. Cancer Res.* **26**, 3908–3917 (2020).

60. Almuradova, E. & Cicin, I. Cancer-related hypercalcemia and potential treatments. *Front. Endocrinol.* **14**, (2023).

61. Chen, L., Dinh, T. A. & Haque, A. Small cell carcinoma of the ovary with hypercalcemia and ectopic parathyroid hormone production. *Arch. Pathol. Lab. Med.* **129**, 531–533 (2005).

62. Matias-Guiu, X. *et al.* Human parathyroid hormone-related protein in ovarian small cell carcinoma. An immunohistochemical study. *Cancer* **73**, 1878–1881 (1994).

63. Tripepi, M., da Costa, A. G., Chi, D. S., Lima, J. & Casanova, J. Small cell carcinoma of the ovary, hypercalcemic type: a mini review. *Front. Oncol.* **15**, (2025).

64. Witkowski, L. *et al.* Germline and somatic SMARCA4 mutations characterize small cell carcinoma of the ovary, hypercalcemic type. *Nat. Genet.* **46**, 438–443 (2014).

65. Belbaraka, R., Taleb, A. & Errihani, H. A Rare Tumor of the Ovary: Carcinosarcoma. *J. Med. Cases* **1**, 55–57 (2010).

66. McCluggage, W. G. Malignant biphasic uterine tumours: carcinosarcomas or metaplastic carcinomas? *J. Clin. Pathol.* **55**, 321–325 (2002).

67. Ammann, L. & Kaseb, H. Carcinosarcoma. *Pathology Outlines* https://www.pathologyoutlines.com/topic/ovarytumormmt.html.

68. Canadian Cancer Society. Malignant mixed Müllerian tumour (MMMT). *Canadian Cancer Society* https://cancer.ca/en/cancer-information/resources/glossary/m/malignant-mixed-mullerian-tumour-mmmt.

69. Jones, S. *et al.* Genomic analyses of gynaecologic carcinosarcomas reveal frequent mutations in chromatin remodelling genes. *Nat. Commun.* **5**, 5006 (2014).

70. Jagtap, S. V., Jagtap, S. S., Gudur, R. & Billawaria, S. Primary ovarian malignant mixed Müllerian tumor: a rare case report. *Ther. Adv. Rare Dis.* **3**, 26330040221107389 (2022).

71. Lisio, M.-A., Fu, L., Goyeneche, A., Gao, Z. & Telleria, C. High-Grade Serous Ovarian Cancer: Basic Sciences, Clinical and Therapeutic Standpoints. *Int. J. Mol. Sci.* **20**, 952 (2019).

72. Gonzalez, A., Nagel, C. I. & Haight, P. J. Targeted Therapies in Low-Grade Serous Ovarian Cancers. *Curr. Treat. Options Oncol.* **25**, 854–868 (2024).

73. Goldberg, R. M. *et al.* Secondary cytoreductive surgery for recurrent low-grade serous ovarian carcinoma: A systematic review and meta-analysis. *Gynecol. Oncol.* **164**, 212–220 (2022).

74. Chien, W. *et al.* Treatment for ovarian clear cell carcinoma with combined inhibition of WEE1 and ATR. *J. Ovarian Res.* **16**, 80 (2023).

75. Rajadevan, N. *et al.* Mucinous ovarian carcinoma: A survey of practice in Australia and New Zealand. *Aust. N. Z. J. Obstet. Gynaecol.* **64**, 319–325 (2024).

76. Brown, J. & Frumovitz, M. Mucinous Tumors of the Ovary: Current Thoughts on Diagnosis and Management. *Curr. Oncol. Rep.* **16**, 389 (2014).

77. Blanc-Durand, F. *et al.* Dose-intensive regimen treatment for small-cell carcinoma of the ovary of hypercalcemic type (SCCOHT). *Gynecol. Oncol.* **159**, 129–135 (2020).

78. Kulkarni, A., Cooke, C., Fung-Kee-Fung, M., May, T. & Zigras, T. EP246/#636 The efficacy of mek inhibitors (MEKI) in the treatment of low-grade serous ovarian cancer (LGSC): a systematic review. *Int. J. Gynecol. Cancer* **32**, A149–A150 (2022).

79. Stover, E. H. *et al.* The RAS-MEK-ERK pathway in low-grade serous ovarian cancer. *Gynecol. Oncol.* **200**, 22–32 (2025).

80. Fernandez, M. L. *et al.* Markers of MEK inhibitor resistance in low-grade serous ovarian cancer: EGFR is a potential therapeutic target. *Cancer Cell Int.* **19**, 10 (2019).

81. Takahashi, K., Takenaka, M., Okamoto, A., Bowtell, D. D. L. & Kohno, T. Treatment Strategies for ARID1A-Deficient Ovarian Clear Cell Carcinoma. *Cancers* **13**, 1769 (2021).

82. Caumanns, J. J., Wisman, G. B. A., Berns, K., van der Zee, A. G. J. & de Jong, S. *ARID1A* mutant ovarian clear cell carcinoma: A clear target for synthetic lethal strategies. *Biochim. Biophys. Acta BBA - Rev. Cancer* **1870**, 176–184 (2018).

83. Hein, K. Z., Stephen, B. & Fu, S. Therapeutic Role of Synthetic Lethality in ARID1A-Deficient Malignancies. *J. Immunother. Precis. Oncol.* **7**, 41–52 (2024).

84. Bitler, B. G., Aird, K. M. & Zhang, R. Epigenetic synthetic lethality in ovarian clear cell carcinoma: EZH2 and ARID1A mutations. *Mol. Cell. Oncol.* **3**, e1032476 (2015).

85. Wang, Q. *et al.* Gastrointestinal-type chemotherapy prolongs survival in an atypical primary ovarian mucinous carcinoma: A case report. *World J. Clin. Cases* **9**, 2533–2541 (2021).

86. Kurnit, K. C. *et al.* Effects of Gastrointestinal-Type Chemotherapy in Women With Ovarian Mucinous Carcinoma. *Obstet. Gynecol.* **134**, 1253–1259 (2019).

87. Kurnit, K. C. & Frumovitz, M. Primary mucinous ovarian cancer: options for surgery and chemotherapy. *Int. J. Gynecol. Cancer* **32**, 1455–1462 (2022).

88. Sajjad, H. *et al.* Cancer models in preclinical research: A chronicle review of advancement in effective cancer research. *Anim. Models Exp. Med.* **4**, 87–103 (2021).

89. ThermoFisher Scientific. Preclinical Studies in Drug Development. *PPD* https://www.ppd.com/what-is-a-cro/preclinical-studies-in-drug-development/.

90. Golebiewska, A. & Fields, R. C. Advancing preclinical cancer models to assess clinically relevant outcomes. *BMC Cancer* **23**, 230 (2023).

91. Tong, L. *et al.* Patient-derived organoids in precision cancer medicine. *Med* **5**, 1351–1377 (2024).

92. Kersten, K., de Visser, K. E., van Miltenburg, M. H. & Jonkers, J. Genetically engineered mouse models in oncology research and cancer medicine. *EMBO Mol. Med.* **9**, 137–153 (2017).

93. Nolan, K. *et al.* Development of syngeneic murine cell lines for use in immunocompetent orthotopic lung cancer models. *Cancer Cell Int.* **20**, 417 (2020).

94. Doty, D. T. *et al.* Modeling Immune Checkpoint Inhibitor Efficacy in Syngeneic Mouse Tumors in an Ex Vivo Immuno-Oncology Dynamic Environment. *Int. J. Mol. Sci.* **21**, 6478 (2020).

95. Eruslanov, E. B., Singhal, S. & Albelda, S. M. Mouse versus human neutrophils in cancer–a major knowledge gap. *Trends Cancer* **3**, 149–160 (2017).

96. Cook, D. P. *et al.* Comparative analysis of syngeneic mouse models of high-grade serous ovarian cancer. *Commun. Biol.* **6**, 1152 (2023).

97. Weiskirchen, R. Misidentified cell lines: failures of peer review, varying journal responses to misidentification inquiries, and strategies for safeguarding biomedical research. *Res. Integr. Peer Rev.* **10**, 12 (2025).

98. Almeida, J. L. & Korch, C. T. Authentication of Human and Mouse Cell Lines by Short Tandem Repeat (STR) DNA Genotype Analysis. in *Assay Guidance Manual* (eds Markossian, S. et al.) (Eli Lilly & Company and the National Center for Advancing Translational Sciences, Bethesda (MD), 2004).

99. Zheng, Y. *et al.* Multi-omics data integration using ratio-based quantitative profiling with Quartet reference materials. *Nat. Biotechnol.* **42**, 1133–1149 (2024).

100. Kukurba, K. R. & Montgomery, S. B. RNA Sequencing and Analysis. *Cold Spring Harb. Protoc.* **2015**, 951–969 (2015).

101. Kwon, Y. W. *et al.* Application of Proteomics in Cancer: Recent Trends and Approaches for Biomarkers Discovery. *Front. Med.* **8**, 747333 (2021).

102. Jiang, Y. *et al.* Comprehensive Overview of Bottom-Up Proteomics Using Mass Spectrometry. *ACS Meas. Sci. Au* **4**, 338–417 (2024).

103. WES – Macrogen. https://www.macrogen.com.au/ngs/wes/.

104. Xu, S. *et al.* Whole-exome sequencing reveals novel genomic signatures and potential therapeutic targets during the progression of rectal neuroendocrine neoplasm. *Cell Death Dis.* **15**, 833 (2024).

105. Warr, A. *et al.* Exome Sequencing: Current and Future Perspectives. *G3 GenesGenomesGenetics* **5**, 1543–1550 (2015).

106. Pei, X. M. *et al.* Targeted Sequencing Approach and Its Clinical Applications for the Molecular Diagnosis of Human Diseases. *Cells* **12**, 493 (2023).

107. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).

108. Sinha, R., Luna, A., Schultz, N. & Sander, C. A pan-cancer survey of cell line tumor similarity by feature-weighted molecular profiles. *Cell Rep. Methods* **1**, 100039 (2021).

109. Létourneau, I. J. *et al.* Derivation and characterization of matched cell lines from primary and recurrent serous ovarian cancer. *BMC Cancer* **12**, 379 (2012).

110. Fleury, H. *et al.* Novel high-grade serous epithelial ovarian cancer cell lines that reflect the molecular diversity of both the sporadic and hereditary disease. *Genes Cancer* **6**, 378–398 (2015).

111. Ouellet, V. *et al.* Characterization of three new serous epithelial ovarian cancer cell lines. *BMC Cancer* **8**, 152 (2008).

112. Samouëlian, V. *et al.* Chemosensitivity and radiosensitivity profiles of four new human epithelial ovarian cancer cell lines exhibiting genetic alterations in BRCA2, TGFbeta-RII, KRAS2, TP53 and/or CDNK2A. *Cancer Chemother. Pharmacol.* **54**, 497–504 (2004).

113. Gamwell, L. F. *et al.* Small cell ovarian carcinoma: genomic stability and responsiveness to therapeutics. *Orphanet J. Rare Dis.* **8**, 33 (2013).

114. Hughes, C. S. *et al.* Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat. Protoc.* **14**, 68–85 (2019).

115. Lai, K. *et al.* Genomic analysis of atypical fibroxanthoma. *PloS One* **12**, e0188272 (2017).

116. Hitz, B. C. *et al.* SnoVault and encodeD: A novel object-based storage system and applications to ENCODE metadata. 044578 Preprint at https://doi.org/10.1101/044578 (2016).

117. Gammerdinger, M. P., Meeta Mistry, Radhika Khetani, Will. Experimental design considerations. *Introduction to RNA-seq using high-performance computing*

https://hbctraining.github.io/Intro-to-rnaseq-fasrc-salmon-flipped/lessons/02_experimental_planning_considerations.html (2021).

118. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).

119. Vogel, C. & Marcotte, E. M. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* **13**, 227–232 (2012).

120. Ponomarenko, E. A. *et al.* Workability of mRNA Sequencing for Predicting Protein Abundance. *Genes* **14**, 2065 (2023).

121. Cope, L., Wu, R.-C., Shih, I.-M. & Wang, T.-L. High level of chromosomal aberration in ovarian cancer genome correlates with poor clinical outcome. *Gynecol. Oncol.* **128**, 500–505 (2013).

122. McAlpine, J. N. *et al. BRCA1* and *BRCA2* mutations correlate with *TP53* abnormalities and presence of immune cell infiltrates in ovarian high-grade serous carcinoma. *Mod. Pathol.* **25**, 740–750 (2012).

123. Choi, J. Y. *et al.* Ovarian Clear Cell Carcinoma Sub-Typing by ARID1A Expression. *Yonsei Med. J.* **58**, 59–66 (2017).

124. Rousseeuw, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53–65 (1987).

125. Selecting the number of clusters with silhouette analysis on KMeans clustering. *scikit-learn* https://scikit-learn/stable/auto_examples/cluster/plot_kmeans_silhouette_analysis.html.

126. Gao, P. *et al.* Key Role of MCUR1 in Malignant Progression of Breast Cancer. *OncoTargets Ther.* **14**, 4163–4175 (2021).

127. Fan, L., Yang, H., Zhang, B. & Ding, H. MCUR1 is a prognostic biomarker for ovarian cancer patients. *Cancer Biomark.* **33**, 311–316 (2022).

128. Ren, T. *et al.* MCUR1-Mediated Mitochondrial Calcium Signaling Facilitates Cell Survival of Hepatocellular Carcinoma via Reactive Oxygen Species-Dependent P53 Degradation. *Antioxid. Redox Signal.* **28**, 1120–1136 (2018).

129. Burgess, R. J. & Zhang, Z. Histone chaperones in nucleosome assembly and human disease. *Nat. Struct. Mol. Biol.* **20**, 14–22 (2013).

130. Takaya, H., Nakai, H., Takamatsu, S., Mandai, M. & Matsumura, N. Homologous recombination deficiency status-based classification of high-grade serous ovarian carcinoma. *Sci. Rep.* **10**, 2757 (2020).

131. van Wilpe, S. *et al.* Homologous Recombination Repair Deficiency and Implications for Tumor Immunogenicity. *Cancers* **13**, 2249 (2021).

132. Chen, Q., Zhang, Y., Wang, C., Ding, H. & Chi, L. Integrated analysis of single-cell and bulk transcriptome reveals hypoxia-induced immunosuppressive microenvironment to predict immunotherapy response in high-grade serous ovarian cancer. *Front. Pharmacol.* **15**, (2024).

133. More, M. H. *et al.* A Multistep Tumor Growth Model of High-Grade Serous Ovarian Carcinoma Identifies Hypoxia-Associated Signatures. *Cells Tissues Organs* **213**, 79–95 (2022).

134. Qie, S. & Sang, N. Stanniocalcin 2 (STC2): a universal tumour biomarker and a potential therapeutical target. *J. Exp. Clin. Cancer Res. CR* **41**, 161 (2022).

135. Huang, Y. *et al.* Mutant p53 drives cancer chemotherapy resistance due to loss of function on activating transcription of PUMA. *Cell Cycle* **18**, 3442–3455 (2019).

136. Xu, J. *et al.* Metabolic enzyme PDK3 forms a positive feedback loop with transcription factor HSF1 to drive chemoresistance. *Theranostics* **9**, 2999–3013 (2019).

137. Song, I.-S. *et al.* Regulation of glucose metabolism-related genes and VEGF by HIF-1α and HIF-1β, but not HIF-2α, in gastric cancer. *Exp. Mol. Med.* **41**, 51–58 (2009).

138. Černe, K., Polajžer, S., Kobal, B., Lukanović, D. & Škof, E. Integrating Chronic Inflammation and Hypoxia: The Potential Role of HIF-1α in Tumor Behavior and Therapy Response in High Grade Ovarian Carcinoma. *Front. Immunol.* **17**, (2026).

139. Wang, Y. *et al.* Integrated analysis of MIOX gene in prognosis of clear-cell renal cell carcinoma. *Cell Death Dis.* **16**, 368 (2025).

140. Ji, J. X., Wang, Y. K., Cochrane, D. R. & Huntsman, D. G. Clear cell carcinomas of the ovary and kidney: clarity through genomics. *J. Pathol.* **244**, 550–564 (2018).

141. Han, S. *et al.* Myo-Inositol Oxygenase (MIOX): A Pivotal Regulator and Therapeutic Target in Multiple Diseases. *Curr. Issues Mol. Biol.* **47**, 745 (2025).

142. Asada, K. *et al.* Abstract A19: Comparison of genomic alteration patterns between ovarian clear cell carcinoma and renal cell carcinoma. *Clin. Cancer Res.* **19**, A19 (2014).

143. Thuya, W. L. *et al.* Insights into IL-6/JAK/STAT3 signaling in the tumor microenvironment: Implications for cancer therapy. *Cytokine Growth Factor Rev.* **85**, 26–42 (2025).

144. Kato, N., Takeda, J., Fukase, M. & Motoyama, T. Hyalinized stroma in clear cell carcinoma of the ovary: how is it formed? *Hum. Pathol.* **43**, 2041–2046 (2012).

145. Ji, J. X. *et al.* The proteome of clear cell ovarian carcinoma. *J. Pathol.* **258**, 325–338 (2022).

146. Varghese, P. M. *et al.* C4b Binding Protein Acts as an Innate Immune Effector Against Influenza A Virus. *Front. Immunol.* **11**, (2021).

147. Heurich, M., Föcking, M. & Cotter, D. Complement C4, C4A and C4a – What they do and how they differ. *Brain Behav. Immun. - Health* **39**, 100809 (2024).

148. Hok, K. D. *et al.* Functional Roles of the Complement Immune System in Cardiac Inflammation and Hypertrophy. *Int. J. Mol. Sci.* **26**, 9931 (2025).

149. Hollingsworth, M. A. & Swanson, B. J. Mucins in cancer: protection and control of the cell surface. *Nat. Rev. Cancer* **4**, 45–60 (2004).

150. Pinho, S. S. & Reis, C. A. Glycosylation in cancer: mechanisms and clinical implications. *Nat. Rev. Cancer* **15**, 540–555 (2015).

151. Zhang, Y. *et al.* A Sweet Warning: Mucin-Type O-Glycans in Cancer. *Cells* **11**, 3666 (2022).

152. Burchell, J. M., Beatson, R., Graham, R., Taylor-Papadimitriou, J. & Tajadura-Ortega, V. O-linked mucin-type glycosylation in breast cancer. *Biochem. Soc. Trans.* **46**, 779–788 (2018).

153. Chandran Manimegalai, S. *et al.* An investigative study on the impact of DLK1 and NCoR1 knockdown by siRNA transfection on endometrial cancer proliferation: unveiling notch interactions. *Med. Oncol.* **42**, 124 (2025).

154. Cui, D. *et al.* Identification of key genes and pathways in endometriosis by integrated expression profiles analysis. *PeerJ* **8**, e10171 (2020).

155. Liao, L. & Pan, Z. DLK1 as a Potential Biomarker and shows NOTCH signaling could be the potential target for Endometriosis: A Machine Learning Approach. Preprint at https://doi.org/10.21203/rs.3.rs-3990509/v1 (2024).

156. Chen, S. *et al.* A Review of the Clinical Characteristics and Novel Molecular Subtypes of Endometrioid Ovarian Cancer. *Front. Oncol.* **11**, 668151 (2021).

157. Bourgot, I., Primac, I., Louis, T., Noël, A. & Maquoi, E. Reciprocal Interplay Between Fibrillar Collagens and Collagen-Binding Integrins: Implications in Cancer Progression and Metastasis. *Front. Oncol.* **10**, (2020).

158. Huang, J. *et al.* Extracellular matrix and its therapeutic potential for cancer treatment. *Signal Transduct. Target. Ther.* **6**, 153 (2021).

159. Mariotti, L., Pollock, K. & Guettler, S. Regulation of Wnt/β-catenin signalling by tankyrase-dependent poly(ADP-ribosyl)ation and scaffolding. *Br. J. Pharmacol.* **174**, 4611–4636 (2017).

160. Deng, F. *et al.* YAP triggers the Wnt/β-catenin signalling pathway and promotes enterocyte self-renewal, regeneration and tumorigenesis after DSS-induced injury. *Cell Death Dis.* **9**, 153 (2018).

161. Tuo, H. *et al.* MiR-324-3p promotes tumor growth through targeting DACT1 and activation of Wnt/β-catenin pathway in hepatocellular carcinoma. *Oncotarget* **8**, 65687–65698 (2017).

162. DKK2 dickkopf Wnt signaling pathway inhibitor 2 [Homo sapiens (human)] - Gene - NCBI. https://www.ncbi.nlm.nih.gov/gene/27123.

163. Mitrić, A. & Castellano, I. Targeting gamma-glutamyl transpeptidase: A pleiotropic enzyme involved in glutathione metabolism and in the control of redox homeostasis. *Free Radic. Biol. Med.* **208**, 672–683 (2023).

164. Bansal, A. *et al.* Gamma-glutamyltransferase 1 promotes clear cell renal cell carcinoma initiation and progression. *Mol. Cancer Res. MCR* **17**, 1881–1892 (2019).

165. Engqvist, H. *et al.* Immunohistochemical validation of COL3A1, GPR158 and PITHD1 as prognostic biomarkers in early-stage ovarian carcinomas. *BMC Cancer* **19**, 928 (2019).

166. Moore, R. G. *et al.* HE4 (WFDC2) gene overexpression promotes ovarian tumor growth. *Sci. Rep.* **4**, 3574 (2014).

167. Hasanbegovic, L. & Sljivo, N. Determination of the Reference Values of the Tumor Marker HE4 in Female Population of Canton Sarajevo. *Mater. Socio-Medica* **30**, 15–19 (2018).

168. James, N. E., Gura, M., Woodman, M., Freiman, R. N. & Ribeiro, J. R. A bioinformatic analysis of WFDC2 (HE4) expression in high grade serous ovarian cancer reveals tumor-specific changes in metabolic and extracellular matrix gene expression. *Med. Oncol.* **39**, 71 (2022).

169. Liu, Q. *et al.* Prognostic significance of carboxypeptidase Q and its methylation in glioblastoma. *Transl. Cancer Res.* **12**, 1073–1087 (2023).

170. Im, J.-Y. *et al.* CYB5R3 functions as a tumor suppressor by inducing ER stress-mediated apoptosis in lung cancer cells via the PERK-ATF4 and IRE1α-JNK pathways. *Exp. Mol. Med.* **56**, 235–249 (2024).

171. Lund, R. R. *et al.* NADH-Cytochrome b5 Reductase 3 Promotes Colonization and Metastasis Formation and Is a Prognostic Marker of Disease-Free and Overall Survival in Estrogen Receptor-Negative Breast Cancer. *Mol. Cell. Proteomics MCP* **14**, 2988–2999 (2015).

172. Zhao, R. *et al.* PTPN1 is a prognostic biomarker related to cancer immunity and drug sensitivity: from pan-cancer analysis to validation in breast cancer. *Front. Immunol.* **14**, 1232047 (2023).

173. Chen, X. *et al.* Identification of adhesion-associated extracellular matrix component thrombospondin 3 as a prognostic signature for clear cell renal cell carcinoma. *Investig. Clin. Urol.* **63**, 107–117 (2022).

174. Otte, A. *et al.* c-Met inhibitors attenuate tumor growth of small cell hypercalcemic ovarian carcinoma (SCCOHT) populations. *Oncotarget* **6**, 31640–31658 (2015).

175. Shvartsur, A. & Bonavida, B. Trop2 and its overexpression in cancers: regulation and clinical/therapeutic implications. *Genes Cancer* **6**, 84–105 (2015).

176. Zhou, D. *et al.* A new TROP2-targeting antibody-drug conjugate shows potent antitumor efficacy in breast and lung cancers. *Npj Precis. Oncol.* **8**, 94 (2024).

177. Liu, Y. *et al.* Folate receptor alpha for cancer therapy: an antibody and antibody-drug conjugate target coming of age. *mAbs* **17**, 2470309.

178. Varaganti, P., Buddolla, V., Lakshmi, B. A. & Kim, Y.-J. Recent advances in using folate receptor 1 (FOLR1) for cancer diagnosis and treatment, with an emphasis on cancers that affect women. *Life Sci.* **326**, 121802 (2023).

179. Pan, L. *et al.* HER2/PI3K/AKT pathway in HER2-positive breast cancer: A review. *Medicine (Baltimore)* **103**, e38508 (2024).

180. Lin, C.-W., Parveen, R., Yamaguchi, H. & Hung, M.-C. Therapeutic challenges in HER2-targeted antibody therapies: trastuzumab and its ADC derivatives in breast cancer. *Am. J. Cancer Res.* **15**, 3817–3834 (2025).

181. Ko, H. C. *et al.* From tissue-specific to tissue-agnostic: HER2 overexpression and the rise of antibody-drug conjugates. *Front. Oncol.* **15**, 1565872 (2025).

182. Nurgalieva, A. K. *et al.* Sodium-dependent phosphate transporter NaPi2b as a potential predictive marker for targeted therapy of ovarian cancer. *Biochem. Biophys. Rep.* **28**, 101104 (2021).

183. Vlasenkova, R., Nurgalieva, A., Akberova, N., Bogdanov, M. & Kiyamova, R. Characterization of SLC34A2 as a Potential Prognostic Marker of Oncological Diseases. *Biomolecules* **11**, 1878 (2021).

184. Foroutan, M. *et al.* Single sample scoring of molecular phenotypes. *BMC Bioinformatics* **19**, 404 (2018).

185. Sigismund, S., Avanzato, D. & Lanzetti, L. Emerging functions of the EGFR in cancer. *Mol. Oncol.* **12**, 3–20 (2018).

186. Uribe, M. L., Marrocco, I. & Yarden, Y. EGFR in Cancer: Signaling Mechanisms, Drugs, and Acquired Resistance. *Cancers* **13**, 2748 (2021).

187. An, S. J. *et al.* Regulation of EGF-stimulated activation of the PI-3K/AKT pathway by exocyst-mediated exocytosis. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2208947119 (2022).

188. Luo, J., Manning, B. D. & Cantley, L. C. Targeting the PI3K-Akt pathway in human cancer: rationale and promise. *Cancer Cell* **4**, 257–262 (2003).

189. Dhiman, V. K., Kumari, M. & Singh, D. Chemoresistance: The hidden barrier in cancer treatment. *Cancer Pathog. Ther.* **4**, 98–109 (2026).

190. Wang, Y. *et al.* Advances in the molecular regulation mechanism of tumor dormancy and its therapeutic strategy. *Discov. Oncol.* **15**, 184 (2024).

191. Matulonis, U. A. *et al.* High Throughput Interrogation of Somatic Mutations in High Grade Serous Cancer of the Ovary. *PLOS ONE* **6**, e24433 (2011).

192. Żuk, M. *et al.* Significance of the PIK3CA mutations in the differential diagnosis of ovarian epithelial carcinoma. https://ejtcm.gumed.edu.pl/articles/102672.

193. Nikolatou, K. *et al.* PTEN deficiency exposes a requirement for an ARF GTPase module for integrin-dependent invasion in ovarian cancer. *EMBO J.* **42**, EMBJ2023113987 (2023).

194. Teng, P. *et al.* Pharmacologic inhibition of ATR and ATM offers clinically important distinctions to enhancing platinum or radiation response in ovarian, endometrial, and cervical cancer cells. *Gynecol. Oncol.* **136**, 554–561 (2015).

195. Shim, Y. J. Therapeutic Targeting of DNA Damage Response Pathways in TP53- and ATM-Mutated Tumors. *Brain Tumor Res. Treat.* **13**, 73–80 (2025).

196. Maréchal, A. & Zou, L. DNA Damage Sensing by the ATM and ATR Kinases. *Cold Spring Harb. Perspect. Biol.* **5**, a012716 (2013).

197. Anggraeni, T., Tan, M. & Winarto, H. EP221/#1446 The role of ATM ATR gene on resistance of cancer stem cell subpopulations in advanced ovarian cancer: therapy response to in vitro apoptosis and proliferation. *Int. J. Gynecol. Cancer* **33**, A167–A169 (2023).

198. Tang, Q., Wang, X., Wang, H., Zhong, L. & Zou, D. Advances in ATM, ATR, WEE1, and CHK1/2 inhibitors in the treatment of PARP inhibitor-resistant ovarian cancer. *Cancer Biol. Med.* **20**, 915–921 (2023).

199. Chen, C.-W., Buj, R., Dahl, E. S., Leon, K. E. & Aird, K. M. ATM inhibition synergizes with fenofibrate in high grade serous ovarian cancer cells. *Heliyon* **6**, e05097 (2020).

200. Andrikopoulou, A. *et al.* Germline and somatic variants in ovarian carcinoma: A next-generation sequencing (NGS) analysis. *Front. Oncol.* **12**, 1030786 (2022).

201. Weberpals, J. I. *et al.* Tumor genomic, transcriptomic, and immune profiling characterizes differential response to first-line platinum chemotherapy in high grade serous ovarian cancer. *Cancer Med.* **10**, 3045–3058 (2021).

202. Chaluvally-Raghavan, P. *et al.* Copy Number Gain of hsa-miR-569 at 3q26.2 Leads to Loss of TP53INP1 and Aggressiveness of Epithelial Cancers. *Cancer Cell* **26**, 863–879 (2014).

203. Nanjundan, M. *et al.* Amplification of MDS1/EVI1 and EVI1, located in the 3q26.2 amplicon, is associated with favorable patient prognosis in ovarian cancer. *Cancer Res.* **67**, 3074–3084 (2007).

204. Radosa, J. C. *et al.* Effect of the 3q26-coding oncogene SEC62 as a potential prognostic marker in patients with ovarian neoplasia. *Front. Physiol.* **13**, 1054508 (2023).

205. Loh, W. E. *et al.* Human chromosome 11 contains two different growth suppressor genes for embryonal rhabdomyosarcoma. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 1755–1759 (1992).

206. Bepler, G. & Garcia-Blanco, M. A. Three tumor-suppressor regions on chromosome 11p identified by high-resolution deletion mapping in human non-small-cell lung cancer. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 5513–5517 (1994).

207. Goldberg, E. K. *et al.* Localization of Multiple Melanoma Tumor–Suppressor Genes on Chromosome 11 by Use of Homozygosity Mapping-of-Deletions Analysis. *Am. J. Hum. Genet.* **67**, 417–431 (2000).

208. Gabra, H. *et al.* Chromosome 11 allele imbalance and clinicopathological correlates in ovarian tumours. *Br. J. Cancer* **72**, 367–375 (1995).

209. Ohsuga, T. *et al.* Distinct preoperative clinical features predict four histopathological subtypes of high-grade serous carcinoma of the ovary, fallopian tube, and peritoneum. *BMC Cancer* **17**, 580 (2017).

210. Davidson, N. R. *et al.* Molecular subtypes of high-grade serous ovarian cancer across racial groups and gene expression platforms. *bioRxiv* 2023.11.01.565179 (2023) doi:10.1101/2023.11.01.565179.

211. Chen, G. M. *et al.* Consensus on Molecular Subtypes of High-grade Serous Ovarian Carcinoma. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **24**, 5037–5047 (2018).

212. Khashaba, M., Fawzy, M., Abdel-Aziz, A., Eladawei, G. & Nagib, R. Subtyping of high grade serous ovarian carcinoma: histopathological and immunohistochemical approach. *J. Egypt. Natl. Cancer Inst.* **34**, 6 (2022).

213. Tessari, M. A. *et al.* Transcriptional activation of the cyclin A gene by the architectural transcription factor HMGA2. *Mol. Cell. Biol.* **23**, 9104–9116 (2003).

214. Razia, S. *et al.* Histological and Genetic Diversity in Ovarian Mucinous Carcinomas: A Pilot Study. *Curr. Oncol.* **30**, 4052–4059 (2023).

215. Mabuchi, S., Sugiyama, T. & Kimura, T. Clear cell carcinoma of the ovary: molecular insights and future therapeutic perspectives. *J. Gynecol. Oncol.* **27**, e31 (2016).

216. Obeagu, E. I. Clear Cell Ovarian Carcinoma and Its Distinct Coagulopathy Profile: Molecular Drivers and Clinical Implications. *Cancer Manag. Res.* **17**, 2459–2467 (2025).

217. Patel, R., Kalthur, G., Datta, R., Shah, S. & Dutta, R. The Nexus of Iron, Senescence, and Fibrosis in Endometriosis: A Narrative Review. *Reprod. Sci.* **32**, 3783–3806 (2025).

218. Blanc-Durand, F., Ngoi, N., Lim, D., Ray-Coquard, I. & Tan, D. S. Clearer Horizons: The latest advances in clear cell ovarian cancer treatment. *Cancer Treat. Rev.* **138**, 102977 (2025).

219. Passarelli, A. *et al.* The immunotherapy era in ovarian clear cell carcinoma: current evidence and future perspective. *Front. Immunol.* **16**, (2025).

220. Zhou, B. *et al.* Notch signaling pathway: architecture, disease, and therapeutics. *Signal Transduct. Target. Ther.* **7**, 95 (2022).

221. Wang, Z., Li, Y., Banerjee, S. & Sarkar, F. H. Emerging Role of Notch in Stem Cells and Cancer. *Cancer Lett.* **279**, 8–12 (2009).

222. Auguste, A. *et al.* Small Cell Carcinoma of the Ovary, Hypercalcemic Type (SCCOHT) beyond SMARCA4 Mutations: A Comprehensive Genomic Analysis. *Cells* **9**, 1496 (2020).

223. Wiegand, K. C. *et al.* ARID1A Mutations in Endometriosis-Associated Ovarian Carcinomas. *N. Engl. J. Med.* **363**, 1532–1543 (2010).

224. Orlando, K. A. *et al.* Re-expression of SMARCA4/BRG1 in small cell carcinoma of ovary, hypercalcemic type (SCCOHT) promotes an epithelial-like gene signature through an AP-1-dependent mechanism. *eLife* **9**, e59073 (2020).

225. Hao, F., Zhang, Y., Hou, J. & Zhao, B. Chromatin remodeling and cancer: the critical influence of the SWI/SNF complex. *Epigenetics Chromatin* **18**, 22 (2025).

226. Argelaguet, R. *et al.* Multi-Omics Factor Analysis—a framework for unsupervised integration of multi-omics data sets. *Mol. Syst. Biol.* **14**, e8124 (2018).

227. Meng, C., Kuster, B., Culhane, A. C. & Gholami, A. M. A multivariate approach to the integration of multi-omics datasets. *BMC Bioinformatics* **15**, 162 (2014).