Medical University of Bialystok

Faculty of Medicine with the Division of Dentistry and Division of Medical Education in English

Doctoral dissertation in Medical Sciences

Transcriptomic analysis of High-Grade Serous Ovarian Cancer uncovers molecular markers with potential use in therapy or diagnosis.

Poszukiwanie molekularnych markerów terapeutyczno-diagnostycznych opartych o transkryptomiczne analizy niskozróżnicowanego surowiczego nowotworu jajnika

Anna Erol

Supervisor: Prof. dr hab. Adam Jacek Krętowski

Clinical Research Centre

Medical University of Bialystok

Białystok 2022

Pamięci cioci Lili i cioci Halinki

TABLE OF CONTENTS

Acknowledgements

Pragnę złożyć podziękowania mojemu promotorowi **Profesorowi Adamowi Krętowskiemu**, za umożliwienie mi pracy w Centrum Badań Klinicznych oraz swobodny wybór kierunku badań naukowych oraz **dr Magdalenie Niemirze**, która wsparła te badania swoim doświadczeniem oraz pomogła mi w redagowaniu tej pracy. Serdeczne podziękowania składam również pracownikom Centrum Badań Klinicznych, w szczególności zespołowi **Laboratorium Genomiki i Analiz Epigenetycznych**

Cieszę się również, że mogłam współpracować z bioinformatykami z Centrum Bioinformatyki i Analizy Danych, dziękuję **dr Karolinie Chwiałkowskiej**, która przygotowała część analiz bioinformatycznych. Dziękuję również wszystkim **lekarzom** oraz **pracownikom Biobanku**, którzy byli zaangażowani w zbieranie materiału.

Bardzo dziękuję również **pacjentkom**, które wyraziły zgodę na zebranie materiału do badań. Bez nich, nie powstałaby ta praca oraz badania, które mam nadzieję pomogą w przyszłości kolejnym chorym.

Kolejne podziękowania kieruję do kierowników studiów doktoranckich oraz projektu ImPress **Profesor Ani Moniuszko-Malinowskiej oraz Profesorowi Marcina Moniuszko**. Bardzo dziękuję za wsparcie w sprawach administracyjnych **Marii Szlachcie, Urszuli Ullmann, Magdzie Bogdan** oraz **Wandzie Banach.**

Dziękuję wszystkim **wykładowcom,** z którymi mieliśmy szansę mieć zajęcia podczas studiów. I would like to thank all the **lecturers**; I was honoured to meet throughout the PhD studies.

I want to also thank my **fellow students**, especially **Sinemyiz, Gladys, Felipe and Mauro**! Thank you, a lot, for all these years.

Chciałabym również podziękować **rodzinie** oraz **przyjaciołom**, za nieustanne wsparcie oraz motywację. **Türkiye'deki aileme** çok teşekkür ederim.

Funding

This research was conducted within the project which has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 754432 and the Polish Ministry of Science and Higher Education, from the financial resources for science in 2018-2023 granted for the implementation of an international co-financed project.

Abstract

Molecular biology is one of the fastest developing technology areas. **Next Generation Sequencing** (NGS) and data science methods allow wide investigation of changes occurring on DNA and RNA level. They are revolutionising cancer research and are helping in understanding the disease allowing researchers to find novel ways of treatment and diagnosis. Ovarian cancer (OC) causes over 200 000 deaths every year. Its most common subtype is **High-Grade Serous Ovarian Cancer** (HGSOC). The disease is difficult to diagnose at the early stage, highly heterogenous, and despite a good first-line treatment response, most patients will develop drug resistance. Understanding the molecular dysregulation of the disease could help us find precise drug targets for effective and longstanding treatment. Transcriptome analysis allows for studying gene expression of protein-coding RNAs (mRNA) and noncoding RNAs (ncRNAs). **Combining the information of coding and noncoding RNAs is important in understanding the posttranscriptional mechanism that regulates cancer development.**

This work aimed to perform a comprehensive analysis of tissue-derived RNA profiles. This study included 33 primary HGSOC tumour tissues and 33 samples of ovary tissues removed during non-oncological procedures. The analysis focused on the crosstalk between competing endogenous RNAs (ceRNAs), including mRNA, long noncoding RNAs (lncRNAs) and microRNAs (miRNAs) and their interaction with transcription factors (TFs) in HGSOC patients. The key identified genes were used as a basis for the drug repurposing approach to predict drug candidates for HGSOC treatment. **PI-103 and ZSTK474, which are phosphoinositide 3-kinase (PI3K) pathway inhibitors, were identified as drug candidates.** The PI3K pathway, which is involved in developing drug resistance in OC, is closely associated with genes identified in the ceRNA network.

The research presented in this paper may contribute to the development of new treatment options for patients with HGSOC.

Streszczenie

Biologia molekularna jest jedną z najszybciej rozwijających się technologicznie dziedzin. **Sekwencjonowanie Nowej Generacji** (ang. Next Generation Sequencing, NGS) oraz nowoczesne metody analizy danych pozwalają na szerokie badanie zmian zachodzących na poziomie DNA i RNA. Metody te rewolucjonizują badania nad nowotworami, pomagają w ich zrozumieniu i znalezieniu nowych sposobów leczenia. Rak jajnika jest przyczyną ponad 200 000 zgonów każdego roku. Najczęstszym podtypem jest **niskozróżnicowany surowiczy nowotwór jajnika** (HGSOC). Choroba ta jest trudna do zdiagnozowania we wczesnym stadium, wysoce zróżnicowana i mimo dobrej odpowiedzi na leczenie pierwszej linii, u większości pacjentów pojawia się lekooporność. Zrozumienie molekularnej dysregulacji choroby może pomóc nam w znalezieniu precyzyjnych celów leczenia. Analiza transkryptomu pozwala nam badać ekspresję genów kodujących białka (mRNA) oraz tych niekodujących (ncRNA). **Połączenie informacji o kodujących i niekodujących RNA jest ważne dla zrozumienia mechanizmu regulacji ekspresji oraz progresji choroby nowotworowej.**

Głównym celem tej pracy było przeprowadzenie kompleksowej analizy profili ekspresji RNA w tkankach guza HGSOC w porównaniu do tkanek nie zmienionych nowotworowo. Badania przedstawione w tej tezie obejmują analizę 33 pierwotnych tkanek guza HGSOC i 33 próbek tkanek jajnika usuniętych podczas procedur nieonkologicznych. Analiza skupiła się na interakcjach między konkurującymi endogennymi RNA (ceRNA), w tym mRNA, długimi niekodującymi RNA (lncRNA) i mikroRNA (miRNA) oraz ich interakcji z czynnikami transkrypcyjnymi (TF) u pacjentów z HGSOC. Kluczowe zidentyfikowane geny posłużyły jako podstawa podejścia polegającego na zmianie przeznaczenia leków do przewidywania kandydatów na leki HGSOC. Zidentyfikowano związki **PI-103 i ZSTK474, które są inhibitorami szlaku kinazy 3-fosfoinozytydowej (PI3K), jako potencjalne leki.** Szlak PI3K, który bierze udział w rozwijaniu lekooporności w OC jest ściśle powiązany z genami zidentyfikowanymi w sieci ceRNA.

Badania przedstawione w tej pracy mogą przyczynić się do opracowania nowych możliwości leczenia pacjentek z HGSOC.

Abbreviations

ANGPTL5	angiopoietin-like 5
AMOs	Anti-miRNA oligonucleotides
Bcl-xL	Bcl ligand
BMI	body mass index
bp	base pairs
BP	biological processes
BRCA1	breast cancer 1
BRCA1	breast cancer 2
CDK12	cyclin-dependent kinase 12
CDK14	cyclin-dependent kinase 14
ceRNAs	competing endogenous RNAs
CLDN3	CLDN3
CMap	ConnectivityMap
CNV	copy number variations
CP	ceruloplasmin
CPM	counts per million
DDR	DNA damage response
DE	differential expression
DEGs	differentially expressed genes
DELs	differentially expressed lncRNAs
DEMs	differentially expressed miRNAs
DNA	deoxyribonucleic acid
ECM	extracellular matrix
EMT	epithelial-mesenchymal transformation
FDA	Food and Drug Administration
FDR	false discovery rate
FIGO	International Federation of Gynaecology and Obstetrics
FU	fluorescence unit
GFs	growth factors
GO	Gene Ontology
HGSOC	High-Grade Serous Ovarian Cancer

IPA	Ingenuity Pathway Analysis
IRS1	insulin receptor substrate 1
IQR	interquartile range
KEGG	Kyoto Encyclopedia of Genes and Genomes
lncRNAs	long noncoding RNAs
logFC	logarithm Fold Change
MF	molecular functions
miRNAs	microRNAs
MREs	miRNA response elements
mRNAs	protein-coding RNAs
MSig DB	Molecular Signatures Database
MUC16	mucin 16
N	number of samples
ncRNAs	non-coding RNAs
NF1	neurofibromin 1
NFkB	nuclear factor-kappa B
NGS	Next Generation Sequencing
NS	non-significant
OC	ovarian cancer
PCDH11X	protocadherin 11 X-linked
PI3K	phosphoinositide 3-kinase
PKR1	prokineticin receptor 1
PROK1	prokineticin 1
RB1	retinoblastoma 1
RIN	RNA integrity number
RNA	ribonucleic acid
RNAi	RNA interference
RNA-seq	RNA-sequencing
RTKs	receptor tyrosine kinases
SLC34A2	solute carrier family 34 member 2
SLPI	secretory leukocyte peptidase inhibitor
STAR	steroidogenic acute regulatory protein
TFs	transcription factors

TK	tyrosine kinase
TMM	trimmed mean of M values
TNF-A	tumour necrosis factor-alpha
TP53	tumour protein P53
UV response DN	genes downregulated by ultraviolet radiation
v	version
WFIKKN2	WAP, follistatin/kazal, immunoglobulin, kunitz and netrin domain containing 2

List of table and figure captions

1.1. Tables

Table 1. Patient characteristics. The age, body mass index (BMI) and International Federation of Gynaecology and Obstetrics (FIGO) stage of patients involved in the study are presented in the table. N is the number of samples for a given variable. The p-value shows the results of the nonparametric Wilcoxon rank-sum test. The interquartile range is given in brackets for age and BMI. The percentage of tumour number samples is given in brackets for FIGO stages.

Table 2. Competing endogenous RNA (ceRNA) network of High-Grade Serous Ovarian Cancer (HGSOC) characteristics. The table includes the number of unique sets of molecules and different molecules obtained separately during the construction of the ceRNA network.

Table 3. Competing endogenous RNA (ceRNA) network of High-Grade Serous Ovarian Cancer (HGSOC). Coding and noncoding RNA molecules involved in the construction of the ceRNA network of HGSOC are shown.

Table 4. Characteristics of the transcription factor (TF)–competing endogenous RNA (ceRNA) interactions High-Grade Serous Ovarian Cancer (HGSOC). The table presents the number of nodes, edges, and the average number of neighbours in the TF–ceRNA network.

Table 5. Drug repurposing analysis results are based on the High-Grade Serous Ovarian Cancer (**HGSOC**) **competing endogenous RNA (ceRNA) network.** The list of compounds identified based on the genes in the ceRNA network of HGSOC with CMap and L100CDS2 after the cut-off was considered.

1.2. Figures

Figure 1. Analysis workflow scheme. High-Grade Serous Ovarian Cancer (HGSOC) tissues and control ovarian tissues were collected, and total RNA with a fraction of microRNA (miRNA) was isolated. Total RNA and small RNA sequencing were performed, and differentially expressed genes (DEGs), differentially expressed miRNAs (DEMs) and differentially expressed lncRNAs (DELs) were identified. A competing endogenous RNA (ceRNA) network was constructed, and transcription factors (TFs) were

identified. A drug repurposing analysis was performed using L100CDS2 and a ConnectivityMap (CMap), and compounds with HGSOC treatment potential were identified. Part of the graph objects is used from biorender.com.

Figure 2. The Tape Station traces of a sample with RIN 8.4. The diagram represents the sample intensity in fluorescence unit (FU) versus the size of the RNA fragment in base pairs (bp).

Figure 3. Top 20 significant results of functional enrichment of differentially expressed genes (DEGs) in High-Grade Serous Ovarian Cancer (HGSOC) (A) Gene Ontology (GO) Biological Processes (BP) enrichment results for the HGSOC DEGs (B) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment results for the HGSOC DEGs.

Figure 4. Differential expression analysis results, competing endogenous RNA (ceRNA) network and results of functional pathway enrichment assays of the ceRNA sets in High-Grade Serous Ovarian Cancer (HGSOC). (A) Volcano plot presenting the differential expression of lncRNAs (DELs), NS, nonsignificant. (B) Volcano plot presenting the differential expression of miRNAs (DEMs). (C) Volcano plot presenting the differential expression of genes (DEGs). (D) ceRNA network of HGSOC visualised in Cytoscape. The logarithm 2-fold change of molecules is represented by the gradually changing colour of the nodes ranging from -11.95 to 11.96. All genes with downregulated expression are presented as blue nodes, and all genes with upregulated expression are presented as red nodes. (E) Gene Ontology (GO) enrichment results for the HGSOC ceRNA network. (F) Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment results for the HGSOC ceRNA network. (G) Molecular Signatures Database (MSig DB) hallmark enrichment results for the HGSOC ceRNA network.

Figure 5. PI3K-AKT pathway in High-Grade Serous Ovarian Cancer (HGSOC). The shading in red represents upregulated genes (or group of genes) and green indicates downregulated genes (or group of genes) for all differentially expressed genes. The red rectangle highlights a group of genes in the HGSOC ceRNA network. The star represents a group of genes whose collective expression is reversed when all DEGs or only ceRNA network-involved mRNAs are considered.

Figure 6. Transcription factor (TF)–competing endogenous RNA (ceRNA) network based on High-Grade Serous Ovarian Cancer (HGSOC) and pathway analysis of hub regulators. (A) A network containing key TFs identified from among differentially expressed genes (DEGs) and their targets within the HGSOC ceRNA network. (B) Reactome terms enriched with hub regulators.

Figure 7. Heatmap of the mechanism of action (MOA) and candidate compounds for High-Grade Serous Ovarian Cancer (HGSOC) repurposed drugs. The two compounds, which were identified by both CMap and L1000CD2 tools, are presented with their chemical structures.

Figure 8. Known molecular targets of PI-103 and ZSTK474 based on a QIAGEN IPA Knowledgebase[1] **analysis.** The diagrams are colour-coded based on the differential expression of targets in the HGSOC dataset. Red indicates an overexpressed gene, and green indicates a gene with downregulated expression. (A) Known molecular targets of PI-103. (B) Known molecular targets of ZSTK-474. The graph was adapted from QIAGEN IPA[1].

1. Introduction

"Cancer begins and ends with people. In the midst of scientific abstraction, it is sometimes possible to forget this one basic fact."

June Goodfield

The most important goal in research is to find a way to apply it. Basic science gives us an understanding of the complex and molecular pathways of cancer, but in the end, there are always patients hopefully waiting, that one day we will be able to help them fully recover. This is my biggest motivation.

1.1. High-Grade Serous Ovarian Cancer in the time of technological revolution in life sciences

Although Ovarian Cancer (OC) mortality rates globally have decreased or levelled in the last years, it is still the most fatal gynaecological malignancy and the fourth leading cause of cancer-related deaths among women in Europe and the United States[2–4]. OC is an umbrella term for histologically and genetically distinct subtypes[5,6]. There are also gaps in the definition of its precursor lesions, which most probably also vary between subtypes[5–7]. Epithelial OC can be divided into the high-grade serous, low-grade serous, endometrioid, clear cell and mucinous OC[7]. Each of the subtypes is varying in presence of different mutations, the outcome of the patient and the response to the treatment. Despite years of research, there is still a big gap in knowledge about underlying mechanisms causing the development, progression, and high heterogeneity of OC.

High-Grade Serous Ovarian Cancer (HGSOC) is the most common and aggressive subtype of OC[6]. Its asymptomatic nature often leads to late diagnosis, so HGSOC is generally present in the advanced stage. The treatment of the primary tumours includes cytoreductive surgery and chemotherapy[8]. For many patients, however, it becomes almost a chronic disease with the times of disease recurrence following resumed chemotherapy[8]. Unfortunately, most patients with advanced diseases are developing chemotherapy resistance. The survival prognosis for drug-resistant patients drops dramatically, and the treatment options are limited[8–10]. Therefore, novel and effective therapeutics are in urgent demand. Understanding the pathways of cancer progression, pathogenesis and development of multidrug resistance will build a foundation for developing new therapies and describing novel molecular targets for treatment.

The development of new high-throughput technologies is revolutionising biological research. Next Generation Sequencing (NGS), and modern bioinformatic methods of analysis are helping to fill these gaps. The modern era of biological research started with human genome sequencing. In 2000 Bill Clinton, Tony Blair, Francis Collins, and Craig Venter announced that the human genome draft had been prepared. It had a strong impact on life science research and medicine. It boosted development in computational biology giving researchers more tools for detangling complicated networks of molecular biology. The Human Genome Project was one of the first "big science" projects, connecting laboratories from all over the world. This has shown how powerful such consortia can be and inspired further projects, including The Cancer Genome Atlas[11]. The collection of patients' genomic, transcriptomic, epigenomic, proteomic and clinical data of oncological patients began in 2006 and continues to this day. Nowadays, when sequencing is cheaper and faster, we have stumbled upon a bottleneck, which is data analysis. The Cancer Genome Atlas data are published for researchers to use in an open-source manner. This gives data mining opportunities for groups with not enough resources to process the samples or without access to patients' samples. On the other hand, it helps in more discoveries and breakthroughs in cancer research. Yet, using the publicly available datasets has some limitations, including limited access to information about patients or planning further experiments from the patients' samples to validate the analysis results. The Medical University of Bialystok started its own biobanking[12]. The presence of the Biobank and strong cooperation with clinicians help better understand the needs of the patients and the challenges of the disease treatment.

In 2011 the genomic analysis of HGSOC by The Cancer Genome Atlas consortium was published. Their integrated analysis included expression profiles of messenger RNA (mRNA) and micro RNA (miRNA), the methylation status of promoter regions, DNA copy number variation (CNV) and the analysis of exons[13]. Soon after, Kanchi *et. al*. published another large-scale exome-wide analysis of OC[14]. The results of both studies have shown loss of function mutations in tumour protein P53 (*TP53*) tumour suppressor gene in almost all (96%) patients[13]. Therefore, the *TP53* mutations are seen as an early event in the HGSOC evolution[15]. The proper functioning of that protein is crucial for many key cellular mechanisms, such as DNA repair, cell cycle regulation, apoptosis and senescence, or regulation of metabolism. The models of development of Invasive Serous Carcinoma, including HGSOC, always begin with a layer of 'P53 Signature', which is

associated with an increase of nuclear TP53 staining[16]. Disruption of *TP53* impairs DNA damage repair, which leads to serious consequences. The cell can further divide, even though mistakes in DNA sequence. This is why subsequential mutational processes are not as homogenous. The Cancer Genome Project results stated recurring somatic mutation events including neurofibromin 1 (*NF1*), BRCA1/2 DNA repair associated (*BRCA1/2*), RB transcriptional corepressor 1 (*RB1*), and cyclin-dependent kinase 12 (*CDK12*), however with lower populational prevalence than *TP53*[13–15]. The further steps in the development of HGSOC include the development of tubal interepithelial carcinoma and invasive serous carcinoma, where the cells first start to strongly proliferate and then became invasive into the tissue[16]. HGSOC has been identified as extremely chromosomally unstable due to its deficiency in the homologous recombination DNA repair pathway[13–15]. This high instability has the potential on acquiring chemoresistance and high inter- and intratumoral heterogeneity[13–15]. There were several pathways identified as altered in HGSOC, such as NOTCH signalling, RB and PI3K/RAS signalling, FOXM1 signalling, and DNA repair pathways[7].

HGSOC is highly heterogenous and there is a need to identify the ways to divide patients in a way, which will allow better assessment of their outcome or more precise treatment. The Cancer Genome Atlas using transcriptomic profiles of patients stratified the HGSOC into four different subgroups: mesenchymal, immunoreactive, differentiated and proliferative, naming it after gene content in the clusters. The subtypes showed differences in overall survival, which confirms the importance of understanding the heterogeneity within HGSOC[13]. With the further development in NGS and progress in computational power, faster processing of enormous amounts of genetic data is possible. Multidimensional studies have started, giving rise to modern multi-omics studies. Omics technologies are collective characterisations and measurements of pools of biological molecules such as DNA, RNA or proteins, which are then translated into the structure, function, and dynamics of organisms. Using the same publicly available data from The Cancer Genome Atlas, Zhang *et. al*. prepared multi-omics-based molecular subtyping of HGSOC patients using an integrated framework. It allowed them to consider not only the transcriptomic profile but also DNA methylation, protein expression, miRNA expression and pathway activity[17]. The high heterogeneity impedes understanding of pathogenesis and resistance development in HGSOC as well as treatment outcomes of the patients.

1.2. RNA crosstalk - gene expression regulation by non-coding RNAs

The analysis of DNA mutation status, RNA and protein expression changes in HGSOC has helped better understand the molecular biology of HGSOC. Sequence variations in coding regions can impact protein structures, and those in the non-coding regions may result in differentiated gene expression and splicing[18,19]. Protein and RNA expression describes the dynamic state of cells and tissues. It helps us detangle, which molecular pathways are disturbed and might be treatment targets. The altered gene expression is however a result of multi-levelled and complex regulation. This can involve promoter or enhancer activities, DNA methylation or the expression of non-coding RNA[20]. Considering a biological system by understanding all its parts will provide us with precise therapeutic targets or lead to developing sensitive diagnostic tools[21].

There are many layers of gene expression regulation. Salmena and colleagues introduced a competing endogenous RNA (ceRNA) network hypothesis, through which they describe the regulatory communication between RNA molecules[22]. This hypothesis suggests that different transcriptome components, including coding and noncoding RNAs, pseudogenes and circular RNAs, compete to bind a limited number of miRNAs. miRNAs impact mRNA stability and translation rate by pairing to its complementary sites, called miRNA response elements (MREs), which act as decoys. A single miRNA can repress hundreds of transcripts, and a single mRNA can contain MREs to which multiple miRNAs can bind[23] Additionally, it has been shown that multiple lncRNAs, especially pseudogene-derived lncRNAs, which show high sequence homology to protein-coding transcripts, also contain MREs and can compete with mRNA for miRNA binding[24,25]. Therefore, after binding to MREs on lncRNA sequences, miRNAs are not available to interact with mRNAs. There have been shown arguments for the strong involvement of miRNAs and lncRNAs in drug resistance development in OC. The studies investigating differentially expressed miRNAs in OC have selected numerous of these short, single-stranded non-coding RNA molecules as important in OC pathogenesis or drug resistance. Strongly downregulated miRNA-542-3p in OC is an example of such molecule[26]. It has been shown that the overexpression of that miRNA suppresses tumour progression, and the *in-vitro* knockdown of this molecule supports tumour development. One of the target molecules of miRNA-542-3p is cyclin-dependent kinase 14 (*CDK14*), known for its crucial role in cell cycle control[27]. This example shows that it is not enough to understand single molecule aberration. It is the network they are involved in, which is

important to put together the picture of the molecular loops driving OC. Another supporting example of the importance of gene expression regulation by miRNA has been provided by Wang *et al.*[28]. They have constructed a network of miRNAs coordinating DNA damage response (DDR). These allowed them to select miRNAs, which were used with huge success for OC outcome prediction. Long (>200nt) non-coding transcripts are called lncRNA and have been previously described to play an important role in OC pathogenesis and progression, especially in epithelial-mesenchymal transformation (EMT)[9,29]. miRNAs and lncRNAs are not alone responsible for changes in cell molecules. Therefore, constructing a network including miRNAs, lncRNAs and mRNAs will provide us with a fuller view of the cell's internal network. The ceRNA network hypothesis suggests that RNAs can regulate each other by competing for a limited number of miRNAs. This ceRNA network consists of built-in regulatory loops, allowing cis- and trans-regulatory crosstalk throughout the whole transcriptome[22].

Transcriptional Factors (TFs) are another important regulator of gene expression. These specialised proteins bind to the promoter regions of DNA and trigger mRNA expression, making them crucial for many critical life processes[30,31]. TFs and ceRNAs are strongly intertwined: some mRNAs encoding TFs are even included in the ceRNA network itself. Moreover, TFs can affect the expression of mRNAs, playing a role in altering the ceRNA network balance[31,32].

By constructing the ceRNA network, it is possible to look closer at the tight-regulated pathways and genes in the disease. Deregulation of the key pathways drives the disease. It is possible, that restoration of the ceRNA balance might lead even to disease remission.

1.3. Precise medicine and drug repurposing.

In search of precise OC treatment, we cannot rely only on static genomic data, because there the information between genotype and phenotype of the cancer is lost. Progress is being made nowadays. Studies focus on aberrations in the genome, transcriptome, and epigenome between tumour and control samples individually and in combination. There are methods available to study gene expression on a single-cell level and even spatially[33]. The data can be integrated with patients' clinical information, such

as progression-free survival[a] or overall survival[b]. This helps to understand the crucial disturbing molecular signalling and cellular internal networks in cancer development and progression. The investigation of differences between primary tumour and metastatic sites brings researchers closer to understanding tumour evolution. However, the goal is not a pure understanding of the disease, but it brings us closer to increasing the number of safe and effective cancer therapies[34]. The standard procedure for OC treatment includes cytoreductive surgery and chemotherapy. Unfortunately, patients with the disseminated disease will relapse and become drug-resistant[8]. Knowledge regarding the development of drug resistance in OC is limited. The factors influencing developing resistance are divided into three main groups: pharmacokinetic, tumour-microenvironment and cancer-cell specific. The differences in individuals' metabolism and vascularity of the tumour strongly influence drug intake[35]. A cancer cell can alter drug efflux and influx by influencing cell-membrane transport protein[35]. When a drug already reaches a cancer cell, there are numerous mechanisms used by the cancer cell to inactivate the drug, hinder the apoptosis, and survive[35]. In a healthy organism, the immune system should help by removing unhealthy cells, but the tumour is organising its tumour microenvironment. It supports the cancer cell in survival and growth[35]. To overcome the drug-resistance there is a necessity to consider as many factors as possible. There has been shown, that altered miRNA and lncRNA expression levels are coupled with relapse and developing drug resistance in OC[9,29]. miRNA and lncRNA as regulatory RNAs are potentially targeting many of the components causing the development of therapy resistance.

Connecting selected genetic features with drugs is not an easy task. More and more databases are being built to help clinicians and researchers. For example, there is a database of genomic biomarkers for cancer drugs and clinical targetability[36]. It includes clinically "actionable"[c] and clinically "targetable"[d] features. Drug repurposing methods allow new applications for previously developed and/or approved drugs. This strategy is cost- and time-effective because the expensive steps of drug development and the first

[a] It is the time interval from the start of treatment to disease progression used as a measure of the clinical benefit of therapy

[b] It is the time interval from the start of treatment to death of the patient. It is a measure of clinical benefit. It is affected by all the treatments patient is given.

[c] That means, there are genomic features with known predictive, prognostic and or diagnostic association

[d] That means, there are genomic features linked to approved drugs, inclusive off-label use of targeted drugs, or therapeutics investigated in clinical trials

clinical phases are unnecessary[37,38]. With increasing computational power, new tools and databases can be constructed for developing a systematic approach to drug repurposing. In this study, we applied Connectivity Map (CMap) and L100CDS2 web tools to identify drug candidates for HGSOC[39]. They are based on a database of perturbation profiles by various molecules. This way we can accelerate the discovery of new use for old therapeutics[39–41].

2. Goals of the doctoral thesis project

The main goal of the doctoral thesis is to investigate the complex **multilevel network of transcriptional regulation consisting of the ceRNA network (mRNA, miRNA and lncRNA) and ceRNA-related TFs, in HGSOC**. This comprehensive analysis is based on tissue-derived RNA profiles from 66 patients, including 33 samples derived from HGSOC primary tumours from ovary tissues and 33 control samples of ovarian tissues without oncological changes confirmed by histopathology examination.

The main goal has been divided into detailed aims:

1. **Investigation of differently expressed genes, lncRNAs and miRNAs** between HGSOC primary tumour versus control samples.
2. **Identification of the key disrupted genes**, by restriction of the significant result and choice of highest up- and downregulated genes, which might be in the future used as a molecular target.
3. **Identification of key disturbed molecular mechanisms and pathways in HGSOC and its regulation**, by the construction of ceRNA network, based on DEGs, DEMs and DELs and identification of key involved TFs. The
4. **Search for potential therapeutics** using bioinformatics approaches and literature evidence.

3. Materials and Methods

A comprehensive analysis of tissue-derived RNA profiles, including mRNAs, lncRNAs and miRNAs, has been performed. The graphical representation of the analysis workflow is presented in Figure 1.

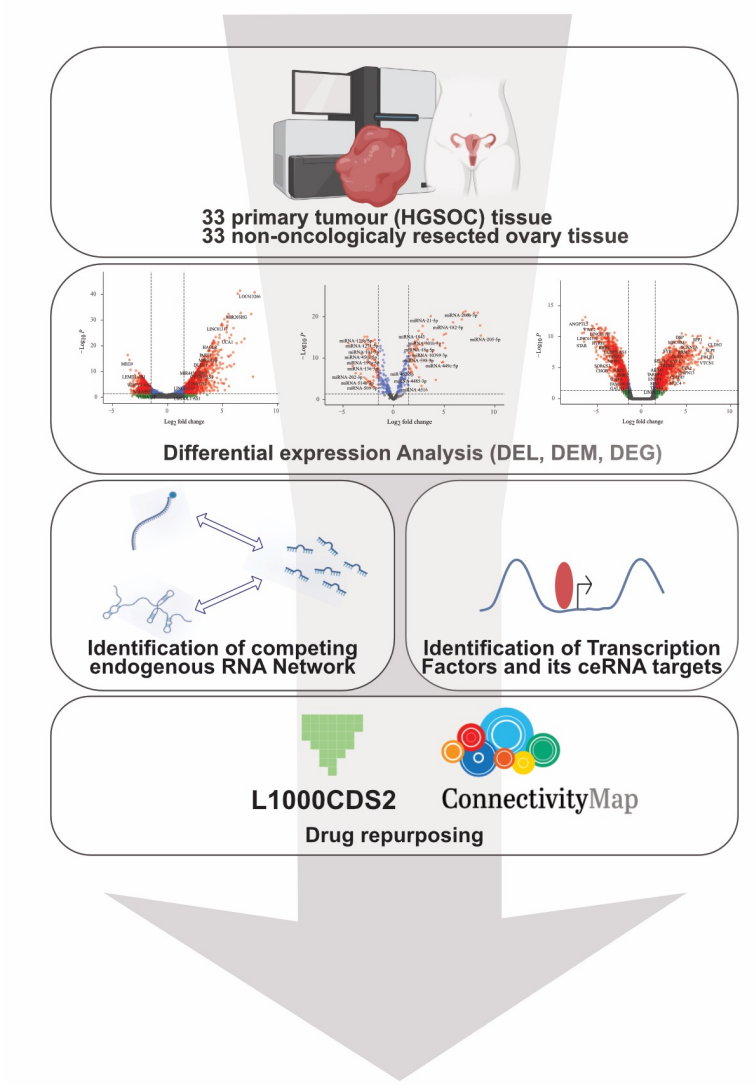

Figure 1. Analysis workflow scheme. High-Grade Serous Ovarian Cancer (HGSOC) tissues and control ovarian tissues were collected, and total RNA with a fraction of microRNA (miRNA) was isolated. Total RNA and small RNA sequencing were performed, and differentially expressed genes (DEGs), differentially expressed miRNAs (DEMs) and differentially expressed lncRNAs (DELs) were identified. A competing endogenous RNA (ceRNA) network was constructed, and transcription factors (TFs) were identified. A drug repurposing analysis was performed using L100CDS2 and a ConnectivityMap (Cmap), and compounds with HGSOC treatment potential were identified. Part of the graph objects is used from biorender.com.

3.1. Sample size estimation

The sample size estimation has been calculated considering similar previous experiments and pilot data. The minimal number of samples per experimental group

(tumour or normal) has been calculated to detect two-fold differences in the relative expression level between groups at the true positive detection powers of 80% and 90%[42]. The RNASeqPower R package[43] was used to apply statistical data covering the obtained true depths and coefficients of variation per group separately because tumour and normal tissue differ in terms of interindividual variations, as well as mRNA, lncRNA and miRNA median lengths, according to the relevant literature and databases[44,45]. Each calculation was performed separately for mRNAs, lncRNAs and miRNAs.

3.2. Sample collection

The Bioethics Committee of the Medical University of Bialystok approved the collection of the samples and the performance of the study was granted (Nr/. R-I-002/492/2019, R-I-002/606/2018). The tissue samples obtained from cytoreductive operations performed at the Clinical Hospital in Bialystok were collected by the Biobank of the Medical University of Bialystok, following the high standards of strict biobanking procedures described further by Nikliński and colleagues[12]. Ovarian tissues removed during non-oncological procedures were used as the control samples for the experiments. Collection of the tissue samples included histopathology examination to confirm pathology in case of HGSOC samples or confirm lack of pathological findings in case of control samples. Written informed consent was obtained from all the patients. Table 1 presents the characteristics of the patients enrolled in the study.

Table 1. Patient characteristics. The age, body mass index (BMI) and International Federation of Gynaecology and Obstetrics (FIGO) stage of patients involved in the study are presented in the table. N is the number of samples for a given variable. The p-value shows the results of the nonparametric Wilcoxon rank-sum test. The interquartile range is given in brackets for age and BMI. The percentage of tumour number samples is given in brackets for FIGO stages.

Variable	N	Non-tumour, N=33[1]	tumour, N=33[1]	p-value[2]
Age	53	54 (46;62)	63 (56;70)	0.011
Unknown		N=13	N=0	
BMI	51	26.7 (23.4;28.5)	26.7 (24.0;32.4)	0.4
Unknown		N=13	N=2	
FIGO	31			
I		-	3 (9.7%)	
III/IV		-	28 (90%)	
Unknown		-	2	

[1] Median (IQR); n (%)

[2] Wilcoxon rank-sum test

3.3. Sample preparation for sequencing and differential expression (DE) analysis.

Total RNA with the miRNA fraction was isolated using a mirVana miRNA Isolation Kit (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's guidelines. RNA concentration, purity, and integrity were assessed by Qubit (Invitrogen) and Tape Station 2200 (Agilent Technologies, Santa Clara, CA, USA). RNA sequencing (RNA-seq) libraries were constructed from 1 μg of total RNA with an RNA integrity number (RIN) >8 with an Illumina TruSeq Stranded Total RNA Library Prep Gold (Illumina, San Diego, CA, USA). Small RNA-sequencing libraries were prepared with TruSeq Small RNA Library (Illumina) following the manufacturer's instructions. Indexed libraries were pooled, clustered using cBot, and sequenced on an Illumina HiSeq 4000 platform, generating 150 base pairs (bp) paired-end reads (2 × 75 bp).

3.4. Differential expression of genes

RNA-sequencing data was assessed for quality with FastQC (Babraham Institute, Cambridge, United Kingdom) and multiQC[46] before and after the trimming and filtering steps. Next, soft trimming and filtering with Bbduk (DOE Joint Genome Institute, Berkeley, CA, USA) were performed. Reads were mapped to the GRCh38 reference genome with the splicing-aware aligner STAR with quantMode for gene counts[47]. Ensemble annotation version (v)GRCh38.98 was used. Raw gene counts were used as the input for differentially expressed gene (DEG) identification. Raw count data were filtered to identify genes expressed at low levels, less than 10 raw counts in the smallest library, considering corresponding equivalent counts per million (CPM) values. The counts were normalised using a weighted trimmed mean of the log expression ratios (trimmed mean of M values (TMM))[48]. The data were transformed for linear modelling using voom[49]. Linear modelling and empirical Bayes moderation were applied using the limma package to assess the DEGs[50]. A false discovery rate (FDR) for multiple testing was applied for the correction of p-values.

3.5. Differential expression of miRNAs

Small RNA-seq data were assessed for quality with FastQC and multiQC[46]. Reads were trimmed from adapter sequences with cutadapt[51]. After adaptor trimming, the reads were filtered for quality and length in the 16-28 bp range with cutadapt. miRNA detection was performed with miRge v2.0 using Bowtie for alignment and 25iRbase v22[52]. Raw miRNA counts were used as input for the differential expression of miRNAs (DEM) analysis. Raw count data were filtered to identify miRNAs expressed at low levels, less than 5 raw counts in the smallest library, considering equivalent CPM values. The counts were normalised using a weighted trimmed mean of the log expression ratios[48]. The data were transformed for linear modelling using voom[49]. Linear modelling and empirical Bayes moderation were applied using the limma package to assess the differential expression of miRNAs[50]. FDR multiple testing was applied for the correction of p-values.

3.6. Differential expression of lncRNAs

Total RNA-seq data were assessed with FastQC; next, they were soft-trimmed and filtered with Bbduk. Bbduk tool was also used to filter fragments that originated from

ribosomal RNAs (rRNAs). Cleaned reads were mapped to the human reference genome GRCh38 using the splice-aware aligner STAR. Counts per gene were estimated by RNA-seq with Expectation-Maximisation (RSEM v1.3). The quality of the mapping was assessed with Qualimap v2.2.1, the flagstat module from Samtools, and RseQC. Mapped reads were also visually inspected through Integrative Genomics Viewer. GENCODE version 19 annotations were used to define lncRNAs. For each cancer status, two filters were applied to eliminate unreliable gene measurements: RPKM values in the 50th percentile larger than 0 and RPKM values in the 90th percentile larger than 0.1 were the criteria. Genes for which the values exceeded these criteria were considered to be detectable in a given cancer status[53]. Differential expression analysis of lncRNAs (DEL)was performed using the lncDiff package[54].

3.7. ceRNA Network construction

DELs, which can be targeted by DEMs, were first identified in the DIANA-LncBase v3 database[55]. Next, DEGs, which are targets of DEMs, were identified using miRTarBase 8.0[56]. The DEMs that overlapped in the DEM-DEL and DEM-DEG subnetworks were established as the core of the ceRNA network. The simple relations between the subnetworks connected through common DEMs were used to construct the network. ceRNA network was visualised using Cytoscape[57].

3.8. Identification of TFs

The comprehensive list of human TFs was downloaded from https://github.com/InesdeSantiago/SeqQC.blog/tree/gh-pages. The list was built based on the most used databases and well-cited publications [58–65]. Specifically, TFs in at least two databases were included in the list of 1924 TFs. The DEG list was searched, and TFs were identified. Next, the targets of the identified TFs were identified through a search of TRRUST database v2[66]. To identify TFs among ceRNAs, the results of the previous step were searched and only genes within the ceRNA network were retained. The identified TFs and their targets obtained from the ceRNA network interconnection were visualised with Cytoscape[57]. The hub regulators were identified with the CytoHubba[67] plugin in Cytoscape. After the calculation of node scores, the 15 nodes with the highest degree of centrality were selected.

3.9. Functional analysis

Functional enrichment analysis was performed using clusterProfiler[68] in the Gene Ontology (GO) database[65], Kyoto Encyclopedia of Genes and Genomes (KEGG)[69], Reactome pathway knowledgebase[70] and Molecular Signature Database (Msig DB)[71,72]. The visualisation of the KEGG pathway was performed using the pathview R/Bioconductor package[73]. An enrichment cut-off FDR<0.05 was the threshold. The identification of upstream regulators was performed with Ingenuity Pathway Analysis QIAGEN (IPA) (QIAGEN Inc., https://digitalinsights.qiagen.com/IPA)[1]. The full list is attached in S1.

3.10. Drug repurposing

The transcriptional signature of the ceRNA network was used for identifying suitable candidates for drug repurposing. ConnectivityMap (Cmap)[41] and L1000CDS2[74] were used to identify potential drugs. Cmap (available at https://clue.io) and L1000CDS2 (available at https://maayanlab.cloud/L1000CDS2) are databases of gene signatures for various cell lines that are expressed in response to various perturbations, including exposure to various compounds. The Cmap Drug Repurposing Hub includes a collection of Food and Drug Administration (FDA)-approved drugs, clinical trial drugs and preclinical compounds. The fundamental score calculation was a non-parametric similarity measure based on the weighted Kolmogorov–Smirnov enrichment statistic. Connectivity scores are a standardised measure between -100 and 100. In this study, a cut-off for the connectivity scores was -90 in Cmap, and the results are presented in S2. L1000CDS2 is based on signatures obtained from the L1000 cost-effective high-throughput gene expression assays, and it includes expression measurements of 978 genes. L1000CDS2 calculates the overlap of gene sets based on input signatures and perturbation signatures. The outcome is the 50 best-matched signatures. The full results are attached in S3. The intersection of the results from these two tools was identified was considered a possible potential drug. The molecular targets of the drugs were identified using the IPA Knowledgebase[1]

4. Results

4.1. Sample size estimation

The power of the study 80% was achieved by 12 samples per mRNA group, 16 samples per lncRNA group and 20 samples per miRNA group. To obtain a high power of 90%, 16 samples per mRNA group, 21 samples per lncRNA group and 27 samples per miRNA group would be needed. Ultimately, the groups consisted of 33 samples per group, allowing for power higher than 90% for each of the analyses performed.

4.2. RNA isolation

To obtain high-quality sequencing results, only the RNA isolates were selected, which fulfilled the criteria of RIN > 8 and at least 1 µg of total RNA. An example of the analysis result by Tape Station 2200, RIN 8.4 of one of the samples included in the analysis is presented in Figure 2.

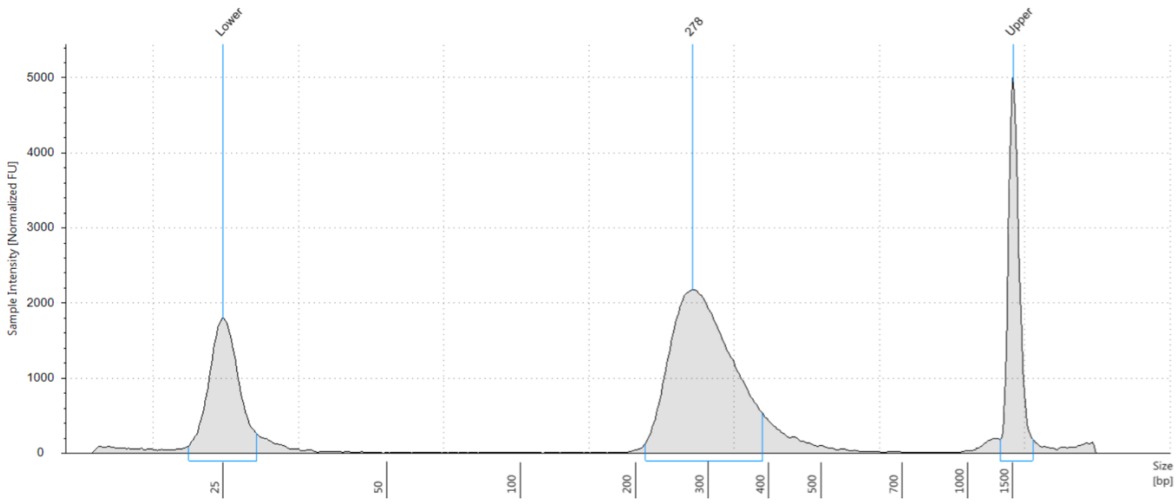

Figure 2 The Tape Station traces of a sample with RIN 8.4. The diagram represents the sample intensity in fluorescence unit (FU) versus the size of the RNA fragment in base pairs (bp).

4.3. Differential expression analysis

Transcriptome and small RNA sequencing of 66 patient tissues, including 33 HGSOC and 33 control samples (hereafter, non-tumour) were performed. The differential expression (DE) analysis revealed 2330 genes with downregulated expression and 2570 genes with upregulated expression (Figure 4C). Analysis of DEMs (Figure 4B) revealed 146 miRNAs with downregulated expression and 132 miRNAs with upregulated

expression. DELs analysis (Figure 4A) revealed 1058 lncRNAs with downregulated expression and 1054 lncRNAs with upregulated expression. The results of pathway analysis based on GO and KEGG databases from all the DEGs are presented in Figure 3. Pathways in cancer and PI3K-Akt signalling pathway were the results with the highest gene ratio of all. Many other significantly enriched KEGG pathways (FDR <0.05) are immune system-related pathways, such as Epstein-Barr virus infection, cytokine-cytokine receptor interaction, systemic lupus erythematosus or leucocyte transendothelial migration. GO BP enrichment significant (FDR <0.05) results are mostly immune system and stroma-related processes and pathways, including leukocyte migration, extracellular structure organisation, T cell activation or leukocyte proliferation. From the significantly changed (FDR<0.05) the most underexpressed genes in comparison to control tissue in HGSOC were prokineticin 1 (*PROK1*) (logarithm Fold Change (logFC) -6.71), angiopoietin-like 5 (*ANGPTL5*) (logFC -6.66), protocadherin 11 X-linked (*PCDH11X*) (logFC -6.49), steroidogenic acute regulatory protein (*STAR*) (logFC -6.48), WAP, follistatin/kazal, immunoglobulin, kunitz and netrin domain containing 2 (*WFIKKN2*) (logFC -6.46) and the most overexpressed secretory leukocyte peptidase inhibitor (*SLPI*) (logFC 7.89), solute carrier family 34 member 2 (*SLC34A2*) (logFC 8.43), mucin 16 (*MUC16*) (logFC 8.47), ceruloplasmin (*CP*) (logFC 8.51), claudin 3 (*CLDN3*) (logFC 8.52).

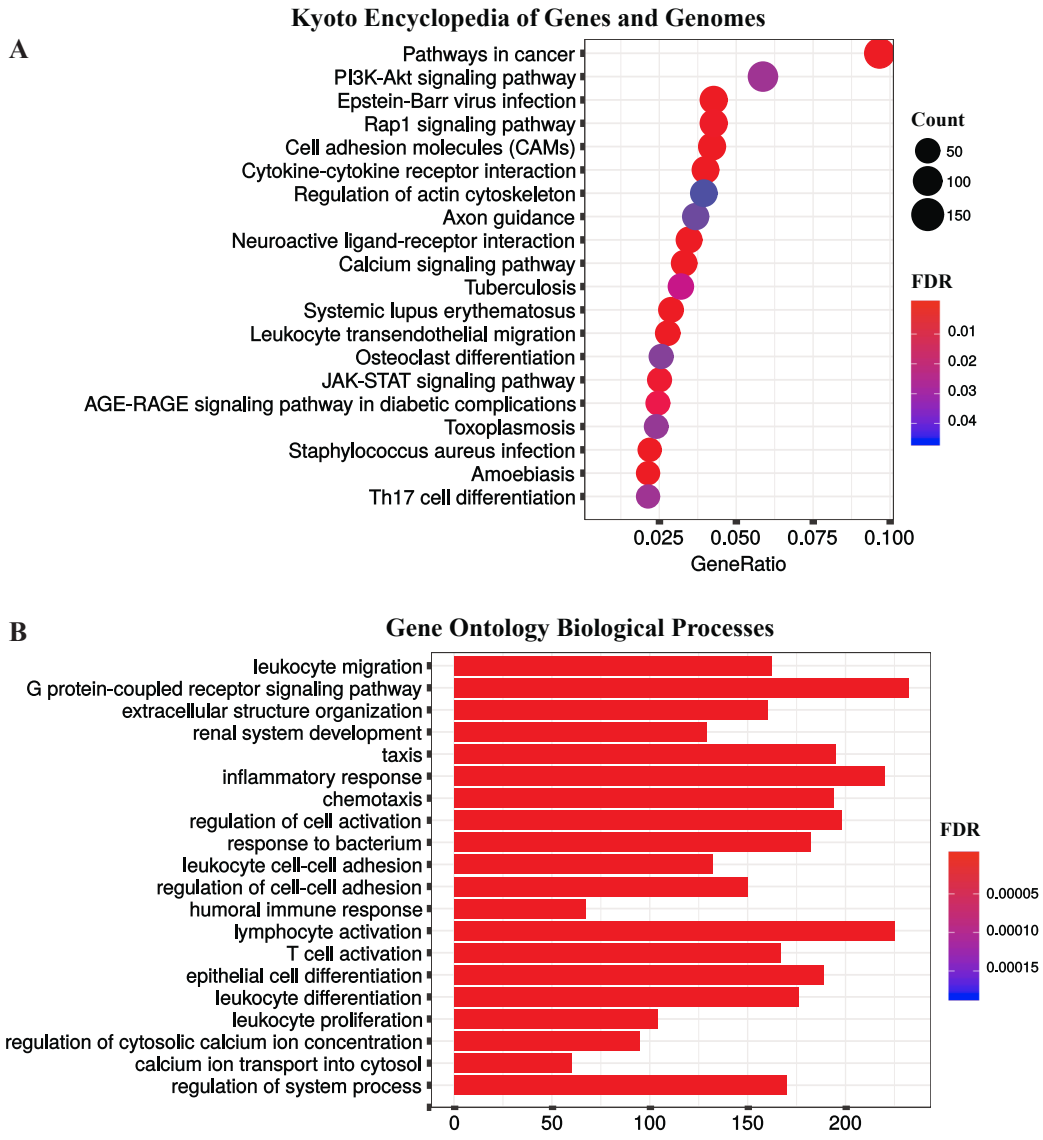

Figure 3 Top 20 significant results of functional enrichment of differentially expressed genes (DEGs) in High-Grade Serous Ovarian Cancer (HGSOC). (A) Gene Ontology (GO) Biological Processes (BP) enrichment results for the HGSOC DEGs (B) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment results for the HGSOC DEGs.

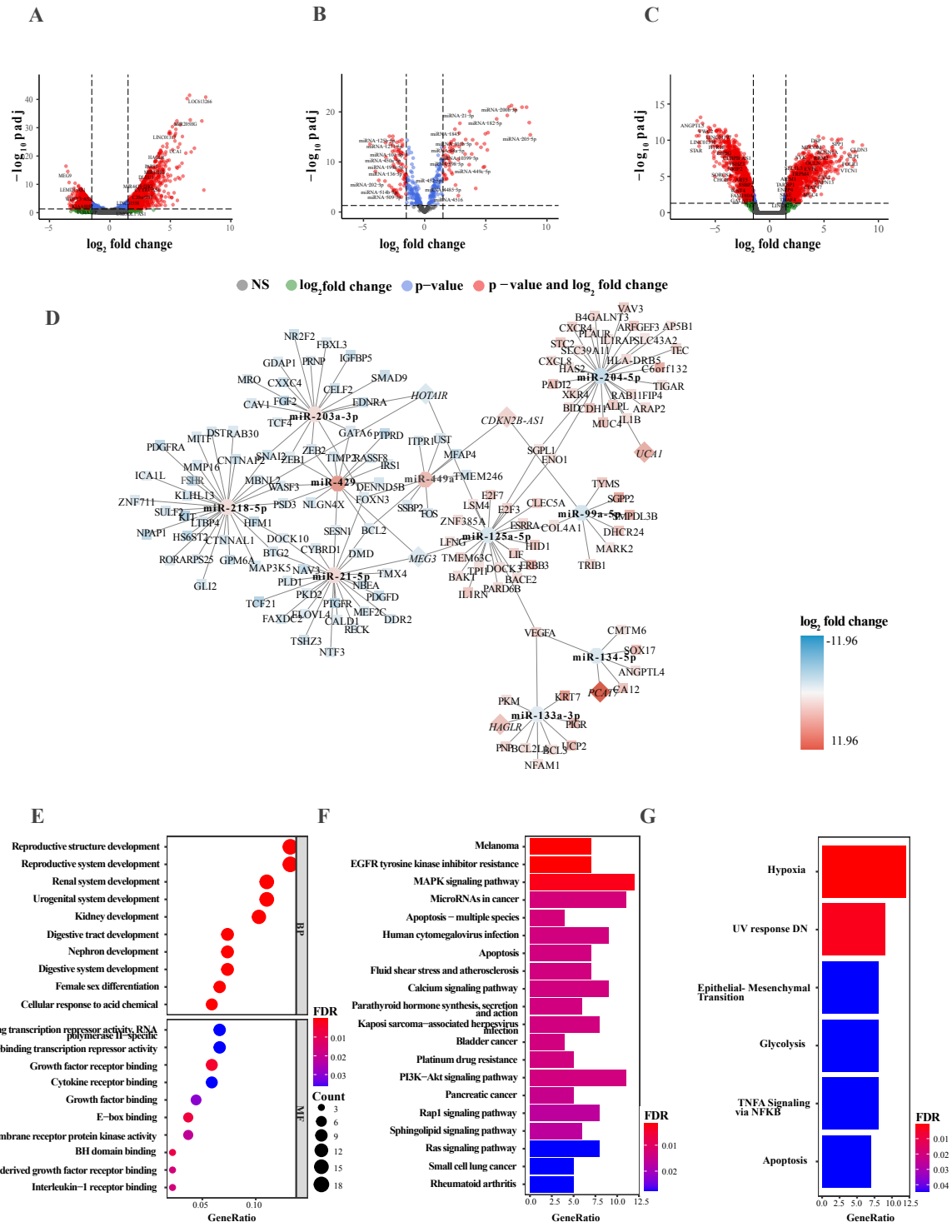

Figure 4 Differential expression analysis results, competing endogenous RNA (ceRNA) network and results of functional pathway enrichment analysis of the ceRNA sets in High-Grade Serous Ovarian Cancer (HGSOC). (A) Volcano plot presenting the differential expression of lncRNAs (DELs), NS, non-significant. (B) Volcano plot presenting the differential expression of miRNAs (DEMs). (C) Volcano plot presenting the differential expression of genes (DEGs). (D) ceRNA network of HGSOC visualised in Cytoscape. The logarithm 2-fold change of molecules is represented by the gradually changing colour of the nodes ranging from -11.95 to 11.96. All genes with downregulated expression are presented as blue nodes, and all genes with upregulated expression are presented as red nodes. I Gene Ontology (GO) enrichment results for the HGSOC ceRNA network. (F) Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment results for the HGSOC ceRNA network. (G) Molecular Signatures Database (Msig DB) hallmark enrichment results for the HGSOC ceRNA network.

4.4. Competing endogenous RNA Network

To construct a ceRNA network, the experimentally observed and/or highly predicted DEM targets, including mRNAs and lncRNAs have been obtained. Table 2 shows the number of molecules involved in each interaction set. Twelve unique pairs of lncRNAs targeted by DEMs were identified. A total of 1258 unique DEMs and their target mRNAs were identified. The sets of overlapping DE lncRNA–miRNA and miRNA–mRNA allowed the construction of a ceRNA network. The ceRNA network consists of six unique lncRNAs, ten unique miRNAs and 141 unique mRNAs, which are listed in Table 3. The ceRNA network was visualised in Cytoscape (Figure 4D). The network included five miRNAs with upregulated expression and five miRNAs with downregulated expression. The lncRNAs *HOTAIR* and *MEG3* are targets for each of the three network internal miRNAs. The lncRNA *CDKN2B-AS1* are targets of two miRNAs involved in the ceRNA network. The lncRNAs *PCAT7*, *HAGLT* and *UCA1* are targets of single miRNAs in the network. All miRNAs targeted at least five mRNAs. The protein-coding *VEGFA, FOXN3* and *BCL2* mRNAs were targeted by more than two miRNAs.
The functional enrichment analysis of GO (Figure 4E), KEGG pathway (Figure 4F) and Msig DB-based hallmark signatures (Figure 4G) was focused on DEGs involved in the HGSOC ceRNA network. Among the statistically significant (FDR<0.05) results of the GO biological processes (BP) analysis, the DEGs were related to reproductive structure and system development, as well as other developmental processes. The GO molecular function (MF) most enriched terms were DNA binding transcription repressor activity and growth factor binding. The highest gene ratio among the 20 most significantly enriched KEGG pathways was observed in the MAPK signalling pathway, miRNAs in cancer and the PI3K-Akt signalling pathway. The hallmark signatures enriched in the ceRNA network were hypoxia, genes downregulated by ultraviolet (UV) radiation (UV response DN), the epithelial-mesenchymal transition (EMT), glycolysis, tumour necrosis factor-alpha (TNF-A) signalling via nuclear factor-kappa B (NFkB) and apoptosis. Hallmark gene sets summarise and represent specific well-defined biological states or processes. These gene sets were generated through a computational methodology based on the identification of sets with overlapped genes and retention of genes that display coordinated expression.

The detailed PI3K-Akt pathway is visualised in Figure 5. Among all DEGs, genes with disrupted expression in the PI3K-Akt signalling pathway included phosphoinositide

3-kinase (PI3K) and Akt serine/threonine kinase 1. Large red rectangles in Figure 5 mark the gene groups differentially expressed in the ceRNA network; these genes encode growth factors (GFs), receptor tyrosine kinases (RTKs), extracellular matrix (ECM) components, and insulin receptor substrate 1 (IRS1), which directly and indirectly influence PI3K class I expression. Furthermore, pro-survival and anti-apoptotic Bcl-xL (a Bcl ligand) and Bcl-2 were also found dysregulated in the HGSOC ceRNA network.

Table 2. Competing endogenous RNA (ceRNA) network of high-grade serous ovarian cancer (HGSOC) characteristics. The table includes the number of unique sets of molecules and different molecules obtained separately during the construction of the ceRNA network.

Relation	No. of unique sets	No. of unique lncRNAs	No. of unique miRNAs	No. of unique mRNAs
lncRNA–miRNA	12	6	10	-
miRNA–mRNA	1258	-	119	622
lncRNA–miRNA–mRNA	203	6	10	141

Table 3. Competing endogenous RNA (ceRNA) network of High-Grade Serous Ovarian Cancer (HGSOC). Coding and noncoding RNA molecules involved in the construction of the ceRNA network of HGSOC are shown.

Unique molecules in the ceRNA network	**Official gene symbol**
lncRNA	*MEG3, HOTAIR, PCAT7, HAGLR, UCA1, CDKN2B-AS1*
miRNA	miRNA-125a-5p, miRNA-203a-3p, miRNA-134-5p, miRNA-133a-3p, miRNA-204-5p, miRNA-21-5p, miRNA-429, miRNA-218-5p, miRNA-99a-5p, miRNA-449a
mRNA	*LIF, VEGFA, ERBB3, BAK1, E2F7, PARD6B, TPI1, ENO1, COL4A1, SGPL1, IL1RN, DOCK3, LFNG, ESRRA, CLEC5A, ZNF385A, LSM4, BACE2, HID1, E2F3, TMEM63C, GDAP1, EDNRA, ZEB2, SNAI2, CAV1, TCF4, ZEB1, SMAD9, CXXC4, MRO, PRNP, FOXN3, FBXL3, NR2F2, GATA6, IGFBP5, CELF2, FGF2, CA12, SOX17, ANGPTL4, CMTM6, UCP2, KRT7, PKM, PNP, BCL2L1, BCL3, PIGR, NFAM1, HAS2, IL1B, IL1RAP, CXCL8, PLAUR, ALPL, TEC, CDH1, XKR4, VAV3, AP5B1, HLA-DRB5, SLC39A11, ARAP2, BID, B4GALNT3, STC2, ARFGEF3, MUC4, C6orf132, SLC43A2, RAB11FIP4, PADI2, TIGAR, CXCR4, BCL2, RECK, SESN1, BTG2, MEF2C, NTF3, TCF21, DDR2, CYBRD1, ELOVL4, NBEA, PDGFD, PKD2, FAXDC2, CALD1, DOCK10, DMD, TMX4, TSHZ3, PLD1, PTGFR, NAV3, WASF3, PTPRD, RASSF8, TIMP2,DENND5B, NLGN4X, PSD3, IRS1, MBNL2, CNTNAP2, KLHL13, GPM6A, ZNF711, RPS25, DST, MITF, PDGFRA, GLI2, SULF2, RORA, MMP16, MAP3K5, LTBP4, ICA1L, HS6ST2, HFM1, FSHR, CTNNAL1, RAB30, KIT, NPAP1, TRIB1, SGPP2, DHCR24, MARK2, TYMS, SMPDL3B, ITPR1, UST, SSBP2, MFAP4, FOS, TMEM246*

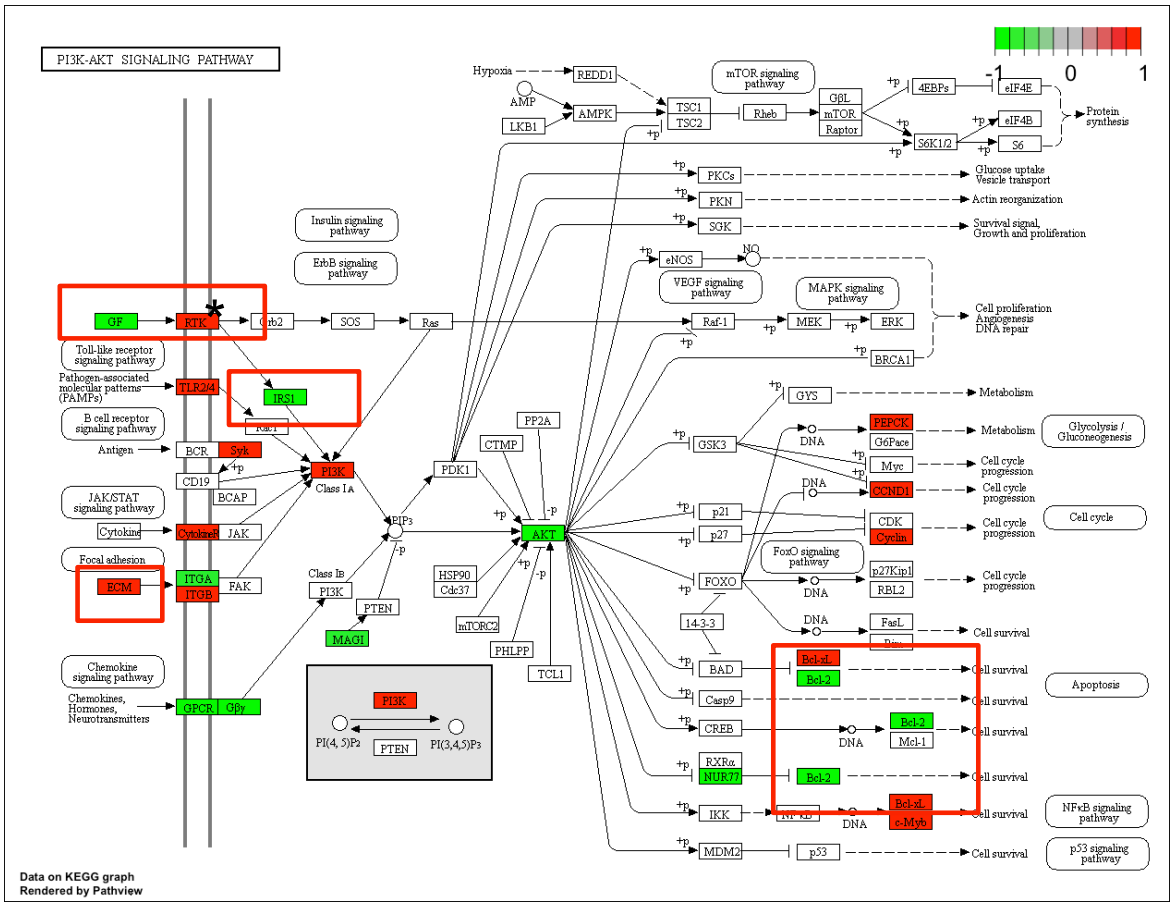

Figure 5 PI3K-Akt pathway in High-Grade Serous Ovarian Cancer (HGSOC). The shading in red represents upregulated genes (or group of genes) and green indicates downregulated genes (or group of genes) for all differentially expressed genes. The red rectangle highlights a group of genes in the HGSOC ceRNA network. The star represents a group of genes whose collective expression is reversed when all DEGs or only ceRNA network-involved mRNAs are considered.

4.5. Identification of regulatory TFs

Based on a curated list containing 1924 TFs, 151 molecules on the DEGs list were identified. The TRRUST database presents information about TF targets. A total of 1464 targets were identified, but to identify the TFs related to the ceRNA network, the list was restricted to include targets only in the ceRNA network. The results are presented as a circular plot with network-related information in Figure 6A, and the characteristics of the network are listed in Table 4. Eight TFs (*FOXP2, EGR1, E2F1, ESR1, SNAI2, FOS, MITF,* and *TCF4*), and seven genes were TF targets (*BCL2, IL1B, MAP3K5, PLAUR, CXCL8, VEGFA,* and *CDH1*) were identified as hub regulators. The Reactome-based functional enrichment of hub genes is illustrated on a circular plot presented in Figure 6B. The enriched terms included oestrogen-related signalling, signalling by interleukins, especially IL-4 and IL-13, and engagement in the SUMOylation process. In the QIAGEN IPA Knowledgebase [1]-based analysis of upstream regulators of ceRNA genes, 561

regulators were identified (S1). The groups of regulators included mainly transcription regulators, miRNAs, cytokines, growth factors, and transmembrane receptors. Among the transcription regulators, 13 overlapped with the TFs identified from among the DEGs and are presented in Figure 6A. These regulators were *SNAI2, E2F1, ZEB2, RUNX3, PKNOX2, MITF, TWIST2, ETV5, KLF4, EGR1, GATA6, IKZF3* and *FOS*. Three of the key ceRNA network miRNAs (miRNA-218-5p, miRNA-204-5p and miRNA-21-5p) were also identified independently as upstream regulators.

Table 4. Characteristics of the transcription factor (TF)–competing endogenous RNA (ceRNA) interactions in High-Grade Serous Ovarian Cancer (HGSOC). The table presents the number of nodes, edges and the average number of neighbours in the TF–ceRNA network.

Characteristics of the network	
Number of nodes	68
Number of edges	107
The average number of neighbours	3

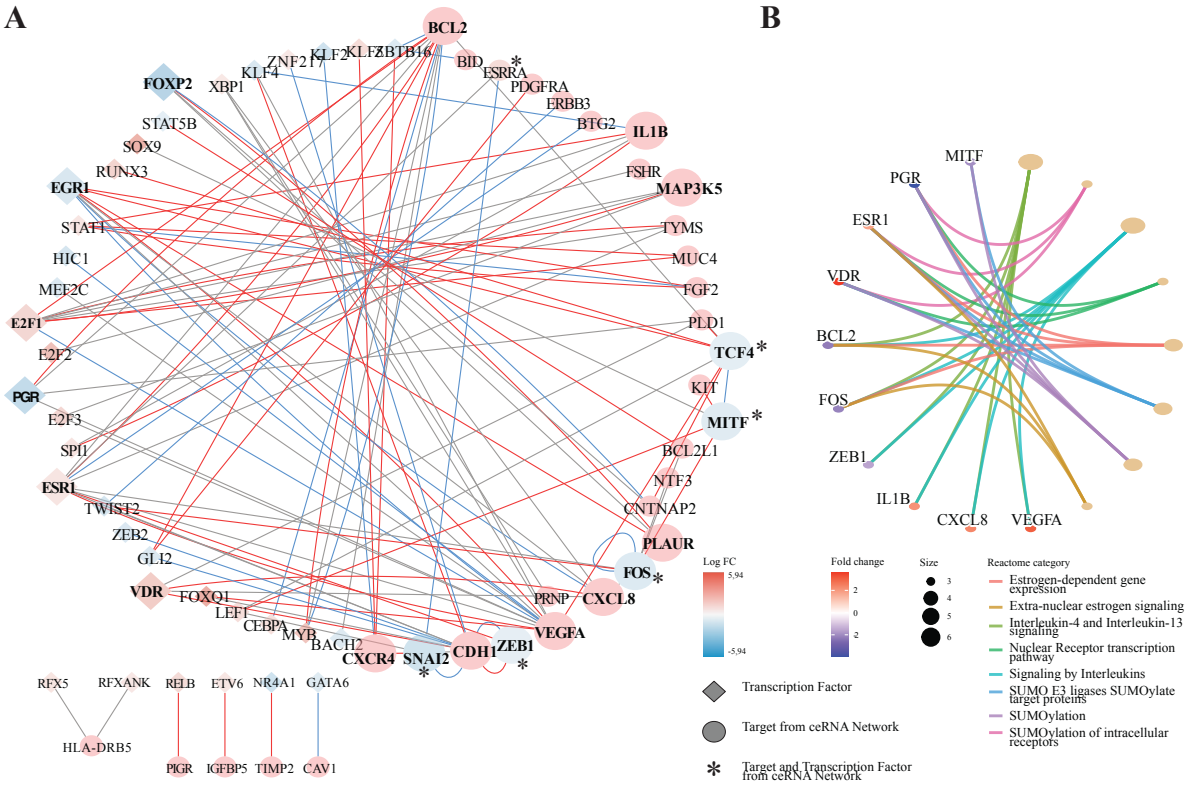

Figure 6 Transcription factor (TF)–competing endogenous RNA (ceRNA) network based on High-Grade Serous Ovarian Cancer (HGSOC) and pathway analysis of hub regulators. (A) A network containing key TFs identified from among differentially expressed genes (DEGs) and their targets within the HGSOC ceRNA network. (B) Reactome terms enriched with hub regulators.

4.6. Drug repurposing

A data-driven systematic approach to drug repurposing is represented by the web tools Cmap and L100CDS2, which were applied to the ceRNA network transcription profile. The Cmap results were restricted to compounds with connectivity scores lower than -90. L1000CDS2 results were based on the top 50 results of a gene set enrichment analysis (GSEA). However, the top 50 results included some of the same compounds at different doses. We combined the results of both analyses and found 53 potential drug candidate compounds and 34 mechanisms of action (MOA) shared by the identified compounds. The most common MOA, shared by eight compounds, was PI3K inhibition. A heatmap including all identified potential drugs is presented in Figure 7, which also includes information on known MOAs. Two compounds found with both algorithms were PI-103 and ZSTK-474, which are known PI3K inhibitors. Their chemical structures are presented in Figure 7, and the list of identified potential drug candidates is also listed in Table 5. Figure 8 presents the results of an IPA Knowledgebase [1] screening for known molecular targets of PI-103 and ZSTK-474.

Table 5. Drug repurposing analysis results based on the High-Grade Serous Ovarian Cancer (HGSOC) competing endogenous RNA (ceRNA) network. The list of compounds identified based on the genes in the ceRNA network of HGSOC with Cmap and L100CDS2 after the cut-off was considered.

	Cmap	**L100CDS2**
Candidate drugs	PD-184352 (S1020), PI-103, ZSTK-474, AS-703026, cefalexin, dactolisib, TPCA-1, selumetinib, KU-0063794, AZD-8055, PI-828, homoharringtonine, TC-2559	PI-103, ZSTK-474, BRD-K19220233, BRD-K52911425, BRD-K12867552, GSK-1059615, NU-7026, NVP-AUY922, saracatinib, HG-6-64-01, QL-X-138, torin-2, WZ-4-145, BRD-A73909368, busulfan, trichostatin A, spironolactone, manumycin A, demeclocycline, TWS-119, BRD-K77947974, canertinib, KU-55933, celastrol, BRD-K35920785, deprenalin, PFI-1, 598226, I-BET151, (+)-JQ1, NCGC00184834-01, 3544, rythmol, DCC-2036, GSK-2126458, BRD-A59145032, P0030, GDC-0980, S1170, N-Cyclopropyl-5-(thiophen-2-yl)isoxazole-3-carboxamide, BRD-K74767048, R3904
Overlapping drugs		PI-103, ZSTK-474

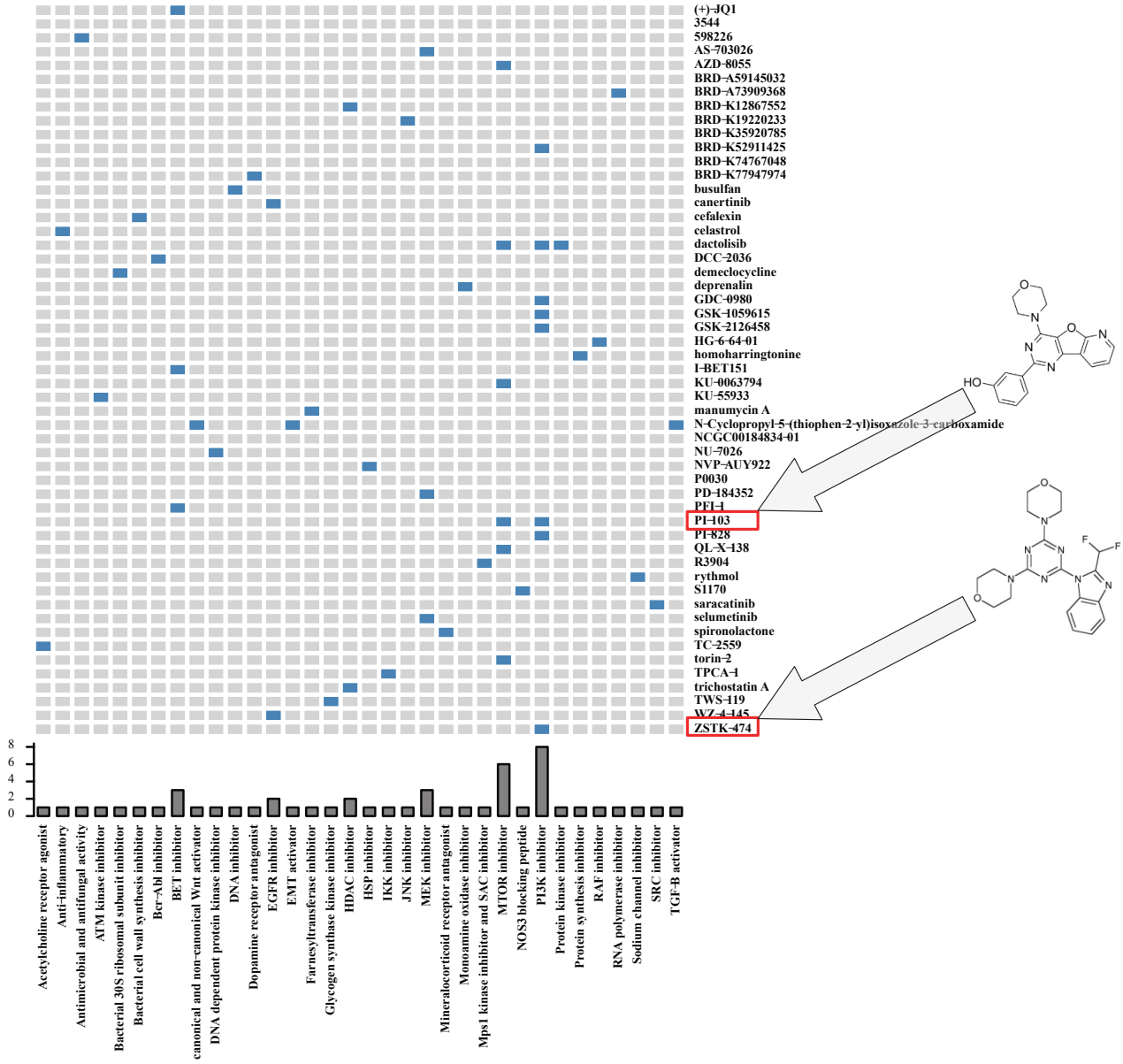

Figure 7 Heatmap of the mechanism of action (MOA) and candidate compounds for High-Grade Serous Ovarian Cancer (HGSOC) repurposed drugs. The two compounds, which were identified by both Cmap and L1000CD2 tools, are presented with their chemical structures.

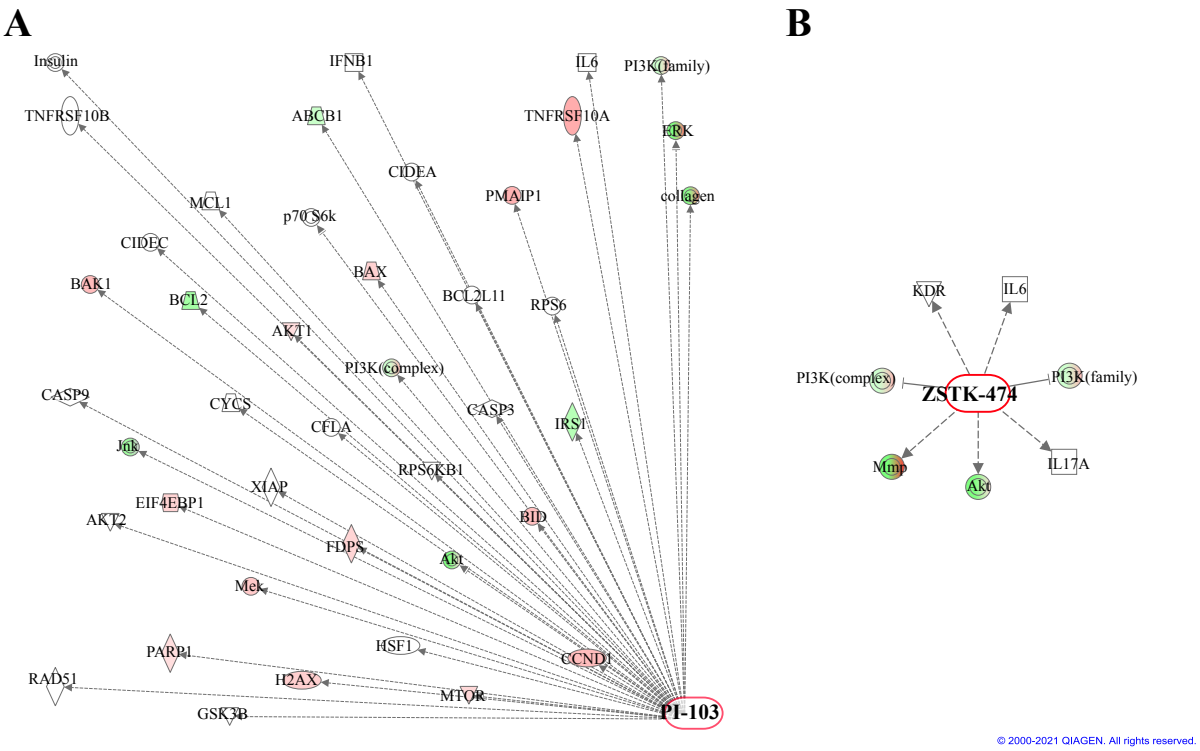

Figure 8 Known molecular targets of PI-103 and ZSTK474 based on a QIAGEN IPA Knowledgebase [1] analysis. The diagrams are colour-coded based on the differential expression of targets in the HGSOC dataset. Red indicates an overexpressed gene, and green indicates a gene with downregulated expression. (A) Known molecular targets of PI-103. (B) Known molecular targets of ZSTK-474. The graph was adapted from QIAGEN IPA[1].

5. Discussion

This study was aimed at investigating the complex topic of RNA crosstalk and gene expression regulation in High-Grade Serous Ovarian Cancer (HGSOC). A competing endogenous RNA (ceRNA) network was constructed, and it revealed the key mechanisms and axes of deregulation in this disease. The results of the study confirmed that the PI3K/Akt pathway is a main HGSOC-related pathway. Moreover, the results showed that this pathway is also highly dependent on non-coding RNA (ncRNA) and protein-coding RNA (mRNA) interactions. Adding another layer of complexity, transcription factors (TFs) were found to be directly involved in the ceRNA network. The function of these identified TFs were further explored and showed influence on immune system signalling, EMT and post-translational protein modification. Studying these communication partners and characterising the mechanism of their cooperation is crucial to understanding the complexity of HGSOC. However, studies that have sequenced and investigated mRNAs, micro RNAs (miRNAs) and long ncRNAs (lncRNAs) simultaneously in HGSOC patients are rare.

The differential expression (DE) analysis resulted in over 4000 DEGs, over 200 DEMs and over 2000 DELs. The functional analysis of identified DEGs, showed clearly, that they are strongly cancer-related. All of them could potentially be biomarkers for HGSOC. Further integrated analysis of the expression of these DEGs with clinical outcome of the patients, the functional assays *in-vitro* of the gene knock-out, under- and overexpression and understanding of the involvement of these genes in tangled cell's internal network would be necessary for confirmation of such potential. *PROK1* has been previously described as a colorectal cancer plasma-derived biomarker[75]. In OC prokineticin receptor 1 (*PKR1*) positive exosomes *PROK1* itself showed to be involved in angiogenesis[76,77]. *ANGPTL5* belongs to the angiopoietin-like family, which has been described to have an autocrine and paracrine activity influencing different stages of cancer development, including angiogenesis[78,79]. Protocadherins, including *PCDH11X*, are cell adhesion molecules, which altered expression has been found in a variety of cancers. Exact mechanisms of protocadherins aberrated expression influencing cancer development and progression are however not yet described[80]. There is not much known also about the mechanisms of *STAR* proteins in cancer cells either, however its high expression has been described to correlate with an unfavourable outcome for breast cancer patients[81]. *WFIKKN2*, together with *WFIKKN1* play important role in regulating

several growth factors in an organism. Disturbance of the balanced and tight regulation might be one of the reasons for developing diseased phenotype, including the development of cancer[82,83]. The described proteins were the five most underexpressed genes in HGSOC samples versus control samples identified in the study. Unfortunately, there are still many unknown functions and interconnections with other important molecules. The most overexpressed genes identified by DE analysis are *SLPI*, *SLC34A2*, *MUC16* and *CLDN3*. The upregulation of *SLPI* expression has been described in many cancers. It has been called a survival and proliferation factor and is known to influence inflammatory response by regulation of NF-κB activation and neutrophil proteases inhibition[84]. The molecule draws attention also as a possible target against tumour metastasis, as it increases TF – *FoxM1* binding to its target genes and has been described to physically interact with known tumour suppressor protein *Rb*[84–86]. The overexpression of *SLC34A2* in different cancer has been already often described. It is known to be a potential biomarker and possible precise drug target. The antibody-drug conjugate treatment directed to the gene product of *SLC34A2* was effective in preclinical studies on mouse models[87]. *MUC16* encodes CA-125, which has high expression in patients' blood and is the main biomarker used for OC diagnosis nowadays[88]. Tight-junction protein *CLDN3* has been also already previously described to be overexpressed in OC, moreover and its' silencing by siRNAs showed potential therapeutic effects *in-vitro*[89–92].

The pathways and processes including Epstein-Barr virus infection, cytokine-cytokine receptor interaction, systemic lupus erythematosus or leucocyte transendothelial migration, leukocyte migration, extracellular structure organisation, T cell activation and leukocyte proliferation are strongly immune system and stroma-related processes and pathways were identified as the enriched pathways, based on DEGs. The result confirms strong tumour dependence on tumour microenvironment and stroma cells interaction. These complex cell-cell interactions and communication are strictly regulated, often also by ncRNAs[93,94]. Identification of crucial ncRNAs playing role in HGSOC and restriction of the DEGs to the strongly ncRNAs regulated might have also an important clinical implication.

ceRNA network construction allowed for the identification of key ncRNAs and TFs that regulate gene expression and ncRNAs-regulated mRNAs in HGSOC. The ceRNA network evaluated in this study included ten key miRNAs, miRNA-125a-5p, miRNA-203a-3p, miRNA-134-5p, miRNA-133a-3p, miRNA-204-5p, miRNA-21-5p,

miRNA-429, miRNA-218-5p, miRNA-99a-5p, and miRNA-449a, and six lncRNAs, *MEG3, HOTAIR, PCAT7, HAGLR, UCA1* and *CDKN2B-AS1*. Both the lncRNAs and mRNAs have multiple miRNA response elements (MREs), and they competed for available miRNAs.

ncRNAs play a regulatory role, and their abnormal expression levels influence key biological pathways. miRNAs constitute a large family of expression regulators that can target multiple genes. They exert an RNA-mediated interference (RNAi) effect on their targets[95,96]. miRNAs are intricately involved in cancer pathogenesis. The tumour suppressor function of miRNAs can lead to oncogene translation suppression, and miRNAs, which can function as oncogenes, inhibit tumour suppressor gene action[95,97]. Izumiya and colleagues[98] recognised the great importance of miRNAs and collected information obtained through functional assays of cancer-associated miRNAs. They have gathered evidence of miRNAs contributing to an altered cell cycle, apoptosis, cell invasion, angiogenesis and general oncogene-associated miRNAs and tumour suppressor-associated miRNAs. Many of the key miRNAs identified in the ceRNA network (e.g., miRNA125a-5p, miRNA-21) in our study have been previously recognised by Izumiya and colleagues as crucial to cancer[98]. miRNA-125a has been described as a suppressor of cancer cell invasion through the inhibition of *ERBB2/3* expression[99]. miRNA-203a-3p has been previously described as a candidate tumour suppressor in colorectal carcinoma, and its overexpression has been shown to inhibit cell invasion and migration and to increase the apoptosis rate in colorectal cancer cell lines[100]. Literature corroborating the importance of miRNA-204-5p has indicated that this miRNA is a well-characterised tumour suppressor[101,102]. miRNA-204-5p is involved in the regulation of the tumour growth and metastasis of many tumour types[101,102]. The other miRNAs identified in the network have also been previously shown to be differentially expressed in other cancer types (miRNA-429[103], miRNA-218-5p[104], miRNA-99a-5p[105], and miRNA-449a[106]). Moreover, the expression of miRNA-99a-5p is elevated in the serum of OC patients and thus may serve as a potential diagnostic biomarker[107].

Studying other ncRNAs, such as lncRNAs, is a challenge, as lncRNAs are expressed at low levels, and their functions are not well understood. LncRNAs can act as miRNA decoys in human development and pathologies, which means that they indirectly regulate transcription by competing for shared miRNAs[53,108,109]. One of the key lncRNAs identified in our study, *PCAT7,* has been previously described as a

regulator of another key molecule in our ceRNA-network, miRNA-134-5p, in nasopharyngeal carcinoma. Decreased *PCAT7* expression has been shown to suppress tumour cell proliferation[110]. The other key lncRNAs in HGSOC identified in our study, *UCA1, HOTAIR,* and *MEG3* may be involved in a causal relationship in many cancer types, such as liver, colon, stomach or breast cancer[111]. *HOTAIR* and *MEG3* were the most intertwined key lncRNAs in our ceRNA network, as they each target three miRNAs in the network. *HAGLR* in lung cancer and colon cancer has been shown to support tumour growth and has been characterised by its miRNA sponging ability[112,113]. Among the HGSOC ceRNAs in our study, the lncRNA *CDKN2B-AS1* is shown to be targeted by two DEMs; however, in the literature, many more miRNAs have been described to target *CDKN2B-AS1*, including let-7 family miRNAs, which are highly involved in cancer[114].

In addition to expanding the functional knowledge about miRNAs and their potential roles as non-invasive biomarkers, research on ncRNAs has clear clinical applications, as ncRNAs can be used in the identification of new targets for precision therapies. The disrupted gene expression of ceRNA-involved ncRNAs described in various cancer-related reports supports the supposition that ncRNAs exert an impact on cancer hallmarks. Therefore, it is important to understand the roles played by ncRNAs in cancer. Notably, functional screenings have shown that the disruption of a single miRNA can lead to abnormal cell proliferation, invasive cell behaviour or cancer cell apoptosis because miRNAs are at the centre of transcription regulation[102,115,116]. Knowledge about tumour-driving miRNAs may lead to the design of synthetic anti-sense oligonucleotides that might be used in treatments. Saini *et al.* reported that reconstitution of miRNA-708 in prostate cancer cell lines led to decreased tumorigenicity[115]. Anti-miRNA oligonucleotides (AMOs) or antagomirs can inactivate oncogenic miRNAs through RNA interference (RNAi), which can help slow tumour progression[117,118]. On the other hand, tumour suppressor miRNA expression can be disrupted in patients through transient expression systems applied through viral or liposomal delivery[117,119]. These systems might be useful for administering large quantities of miRNAs. Tremendous progress has been made in miRNA treatments; however, further studies are needed before miRNAs can be introduced for clinical management[117–120]. Understanding the complex network of RNA interactions might help in developing miRNA sponges, which are artificial transcripts with multiple MRE copies that are expressed at high levels. They can inhibit miRNA function specifically and effectively, making them

future therapeutic candidates[121]. Understanding the ceRNA network is important for designing better disease and gene knockout models. Epigenetic interactions and posttranscriptional regulation are not often considered in model design; however, they may be crucial for the success of model development. The study shows that analysis of the regulation of transcription is a complex matter and should take into consideration mRNAs and ncRNAs.

Gene expression regulation may include lncRNAs and miRNAs, but it also greatly depends on TFs. Therefore, the next step in the study was to identify the ceRNA network-related TFs in HGSOC. 15 ceRNA-related TFs were identified based on their degree of centrality as hub regulators and their ceRNA target molecules. Eight TF hub regulators: *FOXP2, EGR1, E2F1, ESR1, SNAI2, FOS, MITF,* and *TCF4* and seven TF target gene hub regulators (*BCL2, IL1B, MAP3K5, PLAUR, CXCL8, VEGFA, and CDH1*) were identified. *VEGFA* is a key mediator of tumour angiogenesis, which is a hallmark of cancer[122]. The identified TFs also overlapped with those discovered through another independent analysis of upstream regulators based on the QIAGEN IPA Knowledgebase[1]. The common TFs found in both analyses were *SNAI2, E2F1, ZEB2, RUNX3, PKNOX2, MITF, TWIST2, ETV5, KLF4, EGR1, GATA6, IKZF3* and *FOS*. Three TFs, *E2F1, SNAI2,* and *FOS,* were hub regulators and were identified separately by both aforementioned TF screenings. *E2F1* is an important regulator in OC and is useful as a molecular drug target in this disease[123,124]. *E2F1* influences many cell cycle genes and can activate apoptosis. Notably, *E2F1* is expressed at higher levels in more aggressive OC forms[125]. Additionally, during oncogenesis, *E2F1* upregulates the expression of *BCL2*, an anti-apoptotic gene product[126]. Many of the TFs identified in this study targeted *BCL2*. Another identified hub regulator, *FOS,* was a TF enriched in the following Reactome terms: Interleukin-4, Interleukin-13 signalling, oestrogen-dependent expression and SUMOylation. Dysregulation of IL-4/IL-13 signalling contributes to inflammation and initiation of the immune response, indicating that these interleukins contribute to tumour growth, cell adhesion changes and, ultimately, metastasis[127]. The immune system and microenvironment are generally very important components of cancer development and progression, and interactions between immune cells and cancer cells are necessary for establishing a niche for cancer stem cells and other cancer cells in an organism[128]. Steroid nuclear receptors, such as *ESR1*, influence oestrogen-dependent gene expression[129]. Dysregulation of steroid levels or the expression of their receptors might greatly impact cell functioning, leading to

tumorigenesis[129]. The SUMOylation process as it relates to cancer is garnering attention. It is a posttranslational protein modification that regulates processes such as DNA damage repair, immune responses, carcinogenesis, cell cycle progression and apoptosis [130] SUMOylation is often called the 'master repressor' of the immune system response and is considered a promising target to activate the immune system in the fight against cancer[131]. *SNAI2* has often been described by the important role it plays in the EMT, cell invasion and metastasis[132,133]. Interestingly, *SNAI2* has also been shown to be involved in the miRNA-regulatory network of the mesenchymal cell subtypes in serous ovarian cancer. miRNA-506 has been identified as an EMT inhibitor through its targeting of *SNAI2*[134].

The study also identified 141 mRNAs involved in the HGSOC ceRNA network. *VEGFA, BCL2, and FOXN3* are genes that were profoundly regulated in the network. They were highly intertwined through interactions with miRNAs and lncRNAs. Each of these mRNAs can be targeted by three of the HGSOC key miRNAs identified in this study. These multiple regulatory mechanisms indicate an important role played by the proteins expressed by these transcripts. Indeed, all the genes identified have been extensively characterised and have long been known to be involved in cancer. As described above, *VEGFA* is a key mediator of tumour angiogenesis, and its expression is necessary for tumour progression[122]. *BCL-2* is known for its antiapoptotic role[135], and *FOXN3* is a transcriptional inhibitor that influences cell proliferation[136]. Further investigation into the molecular and biological functions and pathways influenced by genes, or later, studies into proteins encoded by genes within a ceRNA network indicated the clear involvement of the identified genes in developmental processes. When the expression of genes crucial for development is disrupted, the genes can act as oncogenes[137]. In this study, the enriched molecular functions were DNA binding transcription repressor activity and growth factor binding molecules, which are commonly observed during active cell proliferation. Receptor binding or growth factors are important in internal and cell-to-cell communication, and the proper functioning of these processes is necessary for many fundamental mechanisms. These factors can influence cell proliferation, development, or responses to external stimuli[138]. In addition to the miRNA enrichment terms known to be related to cancer, in the ceRNA network, we identified enrichment in pathways such as EGF receptor tyrosine kinase (RTK) inhibitor resistance, the MAPK signalling pathway, and the PI3K-Akt signalling pathway. Tyrosine kinases (TKs) constitute a large group of important molecules in

cells. Their proper regulation is necessary for cell-to-cell signalling and cell growth, differentiation, motility, and adhesion[139–141]. Many TKs are known oncogenes. TK inhibitors, therefore, are promising cancer treatment options[139]. However, as with many other drugs, the responses to treatment can differ. Many mechanisms of resistance to TK inhibitors have been identified, including the amplification of target receptor expression, mutations in a receptor leading to impaired binding ability, low inhibitor uptake by tumour cells and signalling pathways bypassing through alternative pathways, which can lead to the activation of downstream effectors[140]. An example of alternative downstream pathway activation is constitutive Akt activation or unregulated PI3K-Akt pathway activity[141]. The PI3K-Akt pathway has also been found to be an enriched pathway in our ceRNA network, as mentioned above. Interestingly, a key ceRNA miRNA, miRNA-133a-3p, has been previously described to be directly related to PI3K-Akt signalling in prostate cancer because it directly targets multiple cytokine receptors important in the pathway[142]. This signalling pathway is often dysregulated in cancer and is connected with other pathways, such as the mammalian target of rapamycin (mTOR) and MAPK signalling pathways[143,144]. The PI3K-Akt pathway exerts a direct influence on cell apoptosis, the cell cycle and cell metabolisms, such as glycolysis and gluconeogenesis. Changes in cell metabolism are hallmarks of cancer, and the results of the presented study show functional enrichment of the hallmark signature, which indicated significant enrichment of the term glycolysis.

Treatments for OC led to drug resistance development, which is a challenge that must be overcome. Liu *et al.* suggested that the PI3K-Akt pathway is a key modulator of multidrug resistance in cancers[145]. Dysfunction of the PI3K-Akt pathway influences ABC transporter activity, the mTOR pathway, and, as previously mentioned, tumour metabolism. Drug repurposing enables compounds to be identified for use as treatments in diseases other than those for which it was originally developed. Drug repurposing shortens the timeframe of *de novo* developing drug development and allows faster and cheaper identification of compounds that might be later applied to particular diseases[39,146]. It was not surprising that in the pathway enrichment analysis, most of the identified compounds were PI3K and mTOR inhibitors. The first PI3K inhibitor was approved by the Food and Drug Administration (FDA) in 2014 for relapsed or refractory chronic lymphatic leukaemia. Since then, a few more PI3K inhibitors have been approved, and the effects of many more repurposed drugs have been tested in clinical trials[147]. Additionally, PI3K inhibitors, especially in combination with other drugs, such

as PARP inhibitors, have shown promising results in ovarian cancer studies[148,149]. In our study, we identified two PI3K inhibitors, ZSTK474 and PI-103, with both CMap and L1000CDS2 tools. ZSTK474 selectively inhibits class I PI3K isoforms, mostly PI3Kδ. PI-103 is a multitargeted PI3K inhibitor. We screened the QIAGEN IPA Knowledgebase[1] to identify known molecular targets of ZSTK474 and PI-103. PI-103 targets many molecules, such as BCL2, BAK and BAX (which influence cell apoptosis), PARP1 (PARP inhibitors are gaining attention as potential treatment alternatives), insulin, and IRS1 (which profoundly influences cellular metabolism)[150]. In addition, other obvious molecular targets are in the PI3K family, namely, mTOR and Akt. In ovarian cancer cell lines, PI-103 enhanced sensitivity to cisplatin[151]. Most of the known molecular targets of ZSTK474 are directly involved in the PI3K signalling pathway. The effect of ZSTK474 on ovarian cancer has also been previously investigated, especially in combination with other inhibitors, such as EGFR inhibitors. Importantly, ZSTK474 has shown promising results and has been recommended for further clinical trials[152]. Specifically, as a single-agent, ZSTK474 has shown *in-vivo* antitumour activity in xenograft experiments by inhibiting tumour proliferation[153]. Further studies are necessary to confirm the clinical usefulness of these drugs for HGSOC patients. Notably, by restricting the access of DE molecules to key interacting ncRNAs and mRNAs, it was possible to select molecules with true and confirmed effects on OC cell lines. Additionally, through the drug repurposing analysis, we identified compounds never previously mentioned as potentially useful for treating HGSOC, such as cefalexin, a semisynthetic antibiotic, or homoharringtonine, a natural product of evergreen trees that is used in malaria treatment. Further validation studies *in-vitro* and *in-vivo* need to be performed to confirm these results before these compound candidates can be applied in clinical settings.

5.1. Conclusions

Epigenetic transcription regulation allows fine-tuning of gene expression regulation and can control key biological pathways in cancer. Therefore, the presented study aimed to investigate **the complex network of gene expression regulators in HGSOC**. The differential expression analysis of tissue-derived RNA profiles from 66 patients, including 33 samples derived from HGSOC primary tumour from ovary tissues and 33

control samples of ovarian tissues without oncological changes confirmed by histopathology examination has been performed.

The basis for reaching the goal was

1. **Investigation of differentially expressed genes, lncRNAs and miRNAs** between HGSOC primary tumour versus control samples. The analysis revealed 2330 genes with downregulated expression and 2570 genes with upregulated expression. Analysis of DEMs revealed 146 miRNAs with downregulated expression and 132 miRNAs with upregulated expression. DELs analysis revealed 1058 lncRNAs with downregulated expression and 1054 lncRNAs with upregulated expression.

It is not possible to discuss each differentially expressed gene, miRNA and lncRNA, so important was

2. **Identification of the key disrupted genes**, by restriction of the significant result and choice of top five up- and downregulated genes. They might be useful in the future as molecular targets. The up-and downregulated genes have been discussed and confirm previous findings described in the literature. Moreover, further restriction of key disrupted genes was performed by the construction of a ceRNA network based on DEGs, DEMs and DELs.

The main goal was to understand the complex and fine-tuning regulation of gene expression, the focus of the analysis was on

3. **Identification of key disturbed molecular mechanisms and pathways in HGSOC and its regulation** was possible by functional analysis of ceRNA-related mRNAs and ceRNA network-related TFs, including hub regulators of gene expression in HGSOC. The ceRNA network allowed to identify key pathways, including the PI3K pathway, in HGSOC, which were found to be disrupted in many layers of transcriptional regulation. The PI3K pathway has been previously described in the development of drug resistance.

Finding therapeutics which can restore the “healthy” profile of the key disrupted genes, might then be helpful also in overcoming drug resistance. The last goal

4. **Search for potential therapeutics** was achieved using bioinformatics approaches

and literature evidence. CMap and L1000CDS2 tools enabled to identify candidate molecules for the treatment of the HGSOC, namely, PI-103 and ZSTK474, both PI3K inhibitors. The literature provides evidence supporting the effectiveness of PI-103 and ZSTK474 against OC. Additionally, the miRNAs and lncRNAs identified in this study

might be targeted for use in the precision treatment or might have potential applications as biomarkers for HGSOC.

This knowledge might help better understand cancer development and progression and influence the future of treatment development in HGSOC.

5.2. Outlook

Functional *in-vitro* studies of the potential therapeutics for HGSOC will be performed to validate the *in-silico* results.

6. Bibliography

1. Krämer, A., Green, J., Pollard, J. & Tugendreich, S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* **30**, 523–530 (2014).
2. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2016. *CA Cancer J Clin* **66**, 7–30 (2016).
3. Longuespée, R. *et al.* Ovarian cancer molecular pathology. *Cancer and Metastasis Reviews* **31**, 713–732 (2012).
4. Malvezzi, M., Carioli, G., Rodriguez, T., Negri, E. & la Vecchia, C. Global trends and predictions in ovarian cancer mortality. *Annals of Oncology* **27**, 2017–2025 (2016).
5. Jayson, G. C., Kohn, E. C., Kitchener, H. C. & Ledermann, J. A. Ovarian cancer. *The Lancet* **384**, 1376–1388 (2014).
6. Vaughan, S. *et al.* Rethinking ovarian cancer: recommendations for improving outcomes. **11**, (2011).
7. Bowtell, D. D. Rethinking ovarian cancer II: reducing mortality from high-grade serous ovarian cancer. *Nat Rev Cancer* **15**, 668–679 (2015).
8. med Radosław Mądry Katedra Klinika Onkologii, hab *et al.* Recommendation of the Polish Society of Oncological Gynaecology on the diagnosis and treatment of epithelial ovarian cancer. *Oncol Clin Pract* **11**, 233–243 (2015).
9. English, D. P., Menderes, G., Black, J., Schwab, C. L. & Santin, A. D. Molecular diagnosis and molecular profiling to detect treatment-resistant ovarian cancer. *Expert Review of Molecular Diagnostics* Preprint at https://doi.org/10.1080/14737159.2016.1188692 (2016).
10. Birrer, M. J. Ovarian Cancer: Targeting the Untargetable. *American Society of Clinical Oncology Educational Book* **34**, 13–15 (2014).
11. Hood, L. & Rowen, L. The human genome project: Big science transforms biology and medicine. *Genome Med* **5**, 1–8 (2013).
12. Niklinski, J. *et al.* Systematic biobanking, novel imaging techniques, and advanced molecular analysis for precise tumor diagnosis and therapy: The Polish MOBIT project. *Adv Med Sci* **62**, 405–413 (2017).
13. Cancer Genome Atlas Research Network, T. Integrated genomic analyses of ovarian carcinoma. (2012) doi:10.1038/nature10166.

14. Kanchi, K. L. *et al.* Integrated analysis of germline and somatic variants in ovarian cancer. *Nature Communications 2014 5:1* **5**, 1–14 (2014).
15. Lheureux, S., Gourley, C., Vergote, I. & Oza, A. M. Epithelial ovarian cancer. *The Lancet* **393**, 1240–1253 (2019).
16. Levanon, K., Crum, C. & Drapkin, R. New insights into the pathogenesis of serous ovarian cancer and its clinical impact. *J Clin Oncol* **26**, 5284–5293 (2008).
17. Zhang, Z. *et al.* Molecular Subtyping of Serous Ovarian Cancer Based on Multi-omics Data. (2016) doi:10.1038/srep26001.
18. Diederichs, S. *et al.* The dark matter of the cancer genome: aberrations in regulatory elements, untranslated regions, splice sites, non-coding RNA and synonymous mutations. *EMBO Mol Med* **8**, 442–457 (2016).
19. Barrett, L. W., Fletcher, S. & Wilton, S. D. Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. *Cellular and Molecular Life Sciences* **69**, 3613–3634 (2012).
20. Vucic, E. A. *et al.* Translating cancer 'omics' to improved outcomes. *Genome Res* **22**, 188–195 (2012).
21. Chan, I. S. & Ginsburg, G. S. Personalized Medicine: Progress and Promise. *Annu Rev Genomics Hum Genet* **12**, 217–244 (2011).
22. Salmena, L., Poliseno, L., Tay, Y., Kats, L. & Pandolfi, P. P. A ceRNA hypothesis: The rosetta stone of a hidden RNA language? *Cell* vol. 146 353–358 Preprint at https://doi.org/10.1016/j.cell.2011.07.014 (2011).
23. Liu, Y. *et al.* Competitive endogenous RNA is an intrinsic component of EMT regulatory circuits and modulates EMT. *Nat Commun* **10**, (2019).
24. Jeyapalan, Z. *et al.* Expression of CD44 3'-untranslated region regulates endogenous microRNA functions in tumorigenesis and angiogenesis. *Nucleic Acids Res* **39**, 3026–3041 (2011).
25. Poliseno, L. *et al.* A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **465**, 1033–1038 (2010).
26. Li, J., Shao, W. & Feng, H. MiR-542-3p, a microRNA targeting CDK14, suppresses cell proliferation, invasiveness, and tumorigenesis of epithelial ovarian cancer. *Biomedicine & Pharmacotherapy* **110**, 850–856 (2019).
27. Li, J., Shao, W. & Feng, H. MiR-542-3p, a microRNA targeting CDK14, suppresses cell proliferation, invasiveness, and tumorigenesis of epithelial ovarian cancer. *Biomed Pharmacother* **110**, 850–856 (2019).

28. Wang, T. *et al.* The expression of miRNAs is associated with tumour genome instability and predicts the outcome of ovarian cancer patients treated with platinum agents. *Sci Rep* (2017) doi:10.1038/s41598-017-12259-w.
29. Panoutsopoulou, K., Avgeris, M. & Scorilas, A. miRNA and long non-coding RNA: molecular function and clinical value in breast and ovarian cancers. *Expert Rev Mol Diagn* **0**, 14737159.2018.1538794 (2018).
30. Lambert, S. A. *et al.* The Human Transcription Factors. *Cell* vol. 172 650–665 Preprint at https://doi.org/10.1016/j.cell.2018.01.029 (2018).
31. Karreth, F. A., Tay, Y. & Pandolfi, P. P. Target competition: Transcription factors enter the limelight. *Genome Biol* **15**, 1–3 (2014).
32. Ala, U. *et al.* Integrated transcriptional and competitive endogenous RNA networks are cross-regulated in permissive molecular environments. *Proc Natl Acad Sci U S A* **110**, 7154–7159 (2013).
33. Erol, A., Niemira, M. & Krętowski, A. J. Novel Approaches in Ovarian Cancer Research against Heterogeneity, Late Diagnosis, Drug Resistance, and Transcoelomic Metastases. *Int J Mol Sci* **20**, (2019).
34. Friedman, A. A., Letai, A., Fisher, D. E. & Flaherty, K. T. Precision medicine for cancer with next-generation functional diagnostics. *Nat Rev Cancer* **15**, 747–756 (2015).
35. Agarwal, R. & Kaye, S. B. Ovarian cancer: Strategies for overcoming resistance to chemotherapy. *Nature Reviews Cancer* vol. 3 502–516 Preprint at https://doi.org/10.1038/nrc1123 (2003).
36. Dienstmann, R., Jang, I. S., Bot, B., Friend, S. & Guinney, J. Database of genomic biomarkers for cancer drugs and clinical targetability in solid tumors. *Cancer Discov* **5**, 118–123 (2015).
37. Strittmatter, S. M. Overcoming drug development bottlenecks with repurposing: Old drugs learn new tricks. *Nature Medicine* vol. 20 590–591 Preprint at https://doi.org/10.1038/nm.3595 (2014).
38. Pushpakom, S. *et al.* Drug repurposing: Progress, challenges and recommendations. *Nature Reviews Drug Discovery* vol. 18 41–58 Preprint at https://doi.org/10.1038/nrd.2018.168 (2018).
39. Liu, T. P., Hsieh, Y. Y., Chou, C. J. & Yang, P. M. Systematic polypharmacology and drug repurposing via an integrated L1000-based Connectivity Map database mining. *R Soc Open Sci* **5**, (2018).

40. Lamb, J. The Connectivity Map: a new tool for biomedical research. *Nature Reviews Cancer 2007 7:1* **7**, 54–60 (2007).
41. Subramanian, A. *et al.* A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* **171**, 1437-1452.e17 (2017).
42. Hart, S. N., Therneau, T. M., Zhang, Y., Poland, G. A. & Kocher, J. P. Calculating sample size estimates for RNA sequencing data. *Journal of Computational Biology* **20**, 970–978 (2013).
43. Therneau T, Hart S, K. J. Calculating samples size estimates for RNA Seq studies. R package. Preprint at (2021).
44. Piovesan, A. *et al.* Human protein-coding genes and gene feature statistics in 2019. *BMC Res Notes* **12**, 1–5 (2019).
45. Derrien, T. *et al.* The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Res* **22**, 1775–1789 (2012).
46. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
47. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
48. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**, 1–9 (2010).
49. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. Voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* **15**, 29 (2014).
50. Smyth, G. K. *et al.* RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Res* **5**, (2018).
51. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* **17**, 10–12 (2011).
52. Lu, Y., Baras, A. S. & Halushka, M. K. miRge 2.0 for comprehensive analysis of microRNA sequencing data. *BMC Bioinformatics* **19**, 1–12 (2018).
53. Yan, X. *et al.* Comprehensive Genomic Characterization of Long Non-coding RNAs across Human Cancers. *Cancer Cell* **28**, 529–540 (2015).
54. Li, Q. *et al.* LncDIFF: A novel quasi-likelihood method for differential expression analysis of non-coding RNA. *BMC Genomics* **20**, 1–13 (2019).

55. Karagkouni, D. *et al.* DIANA-LncBase v3: Indexing experimentally supported miRNA targets on non-coding transcripts. *Nucleic Acids Res* **48**, D101–D110 (2020).
56. Huang, H. Y. *et al.* MiRTarBase 2020: Updates to the experimentally validated microRNA-target interaction database. *Nucleic Acids Res* **48**, D148–D154 (2020).
57. Shannon, P. *et al.* Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498–2504 (2003).
58. Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. A. & Luscombe, N. M. A census of human transcription factors: Function, expression and evolution. *Nature Reviews Genetics* vol. 10 252–263 Preprint at https://doi.org/10.1038/nrg2538 (2009).
59. Carro, M. S. *et al.* The transcriptional network for mesenchymal transformation of brain tumours. *Nature* **463**, 318–325 (2010).
60. Schmeier, S., Alam, T., Essack, M. & Bajic, V. B. TcoF-DB v2: Update of the database of human and mouse transcription co-factors and transcription factor interactions. *Nucleic Acids Res* **45**, D145–D150 (2017).
61. Lambert, S. A. *et al.* The Human Transcription Factors. *Cell* vol. 172 650–665 Preprint at https://doi.org/10.1016/j.cell.2018.01.029 (2018).
62. Hu, H. *et al.* AnimalTFDB 3.0: A comprehensive resource for annotation and prediction of animal transcription factors. *Nucleic Acids Res* **47**, D33–D38 (2019).
63. Garcia-Alonso, L., Holland, C. H., Ibrahim, M. M., Turei, D. & Saez-Rodriguez, J. Benchmark and integration of resources for the estimation of human transcription factor activities. *Genome Res* **29**, 1363–1375 (2019).
64. Schubert, M., Colomé-Tatché, M. & Foijer, F. Gene networks in cancer are biased by aneuploidies and sample impurities. *bioRxiv* 752816 Preprint at https://doi.org/10.1101/752816 (2019).
65. The Gene Ontology Consortium. The Gene Ontology resource: enriching a GOld mine. doi:10.1093/nar/gkaa1113.
66. Han, H. *et al.* TRRUST v2: An expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res* **46**, D380–D386 (2018).
67. Chin, C. H. *et al.* cytoHubba: Identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol* **8**, 1–7 (2014).
68. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).

69. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* vol. 28 27–30 Preprint at https://doi.org/10.1093/nar/28.1.27 (2000).
70. Jassal, B. *et al.* The reactome pathway knowledgebase. *Nucleic Acids Res* **48**, D498–D503 (2020).
71. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
72. Liberzon, A. *et al.* The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Syst* **1**, 417–425 (2015).
73. Luo, W. & Brouwer, C. Pathview: An R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics* **29**, 1830–1831 (2013).
74. Duan, Q. *et al.* L1000CDS2: LINCS L1000 characteristic direction signatures search engine. *NPJ Syst Biol Appl* **2**, 1–12 (2016).
75. Tagai, N., Goi, T., Shimada, M. & Kurebayashi, H. Plasma Prokineticin 1, a prognostic biomarker in colorectal cancer patients with curative resection: a retrospective cohort study. *World J Surg Oncol* **19**, 1–11 (2021).
76. Monnier, J. & Samson, M. Prokineticins in angiogenesis and cancer. *Cancer Lett* **296**, 144–149 (2010).
77. Zhang, X. Y. *et al.* Ovarian cancer derived PKR1 positive exosomes promote angiogenesis by promoting migration and tube formation in vitro. *Cell Biochem Funct* **39**, 308–316 (2021).
78. Wu, Y., Gao, J. & Liu, X. Deregulation of angiopoietin-like 4 slows ovarian cancer progression through vascular endothelial growth factor receptor 2 phosphorylation. *Cancer Cell Int* **21**, 1–15 (2021).
79. Carbone, C. *et al.* Molecular Sciences Angiopoietin-Like Proteins in Angiogenesis, Inflammation and Cancer. (2018) doi:10.3390/ijms19020431.
80. Pancho, A., Aerts, T., Mitsogiannis, M. D. & Seuntjens, E. Protocadherins at the Crossroad of Signaling Pathways. *Front Mol Neurosci* **13**, 117 (2020).
81. Manna, P. R. *et al.* Genomic Profiling of the Steroidogenic Acute Regulatory Protein in Breast Cancer: In Silico Assessments and a Mechanistic Perspective. *Cancers 2019, Vol. 11, Page 623* **11**, 623 (2019).
82. Monestier, O. & Blanquet, V. WFIKKN1 and WFIKKN2: "Companion" proteins regulating TGFB activity. *Cytokine Growth Factor Rev* **32**, 75–84 (2016).

83. Kondás, K., Szlaḿa, G., Nagy, A., Trexler, M. & Ṕatthy, L. Biological functions of the WAP domain-containing multidomain proteins WFIKKN1 and WFIKKN2. *Biochem Soc Trans* **39**, 1416–1420 (2011).
84. Munn, L. L. & Garkavtsev, I. SLPI: a new target for stopping metastasis. *Aging (Albany NY)* **10**, 13 (2018).
85. Zheng, D. *et al.* Secretory leukocyte protease inhibitor is a survival and proliferation factor for castration-resistant prostate cancer. *Oncogene 2016 35:36* **35**, 4807–4815 (2016).
86. Nugteren, S. *et al.* High expression of secretory leukocyte protease inhibitor (SLPI) in stage III micro-satellite stable colorectal cancer is associated with reduced disease recurrence. *Scientific Reports 2022 12:1* **12**, 1–10 (2022).
87. Lin, K. *et al.* Preclinical development of an anti-NaPi2b (SLC34A2) antibody-drug conjugate as a therapeutic for non-small cell lung and ovarian cancers. *Clinical Cancer Research* **21**, 5139–5150 (2015).
88. Charkhchi, P. *et al.* cancers CA125 and Ovarian Cancer: A Comprehensive Review. *Cancers (Basel)* **12**, 3730 (2020).
89. Honda, H., Pazin, M. J., D'Souza, T., Ji, H. & Morin, P. J. Regulation of the CLDN3 gene in ovarian cancer cells. *http://dx.doi.org/10.4161/cbt.6.11.4832* **6**, 1733–1742 (2007).
90. Heinzelmann-Schwarz, V. A. *et al.* Overexpression of the Cell Adhesion Molecules DDR1, Claudin 3, and Ep-CAM in Metaplastic Ovarian Epithelium and Ovarian Cancer. *Clinical Cancer Research* **10**, 4427–4436 (2004).
91. Huang, Y. H. *et al.* Claudin-3 gene silencing with siRNA suppresses ovarian tumor growth and metastasis. *Proc Natl Acad Sci U S A* **106**, 3426–3430 (2009).
92. Tight Junction Proteins Claudin-3 and Claudin-4 Are Frequently Overexpressed in Ovarian Cancer but Not in Ovarian Cystadenomas | Clinical Cancer Research | American Association for Cancer Research. https://aacrjournals.org/clincancerres/article/9/7/2567/203550/Tight-Junction-Proteins-Claudin-3-and-Claudin-4.
93. Suzuki, H. I., Katsura, A., Matsuyama, H. & Miyazono, K. MicroRNA regulons in tumor microenvironment. *Oncogene 2015 34:24* **34**, 3085–3094 (2014).
94. del Vecchio, F. *et al.* Long non-coding RNAs within the tumour microenvironment and their role in tumour-stroma cross-talk. *Cancer Lett* **421**, 94–102 (2018).

95. Esquela-Kerscher, A. & Slack, F. J. Oncomirs - MicroRNAs with a role in cancer. *Nature Reviews Cancer* vol. 6 259–269 Preprint at https://doi.org/10.1038/nrc1840 (2006).
96. Catalanotto, C. *et al.* MicroRNA in Control of Gene Expression: An Overview of Nuclear Functions. *Int J Mol Sci* **17**, 1712 (2016).
97. Cho, W. C. S. OncomiRs: The discovery and progress of microRNAs in cancers. *Molecular Cancer* vol. 6 1–7 Preprint at https://doi.org/10.1186/1476-4598-6-60 (2007).
98. Izumiya, M., Tsuchiya, N., Okamoto, K. & Nakagama, H. Systematic exploration of cancer-associated microRNA through functional screening assays. *Cancer Sci* **102**, 1615–1621 (2011).
99. Nishida, N. *et al.* MicroRNA-125a-5p Is an Independent Prognostic Factor in Gastric Cancer and Inhibits the Proliferation of Human Gastric Cancer Cells in Combination with Trastuzumab. *Clinical Cancer Research* **17**, 2725–2733 (2011).
100. Qian, Z., Gong, L., Mou, Y., Han, Y. & Zheng, S. MicroRNA-203a-3p is a candidate tumor suppressor that targets thrombospondin 2 in colorectal carcinoma. *Oncol Rep* **42**, 1825–1832 (2019).
101. Lin, Y. C. *et al.* Tumor suppressor miRNA-204-5p promotes apoptosis by targeting BCL2 in prostate cancer cells. *Asian J Surg* **40**, 396–406 (2017).
102. Hong, B. S. *et al.* Tumor Suppressor miRNA-204-5p Regulates Growth, Metastasis, and Immune Microenvironment Remodeling in Breast Cancer. *Cancer Res* **79**, 1520–1534 (2019).
103. Lang, Y. *et al.* MicroRNA-429 induces tumorigenesis of human non-small cell lung cancer cells and targets multiple tumor suppressor genes. *Biochem Biophys Res Commun* **450**, 154–159 (2014).
104. Xu, Y. *et al.* MicroRNA-218-5p inhibits cell growth and metastasis in cervical cancer via LYN /NF-κB signaling pathway. *Cancer Cell International 2018 18:1* **18**, 1–15 (2018).
105. Shinden, Y. *et al.* Molecular pathogenesis of breast cancer: impact of miR-99a-5p and miR-99a-3p regulation on oncogenic genes. *Journal of Human Genetics 2020 66:5* **66**, 519–534 (2020).
106. Chen, H. *et al.* MicroRNA-449a acts as a tumor suppressor in human bladder cancer through the regulation of pocket proteins. *Cancer Lett* **320**, 40–47 (2012).

107. Yoshimura, A. *et al.* Exosomal miR-99a-5p is elevated in sera of ovarian cancer patients and promotes cancer cell invasion by increasing fibronectin and vitronectin expression in neighboring peritoneal mesothelial cells. *BMC Cancer 2018 18:1* **18**, 1–13 (2018).
108. Srijyothi, L., Ponne, S., Prathama, T., Ashok, C. & Baluchamy, S. Roles of Non-Coding RNAs in Transcriptional Regulation. in *Transcriptional and Post-transcriptional Regulation* (InTech, 2018). doi:10.5772/intechopen.76125.
109. Kazemzadeh, M., Safaralizadeh, R. & Orang, A. V. LncRNAs: emerging players in gene regulation and disease pathogenesis. *J Genet* **94**, 771–784 (2015).
110. Liu, Y., Tao, Z., Qu, J., Zhou, X. & Zhang, C. Long non-coding RNA PCAT7 regulates ELF2 signaling through inhibition of miR-134-5p in nasopharyngeal carcinoma. *Biochem Biophys Res Commun* **491**, 374–381 (2017).
111. Carlevaro-Fita, J. *et al.* Cancer LncRNA Census reveals evidence for deep functional conservation of long noncoding RNAs in tumorigenesis. *Communications Biology 2020 3:1* **3**, 1–16 (2020).
112. Guo, X. *et al.* Long non-coding RNA-HAGLR suppressed tumor growth of lung adenocarcinoma through epigenetically silencing E2F1. *Exp Cell Res* **382**, 111461 (2019).
113. Sun, W. *et al.* Lnc HAGLR Promotes Colon Cancer Progression Through Sponging miR-185-5p and Activating CDK4 and CDK6 in vitro and in vivo. *Onco Targets Ther* **13**, 5913 (2020).
114. Huang, Y., Xiang, B., Liu, Y., Wang, Y. & Kan, H. LncRNA CDKN2B-AS1 promotes tumor growth and metastasis of human hepatocellular carcinoma by targeting let-7c-5p/NAP1L1 axis. *Cancer Lett* **437**, 56–66 (2018).
115. Saini, S. *et al.* miRNA-708 Control of CD44+ Prostate Cancer–Initiating Cells. *Cancer Res* **72**, 3618–3630 (2012).
116. Holubekova, V. *et al.* Epigenetic regulation by DNA methylation and miRNA molecules in cancer. *Future Oncol* 2217–2222 (2017) doi:10.2217/fon-2017-0363.
117. Wang, J. *et al.* Exosomal Delivery of AntagomiRs Targeting Viral and Cellular MicroRNAs Synergistically Inhibits Cancer Angiogenesis. *Mol Ther Nucleic Acids* **22**, 153–165 (2020).
118. Song, M. & Rossi, J. J. The anti-miR21 antagomir, a therapeutic tool for colorectal cancer, has a potential synergistic effect by perturbing an angiogenesis-associated miR30. *Front Genet* **4**, 301 (2014).

119. Krützfeldt, J. *et al.* Silencing of microRNAs in vivo with 'antagomirs'. *Nature 2005 438:7068* **438**, 685–689 (2005).
120. Abba, M. L. *et al.* MicroRNAs as novel targets and tools in cancer therapy. *Cancer Lett* **387**, 84–94 (2017).
121. Ebert, M. S. & Sharp, P. A. MicroRNA sponges: Progress and possibilities. *RNA* **16**, 2043 (2010).
122. Claesson-Welsh, L. & Welsh, M. VEGFA and tumour angiogenesis. *J Intern Med* **273**, 114–127 (2013).
123. Zhan, L. *et al.* E2F1: a promising regulator in ovarian carcinoma. *Tumor Biology* doi:10.1007/s13277-015-4770-7.
124. Farra, R., Dapas, B., Grassi, M., Benedetti, F. & Grassi, G. E2F1 as a molecular drug target in ovarian cancer. *Expert Opin Ther Targets* **23**, 161–164 (2019).
125. De Meyer, T. *et al.* E2Fs mediate a fundamental cell-cycle deregulation in high-grade serous ovarian carcinomas. *Journal of Pathology* **217**, 14–20 (2009).
126. Sang, X. *et al.* E2F-1 targets miR-519d to regulate the expression of the ras homolog gene family member C. *ncbi.nlm.nih.gov*.
127. Suzuki, A., Leland, P., Joshi, B. H. & Puri, R. K. Targeting of IL-4 and IL-13 receptors for cancer therapy. *Cytokine* **75**, 79–88 (2015).
128. Ferguson, L. P., Diaz, E. & Reya, T. The Role of the Microenvironment and Immune System in Regulating Stem Cell Fate in Cancer. *Trends Cancer* **7**, 624–634 (2021).
129. Ahmad, N. & Kumar, R. Steroid hormone receptors in cancer development: A target for cancer therapeutics. *Cancer Lett* **300**, 1–9 (2011).
130. Han, Z.-J., Feng, Y.-H., Gu, B.-H., Li, Y.-M. & Chen, H. The post-translational modification, SUMOylation, and cancer (Review). *Int J Oncol* **52**, 1081 (2018).
131. Kroonen, J. S. & Vertegaal, A. C. O. Targeting SUMO Signaling to Wrestle Cancer. *Trends Cancer* **7**, 496–510 (2021).
132. Fan, L. *et al.* Non-canonical signaling pathway of SNAI2 induces EMT in ovarian cancer cells by suppressing miR-222-3p transcription and upregulating PDCD10. *Theranostics* **10**, 5895 (2020).
133. Lu, Z.-Y. *et al.* SNAI1 overexpression induces stemness and promotes ovarian cancer cell invasion and metastasis. *Oncol Rep* **27**, 1587–1591 (2012).

134. Yang, D. *et al.* Integrated Analyses Identify a Master MicroRNA Regulatory Network for the Mesenchymal Subtype in Serous Ovarian Cancer. *Cancer Cell* **23**, 186–199 (2013).
135. Warren, C. F. A., Wong-Brown, M. W. & Bowden, N. A. BCL-2 family isoforms in apoptosis and cancer. *Cell Death & Disease 2019 10:3* **10**, 1–12 (2019).
136. Kong, X. *et al.* Recent Advances in Understanding FOXN3 in Breast Cancer, and Other Malignancies. *Front Oncol* **9**, 234 (2019).
137. Huang, S., Ernberg, I. & Kauffman, S. Cancer attractors: A systems view of tumors from a gene network dynamics and developmental perspective. *Semin Cell Dev Biol* **20**, 869–876 (2009).
138. Calvo, F. & Sahai, E. Cell communication networks in cancer invasion. *Curr Opin Cell Biol* **23**, 621–629 (2011).
139. Gschwind, A., Fischer, O. M. & Ullrich, A. The discovery of receptor tyrosine kinases: targets for cancer therapy. *Nature Reviews Cancer 2004 4:5* **4**, 361–370 (2004).
140. Huang, L. & Fu, L. Mechanisms of resistance to EGFR tyrosine kinase inhibitors. *Acta Pharm Sin B* **5**, 390 (2015).
141. Özvegy-Laczka, C., Cserepes, J., Elkind, N. B. & Sarkadi, B. Tyrosine kinase inhibitor resistance in cancer: role of ABC multidrug transporters. *Drug Resistance Updates* **8**, 15–26 (2005).
142. Tang, Y. *et al.* Downregulation of miR-133a-3p promotes prostate cancer bone metastasis via activating PI3K/AKT signaling. *Journal of Experimental & Clinical Cancer Research 2018 37:1* **37**, 1–16 (2018).
143. Ediriweera, M. K., Tennekoon, K. H. & Samarakoon, S. R. Role of the PI3K/AKT/mTOR signaling pathway in ovarian cancer: Biological and therapeutic significance. *Semin Cancer Biol* **59**, 147–160 (2019).
144. Osaki, M., Oshimura, M. & Ito, H. PI3K-Akt pathway: Its functions and alterations in human cancer. *Apoptosis* **9**, 667–676 (2004).
145. Liu, R. *et al.* PI3K/AKT pathway as a key link modulates the multidrug resistance of cancers. *Cell Death & Disease 2020 11:9* **11**, 1–12 (2020).
146. Pushpakom, S. *et al.* Drug repurposing: progress, challenges and recommendations. *Nature Reviews Drug Discovery 2018 18:1* **18**, 41–58 (2018).

147. Sanchez, V. E., Nichols, C., Kim, H. N., Gang, E. J. & Kim, Y.-M. Targeting PI3K Signaling in Acute Lymphoblastic Leukemia. *International Journal of Molecular Sciences 2019, Vol. 20, Page 412* **20**, 412 (2019).
148. Wang, D. *et al.* Effective use of PI3K inhibitor BKM120 and PARP inhibitor Olaparib to treat PIK3CA mutant ovarian cancer. *Oncotarget* **7**, 13153 (2016).
149. Konstantinopoulos, P. A. *et al.* Olaparib and α-specific PI3K inhibitor alpelisib for patients with epithelial ovarian cancer: a dose-escalation and dose-expansion phase 1b trial. *Lancet Oncol* **20**, 570–580 (2019).
150. Banerjee, S., Kaye, S. B. & Ashworth, A. Making the best of PARP inhibitors in ovarian cancer. *Nature Reviews Clinical Oncology 2010 7:9* **7**, 508–519 (2010).
151. Cai, Y. *et al.* Inhibition of PI3K/Akt/mTOR signaling pathway enhances the sensitivity of the SKOV3/DDP ovarian cancer cell line to cisplatin in vitro. *Chinese Journal of Cancer Research* **26**, 564 (2014).
152. Glaysher, S. *et al.* Targeting EGFR and PI3K pathways in ovarian cancer. *British Journal of Cancer 2013 109:7* **109**, 1786–1794 (2013).
153. Dan, S., Yoshimi, H., Okamura, M., Mukai, Y. & Yamori, T. Inhibition of PI3K by ZSTK474 suppressed tumor growth not via apoptosis but G0/G1 arrest. *Biochem Biophys Res Commun* **379**, 104–109 (2009).

7. Supplementary materials

S1. The full list of upstream regulators of competing endogenous RNA Network (ceRNA) involved genes identified by Ingenuity Pathway Analysis QIAGEN IPA (QIAGEN Inc., https://digitalinsights.qiagen.com/IPA).

STAT3	IL4	TNFAIP3	SNAI2
TGFB1	MEG3	BHLHE40	LOC105372576
F7	PKM	ETV1	let-7
CG	REL	COPA	FUS-DDIT3
CAMP	S100A8	TNFSF10	PRKCE
TP53	IL15	SMAD3	GDF2
SP1	ETV5	TNF	WNT5A
TMPRSS2-ERG	JUNB	CCL2	mir-515
MIF	ATG7	SNAI1	mir-204
RNA polymerase II	BMP7	IKBKB	UPF1
HOXA11-AS	P38 MAPK	CTLA4	PDCD1
HGF	POU5F1	Fcer1	TEAD2
TP63	RNF152	KDM6B	TEAD3
SHC1	IL26	miR-31-5p (and other	ERG
SELPLG	mir-196	miRNAs w/seed	HDAC3
CXCL12	CTSB	GGCAAGA)	IL25
FSH	ADORA2B	S100A9	SPARC
TWIST2	TP53BP2	DNMT3A	Tgf beta
WT1	IKZF3	PTTG1	MED16
TERT	RAB11FIP1	LATS2	SRC
CUL7	DCN	IL18	F2
RELA	ZIC3	CTNNB1	TLR6
mir-8	LILRB1	ERK	TXNIP
GLI1	EWSR1-FLI1	MRTFB	Lh
HDAC1	PPP2R5C	TCR	SOX2
PKNOX2	FCGR2A	HEIH	SMARCA4
PPID	TREM1	PRC2	ERK1/2
NFKB1	CCL5	IL11	KLF6
GSK3B	IL10	ITGAV	MAP3K7
EGFR	CD36	MMP9	IgG
GATA6	ZFAS1	mir-373	IL27
FOS	CCN6	PARP1	TEAD1
DANCR	STK40	FAS	miR-146a-5p (and
JARID2	TNFSF13	ACVR1C	other miRNAs w/seed
FCER1G	ITGAM	PHB	GAGAACU)
ACTN4	TNFSF13B	TBK1	BRD4
Pkc(s)	ITGAX	IL17A	IKBKE
IGF1	MTUS1	EZH2	FN1
KITLG	DEFB103A/DEFB103	EHF	KRAS
miR-218-5p (and other	B	NOTCH1	CAV1
miRNAs w/seed	LATS1	PADI2	HDAC2
UGUGCUU)	HRAS	CHUK	FGFR2
SAA1	IL6	IRAK4	EPCAM
Growth hormone	AURK	WNT7A	TEAD4
PRMT1	GNA12	LEPR	DUX4
JUN	ANLN	ITGB3	MAP2K1/2
IL1B	CHI3L1	CXCL8	STK11
Creb	PTGS2	IFNB1	EGR1
TIFA	Ctnna	IL1A	C5
ESRP2	Smad	OSCAR	INHBA
GNA13	SLC9A3R1	IL13	EPAS1
IQGAP1	VCAN	IFNG	Akt
HULC	ALB	PFKFB3	IGF1R

IL32	SH3BP2	miR-221-3p (and other miRNAs w/seed GCUACAU)	JINK1/2
DDX58	ITK		Rap1
CSF1	TARID		Cpla2
RAF1	TJP2	mir-302	IFNL2
OSM	CDC6	miR-130a-3p (and other miRNAs w/seed AGUGCAA)	GALNT6
Rsk	CDK11B		DGKH
Cbp/p300	NKRF		ZDHHC2
IRAK1/4	ADM	BGN	P2RY6
CRNDE	PDGFC	CCL21	CCND1
TCEAL7	MAP4	MBD2	TNFRSF1B
NUDT6	KRT7-AS	HIF1A-AS1	KL
EIF4A3	Hif	UGDH	NCOR1
THRAP3	MUC2	SIX1	COL2A1
LINC00261	Hif1	PLAU	COL3A1
ADGRG1	NFkB (complex)	EIF4E2	IL1R1
HLA-DR	Interferon alpha	CD74	NDUFA13
SEMA4A	DICER1	GATAD2B	TFAP4
ERBB	MYD88	NLRC4	TIMP2
PLXND1	ARID1A	PROK1	IRF2
EGLN2	TP53COR1	CANX	PSMB8
AFAP1L2	SERPINA1	MAD2L1	FBLN2
Cebp	CREB1	PAEP	FLOT2
EXOC8	RB1	FABP5	EIF3E
PLEC	Nr1h	MC1R	mir-503
GFPT1	IL33	GPR32	mir-154
RCAN1	MITF	LSINCT5	NCOA1
DAB2	AHR	MVP	AXL
mir-221	IRF3	CCAT2	PTGER2
miR-494-3p (miRNAs w/seed GAAACAU)	TGFB2	SRSF6	UBE3A
	COL18A1	IRAK2	HSPB3
miR-29b-3p (and other miRNAs w/seed AGCACCA)	EGF	PTGER1	FGFR3
	HIF1A	BCL10	SRSF3
	SMAD2	CXCR2	PLA2G4A
miR-4651 (and other miRNAs w/seed GGGGUGG)	IL3	PARK7	GPI
	RBPJ	CCR2	MAP2K2
AQUAPORIN	Steroid 5 alpha-Reductase	COL17A1	NRP2
		SRSF5	MGAT3
PIM	Hsp70	CLEC2D	LOX
ZNRD1ASP	C9orf72	SRSF2	HSPB6
EFEMP1	PIP5K1A	PELP1	RARG
REG1A	LAPTM4B	INS	CTBP1
P2RY2	OPN3	NR3C1	HSPB8
TPSD1	MALAT1	PF4	SIN3B
PIAS3	Collagen Alpha1	FGF2	FPR2
FOLR1	C1QBP	IL2	GSK3A
ITGB6	CASP3	CD40	ANPEP
PAX4	ZFP36	SIRT1	CDK2
NLRP3	IDH1	FOXO1	FURIN
CLC	LAMA3	VEGFA	CD300C
TMEM9B	PLA2G10	CD28	BHLHE41
VCP	CTBP2	TLR3	WTAP
CDX1	GDF15	PI3K (family)	Notch
RALA		ANTXR2	NFkB (family)

Tnf receptor	NEDD9	ZNF395	CBL
Fibrinogen	SUPT20H	Pka catalytic subunit	HEY1
Aldose Reductase	IL36A	HLA-DQ	VTN
HOTAIR	NET1	RUNX3	CUX1
NR2F1-AS1	YAP/TAZ	RNF40	IL17F
IL1RL2	LOC105374325	SLC7A11-AS1	CSNK2A1
CACNA1C	KDM5B	PRTN3	CEBPB
GRHL2	TLR7/8	MED1	
SMAD1/5	Ctbp	PLAUR	
estrogen receptor	Vegf	CRHR1	
SCAVENGER	E2f	SRSF1	
receptor CLASS A	HDL	IL6R	
CASP1	BCAR4	MECP2	
TIRAP	SPINDOC	SLIT2	
SMPD2	KIF26A	TAC1	
THPO	CIP2A	THBS1	
AGR2	NCOR2	WAC	
BMP15	HSPB2	IL22RA1	
EZR	TNFRSF18	CRHR2	
CTNND1	mir-322	BIN1	
DAB2IP	miR-27a-3p (and other	EIF2AK2	
mir-9	miRNAs w/seed	ILK	
mir-27	UCACAGU)	PIK3CD	
miR-142-3p (and other	miR-181a-5p (and	CD40LG	
miRNAs w/seed	other miRNAs w/seed	Fc gamma receptor	
GUAGUGU)	ACAUUCA)	IGF2	
F3	miR-125b-5p (and	COL1A1	
SPRY4-IT1	other miRNAs w/seed	MAVS	
GDF9	CCCUGAG)	TCIM	
GJA1	miR-21-5p (and other	CSNK2B	
SUZ12	miRNAs w/seed	PPP1R1B	
IL1R2	AGCUUAU)	AURKA	
MERTK	STK4	AGT	
ZEB2	JUND	HAND1	
GH1	PTGER4	IL24	
KLF4	FCER2	LINC01234	
DUSP1	GHR	MAPK8	
BCL2	SKIL	PRKG1	
MAPKAPK2	TNFRSF12A	LILRB4	
OTUD7B	IKBKG	SH3KBP1	
FGFR1	MAPK7	IGFBP5	
H3-3A/H3-3B	RPS6KA3	HAVCR2	
CD163	CLEC4E	HFE	
IL18R1	NOX4	SPHK1	
RGCC	E2F1	IL1RAP	
INSR	SMAD4	PTH	
PLA2G2A	CLOCK	TMPO	
PTPN6	PI3K (complex)	DPP4	
RTKN	Gsk3	SCUBE3	
EDNRB	ADCY	THY1	
DNMT3B	C1q	HDAC4	
CD47	DPPA3	IRAK1	
NFKBIA	ZMYND10	SDCBP	

S2. The results of the top 50 results of L1000CDS2 for the High-Grade Serous Ovarian Cancer (HGSOC) competing endogenous RNA (ceRNA) network genes. The tool is based on the screening of signatures from the L1000 cost-effective high-throughput gene expression assay, which includes expression measurements of 978 genes. The L1000CDS2 calculates the overlap of gene-sets between input signatures and the perturbation signatures.

Rank	1-cos(α)	Perturbation	MOA
44.0	1.3824	PI 103 hydrochloride	MTOR inhibitor
2	1.4481	ZSTK-474	PI3K inhibitor
9	1.4238	ZSTK-474	PI3K inhibitor
20	1.4054	BRD-K19220233	JNK inhibitor
45	1.3812	BRD-K52911425	PI3K inhibitor
38	1.3872	BRD-K12867552	HDAC inhibitor
16	1.4135	GSK-1059615	PI3K inhibitor
31	1.3924	NU-7026	DNA-dependent protein kinase inhibitor
15	1.4160	NVP-AUY922	HSP inhibitor
34	1.3901	saracatinib	SRC inhibitor
17	1.4126	HG-6-64-01	RAF inhibitor
18	1.4122	HG-6-64-01	RAF inhibitor
29	1.3977	QL-X-138	MTOR inhibitor
19	1.4055	torin-2	MTOR inhibitor
43	1.3828	WZ-4-145	EGFR inhibitor
49	1.3764	BRD-A73909368	RNA polymerase inhibitor
21	1.4035	busulfan	DNA inhibitor
24	1.4001	trichostatin A	HDAC inhibitor
32	1.3919	trichostatin A	HDAC inhibitor
33	1.3915	trichostatin A	HDAC inhibitor
36	1.3880	trichostatin A	HDAC inhibitor
39	1.3864	trichostatin A	HDAC inhibitor
48	1.3781	trichostatin A	HDAC inhibitor
5	1.4421	spironolactone	Mineralocorticoid receptor antagonist
23	1.4007	manumycin A	Farnesyltransferase inhibitor
12	1.4210	demeclocycline	Bacterial 30S ribosomal subunit inhibitor
1	1.4903	TWS-119	Glycogen synthase kinase inhibitor
28	1.3982	TWS-119	Glycogen synthase kinase inhibitor
42	1.3828	BRD-K77947974	Dopamine receptor antagonist
22	1.4018	canertinib	EGFR inhibitor
41	1.3832	KU-55933	ATM kinase inhibitor
30	1.3958	celastrol	Anti-inflammatory
3	1.4471	BRD-K35920785	
4	1.4440	deprenalin	Monoamine oxidase inhibitor
6	1.4332	PFI-1	BET inhibitor
7	1.4314	598226	Antimicrobial and antifungal activity
8	1.4311	I-BET151	BET inhibitor
10	1.4235	(+)-JQ1	BET inhibitor
11	1.4217	NCGC00184834-01	
13	1.4191	3544 (taurine)	
14	1.4179	Rythmol (Propafenone)	Sodium channel inhibitor
25	1.3989	DCC-2036	Bcr-Abl inhibitor
26	1.3988	GSK-2126458 (Omipalisib)	PI3K inhibitor
27	1.3988	BRD-A59145032	
35	1.3890	P0030	
37	1.3879	GDC-0980	PI3K inhibitor
40	1.3838	S1170	NOS3 blocking peptide
46	1.3809	480743.cdx N-Cyclopropyl-5-(thiophen-2-yl)isoxazole-3-carboxamide	TGF-B, EMT, canonical and non-canonical Wnt activator
47	1.3805	BRD-K74767048	
50	1.3762	R3904 (Reversine)	Mps1 kinase inhibitor and SAC inhibitor

S3. The list of compounds identified by ConnectivityMap (CMap) for the High-Grade Serous Ovarian Cancer (HGSOC) competing endogenous RNA (ceRNA) Network genes with Connectivity scores lower than -90.

Score	ID	Name	MOA
-96.58	BRD-K05104363	PD-184352	MEK inhibitor
-96.57	BRD-K67868012	PI-103	MTOR inhibitor
-95.7	BRD-K63068307	ZSTK-474	PI3K inhibitor
-94.84	BRD-K89014967	AS-703026	MEK inhibitor
-94.67	BRD-K90733503	cefalexin	Bacterial cell wall synthesis inhibitor
-94.57	BRD-K12184916	dactolisib	MTOR inhibitor
-94.34	BRD-K51575138	TPCA-1	IKK inhibitor
-94.04	BRD-K57080016	selumetinib	MEK inhibitor
-93.04	BRD-K67566344	KU-0063794	MTOR inhibitor
-92.73	BRD-K69932463	AZD-8055	MTOR inhibitor
-90.78	BRD-K97365803	PI-828	PI3K inhibitor
-90.59	BRD-K76674262	homoharringtonine	Protein synthesis inhibitor
-90.1	BRD-K67352070	TC-2559	Acetylcholine receptor agonist