


RESEARCH

Open Access



Identification of a novel gene signature predicting response to first-line chemotherapy in BRCA wild-type high-grade serous ovarian cancer patients

Marianna Buttarelli^{1,2†}, Alessandra Ciucci^{1,2†}, Fernando Palluzzi³, Giuseppina Raspaglio^{1,2}, Claudia Marchetti^{2,4}, Emanuele Perrone⁴, Angelo Minucci⁵, Luciano Giacò³, Anna Fagotti^{2,4}, Giovanni Scambia^{2,4} and Daniela Gallo^{1,2*} 

Abstract

Background: High-grade serous ovarian cancer (HGSOC) has poor survival rates due to a combination of diagnosis at advanced stage and disease recurrence as a result of chemotherapy resistance. In BRCA1 (Breast Cancer gene 1) - or BRCA2-wild type (BRCAwt) HGSOC patients, resistance and progressive disease occur earlier and more often than in mutated BRCA. Identification of biomarkers helpful in predicting response to first-line chemotherapy is a challenge to improve BRCAwt HGSOC management.

Methods: To identify a gene signature that can predict response to first-line chemotherapy, pre-treatment tumor biopsies from a restricted cohort of BRCAwt HGSOC patients were profiled by RNA sequencing (RNA-Seq) technology. Patients were sub-grouped according to platinum-free interval (PFI), into sensitive (PFI > 12 months) and resistant (PFI < 6 months). The gene panel identified by RNA-seq analysis was then tested by high-throughput quantitative real-time PCR (HT RT-qPCR) in a validation cohort, and statistical/bioinformatic methods were used to identify eligible markers and to explore the relevant pathway/gene network enrichments of the identified gene set. Finally, a panel of primary HGSOC cell lines was exploited to uncover cell-autonomous mechanisms of resistance.

Results: RNA-seq identified a 42-gene panel discriminating sensitive and resistant BRCAwt HGSOC patients and pathway analysis pointed to the immune system as a possible driver of chemotherapy response. From the extended cohort analysis of the 42 DEGs (differentially expressed genes), a statistical approach combined with the random forest classifier model generated a ten-gene signature predictive of response to first-line chemotherapy. The ten-gene signature included: CKB (Creatine kinase B), CTNBL1 (Catenin, beta like 1), GNG11 (G protein subunit gamma 11), IGFBP7 (Insulin-like growth factor-binding protein 7), PLCG2 (Phospholipase C, gamma 2), RNF24 (Ring finger protein 24), SLC15A3 (Solute carrier family 15 member 3), TSPAN31 (Tetraspanin 31), TTI1 (TELO2 interacting protein 1) and UQCC1 (Ubiquinol-cytochrome c

*Correspondence: daniela.gallo@unicatt.it

†Marianna Buttarelli and Alessandra Ciucci contributed equally to this work.

¹ Unità di Medicina Traslationale per la Salute della Donna e del Bambino, Dipartimento Scienze della Salute della Donna, del Bambino e di Sanità Pubblica, Fondazione Policlinico Universitario A. Gemelli, IRCCS, Largo A. Gemelli 8, 00168 Rome, Italy

Full list of author information is available at the end of the article



reductase complex assembly factor). Cytotoxicity assays, combined with gene-expression analysis in primary HGSOc cell lines, allowed to define CTNBL1, RNF24, and TTI1 as cell-autonomous contributors to tumor resistance.

Conclusions: Using machine-learning techniques we have identified a gene signature that could predict response to first-line chemotherapy in BRCAwt HGSOc patients, providing a useful tool towards personalized treatment modalities.

Keywords: HGSOc, Drug-resistance, Patient stratification, Transcriptomic, Biomarkers, Bioinformatics, Random forest classifier model, Primary ovarian cancer cells

Background

Ovarian cancer is the second most common cause of gynecologic cancer death in women worldwide. Nearly 314,000 women were estimated to have been diagnosed with ovarian cancer and 207,000 to have died from disease in 2020, with rates varying across the world [1]. Over 90% of ovarian malignancies are categorized as epithelial ovarian cancers and HGSOc is estimated to be 50–60% of all ovarian malignancies [2]. HGSOc classically presents at advanced stage, and cytoreductive surgery and systemic platinum-taxane combination chemotherapy represent the standard of care for patients. However, most tumors eventually develop drug resistance, with a 5-year survival generally below 30% [3]. Chemosensitivity to platinum-based chemotherapy is a prognostic factor for patient survival and a significant determinant for subsequent treatment, including maintenance therapy with PARP inhibitors (PARPI). Indeed, the use of PARPI has markedly transformed the therapeutic armamentarium in HGSOc over the last 5 years, since women with HRD (homologous recombination deficiency), in particular BRCA1/2-mutated HGSOcs, greatly benefit from PARPI treatment [4]. Actually, PARPI approval has been granted for patients with ovarian cancer, both in the front line setting and as maintenance or treatment in recurrent platinum-sensitive disease regardless of HRD status [4, 5].

Notably, resistance (and progressive disease) occurs more often in BRCA1- or BRCA2wt than in mutated BRCA HGSOc patients. In platinum-resistant OC, the choice of therapy is challenging and its failure is eventually responsible for almost all deaths from this cancer [6]. Accordingly, advanced-stage, BRCA1/2wt HGSOc patients have a worse prognosis, with shorter PFS (progression free survival) than those with BRCA mutations [7]. In the era of personalized cancer therapy, validated predictive biomarkers do not exist for platinum sensitivity in women with BRCAwt ovarian cancers and predicting platinum sensitivity remains a challenge [8]. Despite various approaches have been developed, including mutational signature, transcriptomic signatures, tumor mutational burden, and functional biomarkers, no definitive results have been reached yet to guide precision

treatment for women with BRCAwt ovarian cancers [8, 9]. Actually, the elevated level of intra- and inter-tumor and inter-patient heterogeneity that characterized HGSOc has been shown to be present even prior to exposure to selective pressures of therapy, this having significantly hampered biomarker discovery into actionable clinical parameters [8].

The present study aimed at the identification of a biomarker signature that could predict response to first-line therapy in BRCAwt HGSOc patients. Patient phenotyping would allow patient allocation to personalized treatment procedures, with significant benefits to both patients and healthcare system.

Materials and methods

Study population

This was a retrospective study including 44 women admitted to the Gynecologic Oncology Unit, Fondazione Policlinico Universitario A. Gemelli IRCCS, Roma, from Apr 2014 to Sep 2018. Eligible patients were defined as having histologically confirmed diagnosis of HGSOc, BRCAwt status and pretreatment frozen biopsy available at XBIOGem-Biorep Gemelli Biobank. Pathology-based estimates of the percentage of tumor cells were evaluated from hematoxylin-eosin-stained slides. Samples included in the analysis had tumor purity higher than 70% (Fig. S1). Ethical approval for the study was obtained from Ethics Committee of Fondazione Policlinico Universitario A. Gemelli IRCCS, Roma (Prot. 2552/20 ID 2694).

Patients were treated either with primary cytoreductive surgery followed by 6 conventional cycles of chemotherapy (PDS) or with 3 or 4 cycles of neoadjuvant chemotherapy before surgery and 3 postoperative chemotherapy cycles (NACT + IDS). Patients received platinum-based chemotherapy; introduction of bevacizumab in combination with standard chemotherapy and as maintenance therapy was allowed following its approval in Italy (January, 2014). Response to chemotherapy and progression were defined according to Response Evaluation Criteria in Solid Tumors (RECIST) and Gynecologic Cancer Intergroup (GCIg) criteria [10, 11]. To define chemosensitivity, the common definition

of platinum resistance was used, identifying as “resistant” (PFI < 6 months) patients who relapsed less than 6 months after chemotherapy was stopped, or that progressed during therapy, and as “sensitive”, patients relapsing 6 months or more after chemotherapy [6]. In order to maximize the identification of potential differences in the biomarker profile associated with chemotherapy responsiveness, we decided to focus our analysis on patients with PFI < 6 months (R, resistant) and PFI > 12 months (S, sensitive).

Selected patients were divided in two cohorts. Targeted RNA sequencing was performed on samples from the discovery cohort, including 14 BRCAwt patients (i.e. 7S and 7R). This analysis allowed the identification of a restricted panel of 42 DEGs that could discriminate chemotherapy-sensitive from -resistant patients. This gene set was tested by HT RT-qPCR in an extended cohort including 25S and 19R patients (validation cohort). Patients' characteristics are reported in Table 1.

Generation of primary cancer cells from patient tissues

This study included 7 patients with newly diagnosed, histologically confirmed BRCAwt HGSOs, admitted to the Gynaecologic Oncology Unit, Fondazione Policlinico Universitario A. Gemelli, IRCCS, Roma, between May 2020 and February 2021. The study was approved by the local Ethics Committee and Institutional Review Board (Protocol 19402/18 ID: 2045). All data were managed using anonymous numerical codes (OV.GEM). Patients' characteristics are reported in Table S1.

Primary cancer cell culture and characterization

Samples for cell cultures were collected at surgery from “left over tissues” by using sterile scalpels. Tumour specimens were placed into sterile tubes containing culture medium [DMEM (Dulbecco's modified Eagle's medium) supplemented with 1% kanamycin] pre-warmed to 37°C. Thereafter, tissue samples placed in 60 mm Petri dishes containing PBS (Phosphate-buffered saline) with 10% fetal bovine serum (FBS), were finely minced by surgical blades into small fragments.

Tumour tissues were enzymatically digested using the Tumor Dissociation kit (Miltenyi Biotec Bologna Italy). According to manufacturer's protocol, fragments were incubated in 5 ml DMEM medium containing enzymes mix for 30 min at 37°C. The cell suspension was applied to a MACS SmartStrainer (70 µm), placed on a 50 mL tube and washed with 20 ml of PBS with 10% FBS, centrifuged and resuspended in complete DMEM/F12 (1:1) supplemented with heat inactivated foetal bovine serum (10%), glutamine (2 mM) and kanamycin (2 mM) (Life Technologies, CA, USA). Cells were maintained at 37°C in the presence of 5% CO₂ and 95% humidified air. Medium was changed 48/72 h after initial plating and every 3 days thereafter. Stromal cells, primarily fibroblasts, progressively disappeared during the subcultivation process. Each patient-derived cell line was characterized by morphology, immunocytochemistry, and immunofluorescence.

Morphological features were studied using an inverted microscope DM IL LED (Leica Microsystems, Milan,

Table 1 Clinicopathological characteristics of HGSO patient cohort

Characteristics	Discovery cohort		Validation cohort	
	Sensitive No. (%)	Resistant No. (%)	Sensitive No. (%)	Resistant No. (%)
All cases	7	7	25	19
Median Age, years (range)	55 (48–65)	62 (47–71)	59 (37–73)	59 (32–79)
FIGO Stage				
III	6 (85.7)	6 (85.7)	22 (88.0)	14 (73.7)
IV	1 (14.3)	1 (14.3)	3 (12.0)	5 (26.3)
Primary treatment strategy				
PDS	7 (100)	6 (85.7)	20 (80.0)	8 (42.1)
NACT+IDS	–	–	5 (20.0)	9 (47.4)
NACT	–	1 (14.3)	–	2 (10.5)
Primary chemotherapy				
Platinum/Paclitaxel	2 (28.6)	4 (57.1)	11 (44.0)	11 (57.9)
Platinum/Paclitaxel/Bevacizumab	5 (71.4)	3 (42.9)	14 (56.0)	8 (42.1)
Residual tumor at surgery				
RT = 0	5 (71.4)	4 (57.1)	17 (68.0)	10 (52.6)
RT > 0	1 (14.3)	2 (28.6)	5 (20.0)	5 (26.3)
Not available/applicable	1 (14.3)	1 (14.3)	3 (12.0)	4 (21.1)

HGSO high-grade serous ovarian cancer, PDS primary debulking surgery, NACT Neoadjuvant chemotherapy, IDS interval debulking surgery

Italy). Cells cultures were then characterized using specific epithelial and stromal markers. To this end, cells were centrifuged at 600 rpm for 5 min at cytospin centrifuge, cytopspins were fixed with 4% paraformaldehyde for 20 min at RT and then permeabilized in 0.5% v/v Triton X-100 in PBS for 10 min. The endogenous peroxidase was blocked with 3% H₂O₂ for 5 min. After washing twice with PBS, cells were incubated with a blocking solution containing 20% normal horse serum in PBS for 30 min at RT. After incubation with the primary antibodies, i.e. anti-cytokeratin 7 (Clone OV-TL 12/30, Agilent Dako, Santa Clara, CA USA, ready-to-use) and anti-human fibroblasts (clone TE-7, Sigma Aldrich, Merck KGaA, Darmstadt, Germany dilution 1:100), for 1 h at RT in a humidified chamber, samples were rinsed with PBS and incubated with the secondary antibody, EnVision System-HRP (Dako), for 30 min at RT. Immunoreactivity was detected using the 3,3'-diaminobenzidine substrate (DAB substrate System, Agilent Dako). Slides were counterstained with Mayer's Haematoxylin, dehydrated in ethanol and xylene, and finally mounted.

Finally, expression of cytokeratin 7 was also confirmed by fluorescence microscopy. To this end, cells were seeded in 6-well plates containing coverslip in complete growth medium, fixed in 4% paraformaldehyde for 20 min at 20°C, and permeabilized in 0.5% v/v Triton X-100 in PBS for 10 min, prior to be blocked with 5% v/v goat serum and 0.1% v/v Triton X-100 in PBS for 1 h. Immunofluorescence staining was obtained using anti- anti-cytokeratin 7 (Clone OV-TL 12/30, Agilent Dako, ready-to-use) following overnight incubation at 4°C. After washing, cells were incubated with secondary antibody anti-mouse Alexa Fluor-488 conjugate (1:200) (Thermo Fisher Scientific Waltham, MA), in the dark for 30 min at 20°C. Coverslip was mounted onto slides using an antifade mounting reagent containing DAPI. Slides were observed under a fluorescence microscope (Leica Biosystems, Newcastle, UK) using a 100X oil immersion objective.

Established HGSOc cell lines

The human ovarian carcinoma cell line COV318 was obtained from the European Collection of Cell Cultures (ECACC, Salisbury, UK), while NIH:OVCAR-3 and OV-90 from the American Type Culture Collection (ATCC, Milan, Italy). Susan Horwitz (Albert Einstein Medical College) donated the HEY cell line. COV318 is a human ovarian epithelial-serous carcinoma cell line established from a peritoneal ascites [12]. The NIH:OVCAR-3 line was established from the malignant ascites of a chemotherapy-treated patient with progressive adenocarcinoma of the ovary [13]. OV-90 cells were derived from a chemotherapy-naive grade 3, stage IIIC,

malignant papillary serous adenocarcinoma [14]. HEY cell line was derived from a human ovarian cancer xenograft originally grown from a peritoneal deposit of a patient with moderately differentiated papillary cystadenocarcinoma of the ovary [15].

NIH:OVCAR-3 and HEY cells were cultured in RPMI 1640 (Roswell Park Memorial Institute Medium) and COV318 in DMEM (Sigma Aldrich). OV-90 were cultured in 1:1 mixture of MCDB 105 medium (containing a final concentration of 1.5 g/L sodium bicarbonate) and Medium 199 (containing a final concentration of 2.2 g/L sodium bicarbonate). Medium was supplemented with 10% FBS for COV318, NIH:OVCAR-3 and HEY, and 15% FBS for OV-90, plus MEM (Minimum Essential Medium) Non-Essential Amino Acid 1%, glutamine 1 mM and kanamycin 1%. Cells were grown in a fully humidified atmosphere of 5% CO₂/95% air, at 37°C. Cells were routinely tested for mycoplasma (MycoAlert mycoplasma detection kit, LONZA, Rockland, ME, USA) and validated by STR (Short Tandem Repeat) DNA profiling (BMR Genomics srl, Padua, Italy).

Cytotoxicity assay

Cells were seeded in 96-well plate overnight at a density of 5×10^3 per well, and then treated with different concentrations (0.001 to 100 μ M) of Paclitaxel or Cisplatin (both from Sigma Aldrich) for 72 h. After incubation, 10 μ L Cell Counting Kit-8 (Sigma Aldrich) was added into each well for 1 h. The OD (optical density) of each sample was detected at 450 nm using a microplate reader (Enspire, Perkin Elmer, Waltham, MA, USA). Each treatment was performed in triplicate. The mean optical density in the indicated concentrations was used to calculate the percentage of cell viability.

Isolation of RNA

Total RNA was isolated from HGSOc tumors and primary cells using the AllPrep DNA/RNA/Protein Mini Kit (Qiagen, Milan, Italy) according to the manufacturer's protocol and stored at -80°C until analyzed. Recovered RNA concentration and quality were measured using the Nanodrop (Thermo Fisher Scientific) and Agilent 2100 Bioanalyzer (Agilent Technologies), respectively.

RNA-seq analysis

A targeted RNA sequencing for gene expression and transcriptome analysis was performed using the Ion Ampliseq Transcriptome Human Gene Expression kit (Thermo Fisher Scientific) on Ion S5™ System (Thermo Fisher Scientific), on a discovery cohort (i.e. 7S and 7R BRCAwt HGSOc patients). This kit has been designed to profile over 20,000 distinct human RNA targets using a highly multiplexed amplification method. Briefly, cDNA

was first generated using the SuperScript® VILO™ cDNA Synthesis kit from 10 ng of total RNA obtained from tumor samples. Library preparation was carried out using an Ion Chef System (Thermo Fisher Scientific) according to the manufacturer's instructions. Briefly, barcoded libraries were generated from cDNA using an Ion AmpliSeq Chef Solutions DL8 Kit (Thermo Fisher Scientific) and an Ion AmpliSeq Transcriptome Human Gene Expression kit (Thermo Fisher Scientific). The Ion AmpliSeq™ technology accurately maintains expression levels of all targeted genes as previously described [16]. Amplified cDNA Libraries were quantified using the Qubit™ 3.0 Fluorometer (Thermo Fisher Scientific). Libraries were then diluted to 100 pM and pooled equally, with eight individual samples per pool. Pooled libraries were amplified using emulsion PCR on Ion Chef and enriched following the manufacturer's instructions. Template libraries were then sequenced on an Ion S5™ sequencing system, using Ion 540 Chip and an Ion 540 kit–Chef Kit (ThermoFisher Scientific).

Assessment of mRNA expression profiles using Fluidigm 48.48 dynamic arrays

The 42-gene panel identified by RNA-seq analysis was tested by HT RT-qPCR in an extended cohort including 25S and 19R patients (validation cohort). Analysis also included mRNAs isolated from primary HGSOC cells. Total RNA was tested using 48.48 dynamic array (Fluidigm Corporation, San Francisco, CA, USA) and a Biomark platform, following the manufacturer's protocol, as previously described [17]. The list of 42 genes evaluated in the study, along with primers used, is shown in Table S2. The geometric mean of RPL13A, RPL4 and GAPDH was taken as reference genes, following GeNorm and Normfinder algorithm [18, 19]. Data were analyzed as dCt values ($Ct_{\text{target}} - Ct_{\text{meanref}}$) according to $2^{-\Delta\Delta Ct}$ method [20].

RNA-seq data analysis

Primary analysis for AmpliSeq sequencing data of all samples was performed using the ampliSeqRNA plugin available for Ion Torrent Software Suite version 5.10.1. Raw reads were mapped to a custom reference sequence set (hg19 AmpliSeq Transcriptome v1.1) using the Torrent Mapping Alignment Program (TMAP) [16]. Raw read counts of the targeted genes were carried out using samtools (samtools view -c -F 4 -L bed_file bam_file), as part of ampliSeqRNA plugin. Differential expression analysis was performed using the DEBrowser R package [21], filtering out genes showing average raw counts less than 50. Batch effect correction was applied over samples coming from different runs, using the ComBat algorithm (the original algorithm is implemented in the

sva R package [22]). The DESeq2 algorithm [23] was then used to determine DEGs at a significance threshold of Benjamini-Hochberg (BH) adjusted P value < 0.05 [24]. Fold changes (FC) are reported as resistant over sensitive expression values. A further stringency threshold of BH-adjusted P -value $< 5e-03$ and absolute $\log_2 FC > 1$ lead to the selection of 42 top DEGs, used for further validation through HT RT-qPCR.

RNA-seq/HT RT-qPCR correlation

Expression data reproducibility was assessed by correlating standardized RNA-seq raw counts with HT RT-qPCR standardized $-\Delta Ct$ values. Since standardization does not affect correlation, the accordance between the two techniques was tested with a simple linear regression, using the lm function of the R base environment [25]. The estimated slope of the regression line was interpreted as Pearson correlation coefficient, and a P value < 0.05 was interpreted as a significant correlation between RNA-seq and HT RT-qPCR values. Separate regressions were done for S and R subjects, to highlight possible differences due to the response to the treatment.

Differential expression evaluation on HT RT-qPCR data

High-throughput RT-qPCR $-\Delta Ct$ values of the 42 selected DEGs, from 25S and 19R subjects, were compared using a Wilcoxon rank sum test (two-sided), implemented within the R base environment. The effect size was the estimated shift between negative ΔCt values in R versus S subjects, assuming a significance level of 0.05.

Random Forest Classifier (RFC) for prediction

To validate the significant DEGs found through the Wilcoxon rank sum test, a 4-fold cross-validation (CV) RFC was set to predict sensitive and resistant phenotypes, and to provide a rank of DEGs contribution to the classification process. During the training procedure, mean decrease accuracy (MDA), mean decrease Gini (MDG) impurity and Brier scores were computed. Brier scores were calculated for each subject and used as a measure of classification uncertainty. Subjects with a Brier score greater than 1 were discarded to improve both prediction accuracy and variable ranking. After removing high-uncertainty subjects, a new training process was started, and MDA and MDG were calculated, measuring the contribution that each variable has in the prediction accuracy and in the entropy of the classification process, respectively. Since MDA and MDG have different scales, their values were normalized from 0 to 100, where the worst predictor (gene) scores 0 and the best one scores 100. The ranking score for each gene is then the average of the percent MDA and MDG values.

The overall performance of the classification procedure was described in terms of Sensitivity, Specificity, Precision (or Positive Predictive Value, PPV), F1 score, Accuracy, False Positive rate (FPR), False Discovery Rate (FDR), and False Negative Rate (FNR). The method was implemented using the R packages *randomForest* [26] and *CMA* [27], for RFC building and Brier score calculation, respectively.

RICF-based heuristic analysis for primary cell lines validation

Gene deregulation trends were also validated on RT-qPCR data from primary cell lines (OV.GEM). The analysis was done separately upon two different treatments: Paclitaxel and Cisplatin. For each treatment, experimental data were used to establish sensitivity or resistance (outcome). Given the small number of statistical units (7 OV.GEM lines), a Residual Iterative Conditional Fitting (RICF)-based heuristic was applied to detect genes with significant differences in resistant versus sensitive cell lines. The RICEF method [28] achieves least squares convergence to exact maximum likelihood estimation of model parameters, in case of high dimensionality. Furthermore, the heuristic disables standard error estimation, and generates *P* values by fast randomization of group labels (resistant, sensitive). The entire procedure is implemented in the *SEMrun()* function of R package *SEMgraph* [29]. A significance level of 0.05 was set to determine gene regulation.

Over Representation Analysis (ORA)

To functionally annotate the 42 DEGs, we performed an ORA over the Kyoto Encyclopedia of Genes and Genomes (KEGG) database, using the Enrichr web-based tool [30]. Pathway enrichment was evaluated according to 3 criteria: (i) absolute number of genes from the input set belonging to a given pathway, (ii) two-sided Fisher's exact test Odds Ratio, 95% confidence intervals and *P*-value (including Benjamini-Hochberg adjusted *P*-value), to evaluate the association between the input gene set and the given pathway, (iii) a combined score derived from the product of the $-\log_{10}$ of the Fisher's exact test *P*-value and the z-score transformation of the Fisher statistic.

Figure creation

Unless stated differently, pictures have been generated using the R package *ggplot2* [31].

Analysis of publicly available datasets

In silico prediction of the possible genes association network for the 42 DEGs was performed using the GeneMANIA plugin in Cytoscape v.3.8.2, an integrated interaction

network program that predicts gene functions and possible interaction networks using many large publicly available datasets. Association data include protein and genetic interactions, pathways, co-expression, co-localization and protein domain similarity. For the estimation of the networks weight, the default network weighting method was chosen (Assigned based on query gene). This method maximize connectivity between all input genes. The GeneMANIA plugin in Cytoscape v.3.8.2 was used to generate the interaction networks [32].

cBioPortal (www.cbioportal.org) was used to access the publicly available Ovarian Serous Cystadenocarcinoma (TCGA, PanCancer Atlas) dataset and queried for co-expression of UQCC1 and CDR2 (Cerebellar Degeneration Related Protein 2), with setting using database default.

Statistics

Means and standard deviation (SD) were calculated for all data points, from at least two independent experiments. Dose-response curve-fit was calculated with v6.0 GraphPad Prism (GraphPad software, La Jolla, CA, USA).

Results

Identification of differentially expressed genes between sensitive and resistant HGSOC patients

To identify specific gene markers of therapy response, we compared pretreatment biopsies from 7 sensitive (S, PFI >12 months) and 7 resistant (R, PFI <6 months) patients, using a targeted whole transcriptome sequencing approach.

A total of 10,807 genes were found to be measurably expressed (raw read counts >50) in tumor samples (Table S3). Under a threshold of adjusted *P* value (padj) <0.05 and $|\log_2\text{FC}| > 1$, we identified 258 DEGs between S and R patients, including 55 upregulated and 203 downregulated genes (Fig. 1a, red and blue colored dots). After applying a more stringent filtering approach ($\text{padj} < 0.005$, $|\log_2\text{FC}| > 1$), 42 DEGs were found to be able to discriminate between the two different phenotypes, as shown by heatmap of sample-to-sample distance (Fig. 1b). Among the 42 DEGs, 17 were upregulated and 25 downregulated in R compared to the S patients. Magnitude of fold changes (log₂ scale) was within -2.51 to $+1.95$. Fold changes and molecular functions of the 42 DEGs are reported in Table S4.

Pathway enrichment analysis

To obtain further insight into the function of DEGs, they were uploaded to the KEGG database. As shown in Fig. 1c, KEGG pathway analysis found 4 significantly enriched pathways: Fc Epsilon RI signaling pathway (combined score 184.64; padj 0.027), Inositol phosphate

metabolism (combined score 166.85; padj 0.027), Phosphatidylinositol signaling system (combined score 110.64; padj 0.033) and Chemokine signaling pathway (combined score 80.53; padj 0.027), with the following genes involved: ALOX5AP (Arachidonate 5-lipoxygenase-activating protein), CXCL16 (C-X-C motif chemokine ligand 16), GNG11, INPP5D (Inositol Polyphosphate-5-Phosphatase D), PLCB1 (Phospholipase C beta 1), and PLCG2.

Gene-gene functional interaction network construction

We also used GeneMANIA plugin in Cytoscape v.3.8.2 to analyze the 42 DEGs with functionally similar genes together, to obtain regulatory networks. According to the annotation information recorded in GeneMANIA, a network involving the 42 DEGs and 20 related genes was established (Fig. 1d). Only physical interaction, pathway and co-expression were included in the network visualization. Notably, a strong physical interaction was evident between UQCC1 and CDR2 (Cerebellar degeneration-related protein 2) (Fig. 1d). We focused our attention on this finding, since CDR2 (normally present in cerebellar Purkinje neurons and brainstem neurons and testes) is a tumor antigen ectopically expressed in more than 50% of ovarian tumors [33]. We therefore sought to determine whether any evidence exists for the relationship between UQCC1 and CDR2 in serous ovarian cancer. Analyzing The Cancer Genome Atlas (TCGA) data via cBioPortal, we found that UQCC1 expression significantly ($P < 0.05$) inversely correlated with CDR2.

Validation of RNA-seq data by HT RT-qPCR

The 42 DEGs panel identified by RNA-seq was validated, by HT RT-qPCR, in an extended cohort of 25S and 19R patients (Validation cohort). Notably, analysis based on the Pearson's correlation coefficient indicated the positive correlation in gene expression measurements between RNA-seq and HT RT-qPCR (Fig. 2a), in both conditions (i.e., S and R patients).

Differentially expressed genes were identified via volcano plot, and genes with a P value less than 0.05 were filtered, this allowing the identification of ten genes significantly associated with chemotherapy response (Fig. 2b). The final gene panel included 4 downregulated (i.e. GNG11, IGFBP7, PLCG2 and SLC15A3) and 6 upregulated (i.e. CKB, CTNNB1, RNF24, TSPAN31, TTI1 and UQCC1) genes (Table 2). The random forest classification model based on the ten-gene signature accurately differentiated S and R patients. Specifically, the prediction performance of Random Forest classifier, using the ten-gene expression data, showed classification accuracy of about 93% and precision of 94% (Table 3). These significant results highlight the importance of considering multigene interactions for the generation of a signature, which may be achieved through machine-learning pattern recognition.

Epithelial markers of sensitivity to chemotherapy

It is well known that tumor microenvironment may considerably shape responses to drug treatment, a condition that could partially explain discrepancies between drug sensitivity in vitro and in vivo [34]. In the light of this, results from transcriptomic analysis on surgical resections might have captured also molecular features of context-dependent therapeutic resistance. Therefore, we used in vitro cellular models to reveal, among the predictive biomarkers identified, those genes possibly involved in cell-autonomous mechanisms of resistance [i.e. independent from the tumor microenvironment (TME)]. To this end, ex vivo cultures from tumor tissues of treatment-naïve patients were established. Before testing, cellular morphology and cell-type protein expression analysis were assessed to determine the purity of cells following isolation and culturing. Selected cell lines had typical primary epithelial culture morphology and their epithelial nature was confirmed by cytokeratin 7 staining (Fig. 3a). The absence of stromal fibroblast overgrowth

(See figure on next page.)

Fig. 1 **a** Volcano plot of the RNA-seq data from high grade serous ovarian cancer (HGSOC) patients ($n = 14$, 7 sensitive and 7 resistant). The plot reports the negative \log_{10} of the Benjamini-Hochberg adjusted P -value against the \log_2 fold-change (resistant vs sensitive) of the RNA-seq expression counts. The significance level for gene regulation is set to 0.05. Each gene is a dot, highlighted in blue for down-regulated genes, red for up-regulated, and grey for not-regulated ones. **b** Hierarchical clustering of the top 42 DEGs (Differentially expressed genes). Hierarchical agglomerative clustering based on Euclidean distance of resistant and sensitive subjects, calculated on RNA-seq standardized raw counts, over the top-42 DEGs (RNA-seq BH-adjusted P -value $< 5E-03$ and $|\log_2FC| > 1$). **c** Barplot of the KEGG enrichment analysis. Barplot showing the enrichment score (x axis) and supporting genes (labels over the bars) of the 4 enriched KEGG pathways among the 42 RNA-seq top-DEGs. Enrichment data was extracted from the output of the online tool Enrichr. **d** GeneMania interaction networks for the 42 DEGs. The network was spatially represented using the Cytoscape degree sorted circle layout, in which all nodes with the same numbers of links are located together around the circle. The circle is composed of the 42 query genes (except the lncRNA LOC100133985) found in our study, whereas the genes inside are the result genes identified by GeneMANIA. The colour of the line connecting the genes indicates the type of communication. Our diagram encompasses co-expression in grey lines, physical interactions in red lines, pathway in green lines. The thicknesses of the links (or edges) between the genes are proportional to the interaction strengths, whereas the node size for each gene is proportional to the rank of the gene based on its importance (or gene score) within the network

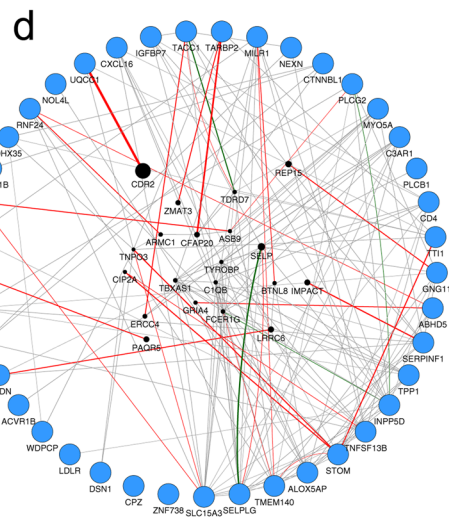
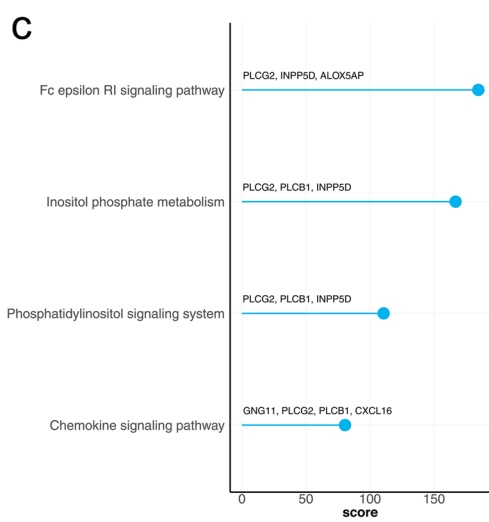
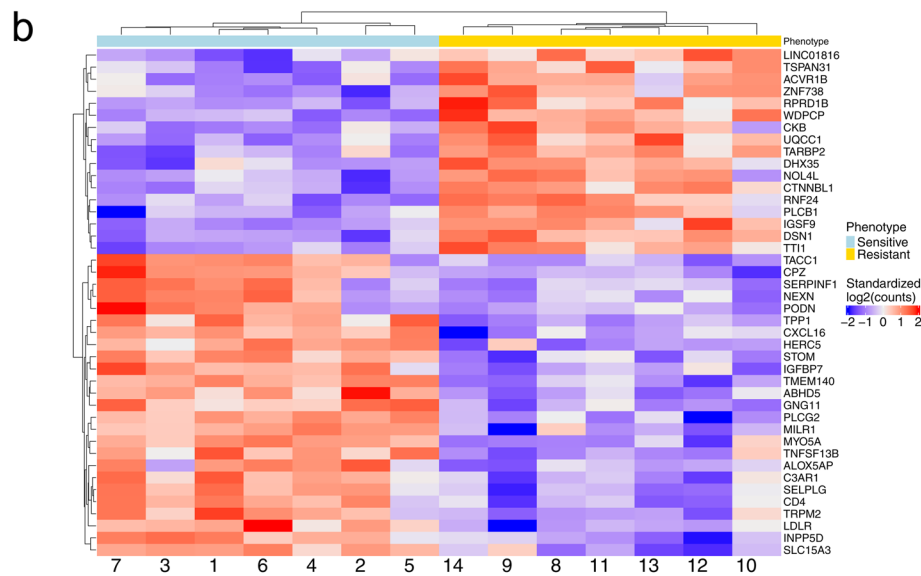
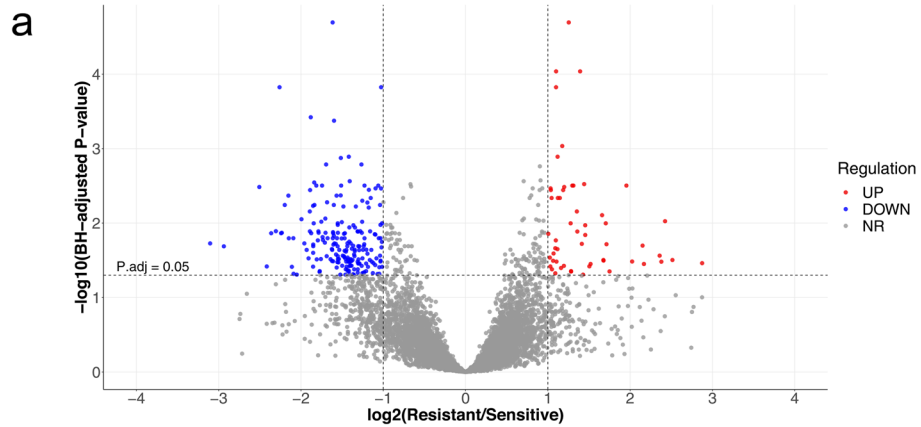


Fig. 1 (See legend on previous page.)

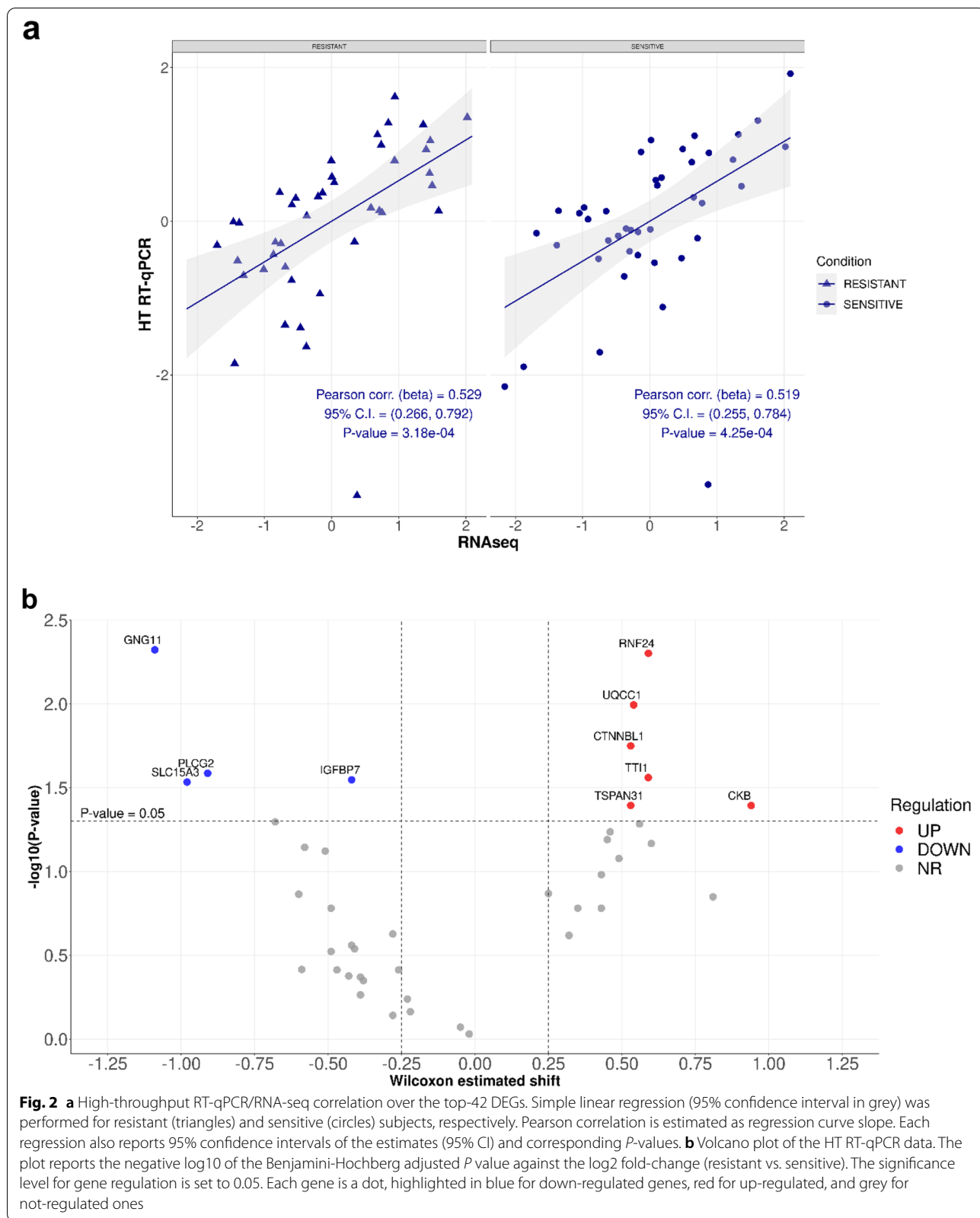


Table 2 Ranking using Random Forest Classifier (RFC) of DEGs identified by Wilcoxon test

Gene symbol	MDA	MDG	MDA (%)	MDG (%)	Mean Percent	Wilcoxon P-value	Estimated shift	95% Conf. Interval
GNG11	20.13	2.25	96.49	100.00	98.25	0.0048	-1.09	-1.71, -0.34
SLC15A3	20.70	1.55	100.00	48.86	74.43	0.0293	-0.98	-1.75, -0.17
PLCG2	20.24	1.57	97.16	50.18	73.67	0.0259	-0.91	-1.70, -0.13
IGFBP7	13.04	1.50	52.42	45.17	48.80	0.0284	-0.42	-0.89, -0.06
CKB	12.57	1.46	49.47	41.73	45.60	0.0404	0.94	0.05, 1.79
RNF24	8.81	1.61	26.12	53.02	39.57	0.0050	0.59	0.20, 1.12
CTNNBL1	9.07	1.47	27.75	43.11	35.43	0.0178	0.53	0.05, 1.03
UQCC1	9.65	1.41	31.33	38.07	34.70	0.0101	0.54	0.15, 0.93
TSPAN31	4.60	1.51	0.00	45.95	22.97	0.0404	0.53	0.03, 0.98
TTI1	5.72	0.89	6.94	0.00	3.47	0.0275	0.59	0.06, 1.21

DEGs Differentially expressed genes, MDA Mean Decrease Accuracy, MDG Mean Decrease Gini impurity

Table 3 Random Forest Classifier (RFC) performances of 10-gene signature

Output	10 DEGs ^a
True Positive	15
True Negative	24
False Positive	1
False negative	2
Sensitivity	88.24
Specificity	96.00
Precision	93.75
F1	90.91
Accuracy	92.86
False Positive Rate	0.0400
False Discovery Rate	0.0625
False Negative Rate	0.0513

^a Two subjects were excluded, given their Brier score > 1

in the primary cell lines generated was verified by TE-7 staining (Fig. S2a).

We then evaluated the cytotoxicity of cisplatin and paclitaxel on a panel of seven different primary cell lines and calculated the half maximal inhibitory concentration (IC₅₀) for each drug/cell line (Table S5). Cells were classified as resistant to either cisplatin or paclitaxel

when the IC₅₀ value was higher than the concentration achievable in patient plasma (concentration steady state/maximum concentration, i.e. C_{ss}/C_{max}), and as sensitive when the IC₅₀ value was lower than C_{ss}/C_{max}, as previously reported [35, 36]. Results obtained showed that two out of the seven primary cell lines were resistant to both drugs (i.e. OV.GEM-8 and OV.GEM-16) (Fig. 3b, Table S5). Notably, at least for OV.GEM-8, drug response of the cell line correlated with patient outcome (Table S1). For the purpose of our analysis, OV.GEM-8 and OV.GEM-16 were thus considered as resistant models, while the remaining five primary cell lines as sensitive models. We next evaluated in these cell lines the expression of the ten DEGs identified as associated to drug response in patients. Interestingly, as depicted in the Polar chart, when considering the specific gene deregulation associated to drug sensitivity, CTNNBL1, RNF24 and TTI1 showed in primary HGSOc cell lines the same trend observed in tumor tissue from HGSOc patients (Fig. 3c). Data obtained give support to the hypothesis that these three genes, may represent tumor cell-derived resistance factors.

We also tested, for comparison, the cytotoxicity of cisplatin and paclitaxel on a panel of established cell lines, selected among those considered as really representative of HGSOc [15, 37]. Representative pictures illustrating

(See figure on next page.)

Fig. 3 a Morphology of primary HGSOc cell lines (scale bars 200 μm) and representative images of cytokeratin 7 immunohistochemical and immunofluorescence staining (scale bars 100 and 20 μm). **b** Waterfall plot for paclitaxel and cisplatin IC₅₀ values (extracted from dose-response curves) of primary and established HGSOc cell lines. The steady state (C_{ss}) or the maximum in vivo plasma concentrations are indicated by the solid line (see Table S5). Bars above the solid line represent the resistant samples and bars below represent the sensitive samples. **c** Polar bar plot of the 10 selected biomarkers. Ten biomarkers have been chosen from the intersection of the top-42 RNA-seq DEGs with BH-adjusted P-value less than 5E-03 and |log₂FC| > 1 and HT RT-qPCR two-sided Wilcoxon rank sum test P value < 0.05. For each of them, the bar colour shading corresponds to the estimated resistant – sensitive shift, and the bar height corresponds to the number of evidences supporting their regulation (1: patients only; 2: patients and OV.GEM cell lines). The black dots show the -log₁₀ of the Wilcoxon's P-value (the further the distance from the center, the stronger the significance). The red circle corresponds to a P value of 0.05

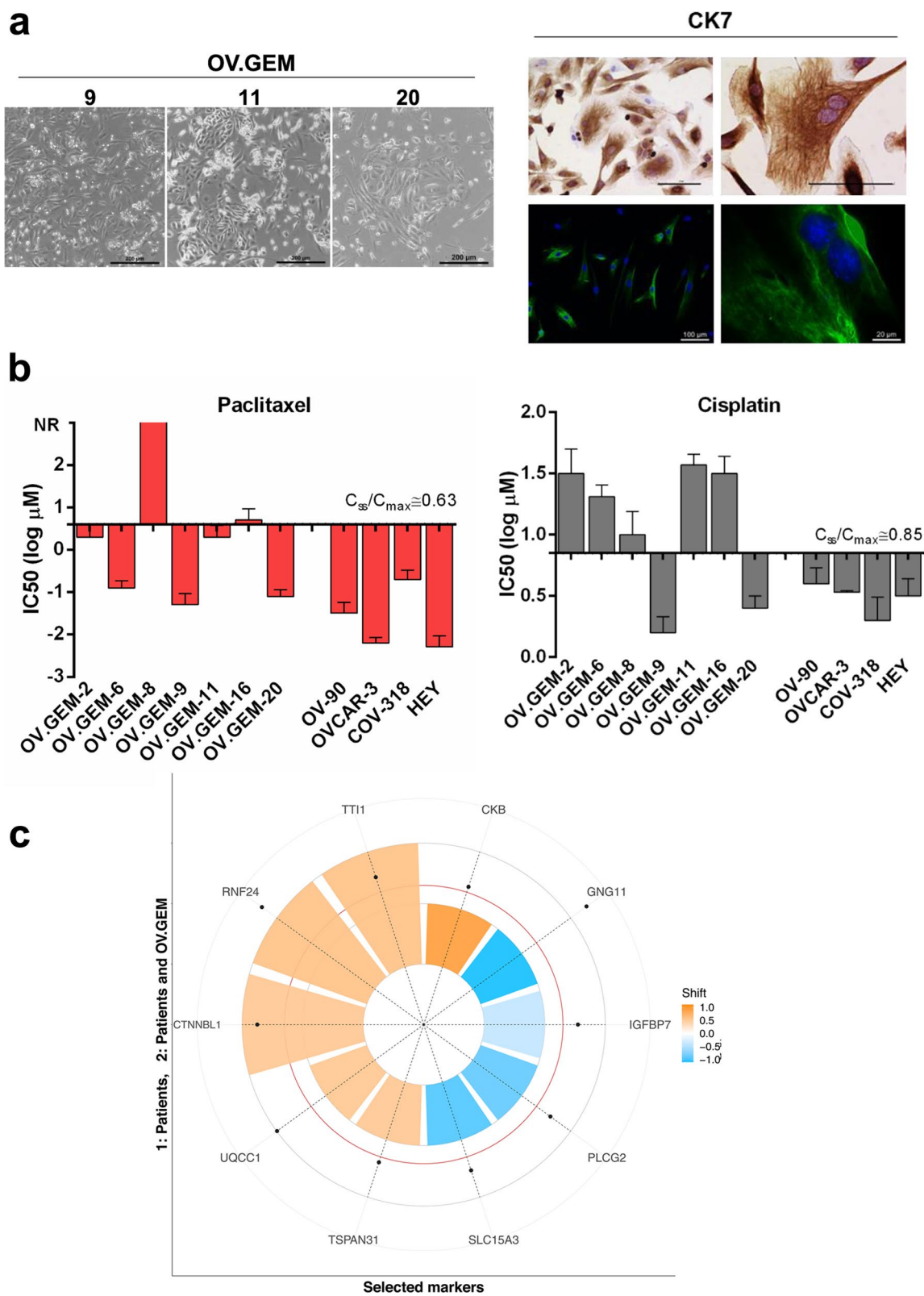


Fig. 3 (See legend on previous page.)

cell morphology and cytokeratin 7 staining of the tested cell lines are shown in Fig. S2b. Notably, all the investigated established cell lines showed IC50 values below the C_{50}/C_{max} and therefore no one could be identified as a resistant model according to the above specified criteria (Fig. 3b, Table S5). Our results confirm the low translational relevance of established cancer cell lines for the identification of clinical anti-cancer drug resistance [38].

Discussion

Here we have identified a ten-gene signature predictive of response to first-line chemotherapy in BRCAwt HGSOC patients. The signature includes downregulated (i.e. GNG11, IGFBP7, PLCG2, and SLC15A3) and upregulated (i.e. CKB, CTNNB1, RNF24, TSPAN31, TTI1 and UQCC1) genes. Importantly, our classification model achieved very high accuracy and precision, demonstrating its potential use as actionable clinical indicator.

Overall results obtained from the present study suggest that both cell-autonomous and TME-driven non-cell-autonomous mechanisms of drug resistance are critical in causing refractoriness of BRCAwt HGSOC patients and antineoplastic treatment failure (Fig. 4). Specifically, only three out of the ten DEGs identified in tumor biopsies were found to be similarly deregulated in primary HGSOC cells, i.e. CTNNB1, RNF24 and TTI1, this implicating a role for these genes in tumor cell exclusive pathways of chemotherapy resistance. For the remaining biomarkers identified, as discussed in detail below for each of them, our findings support a role in different signaling pathways possibly driven by interactions between tumor cells and components of the extracellular matrix. These findings are in line with recent evidences indicating that chemoresistance in HGSOC stems from complex, and not fully understood, mechanisms linked to pre-existing gene expression intrinsic within the chemotherapy-naive tumor cells or

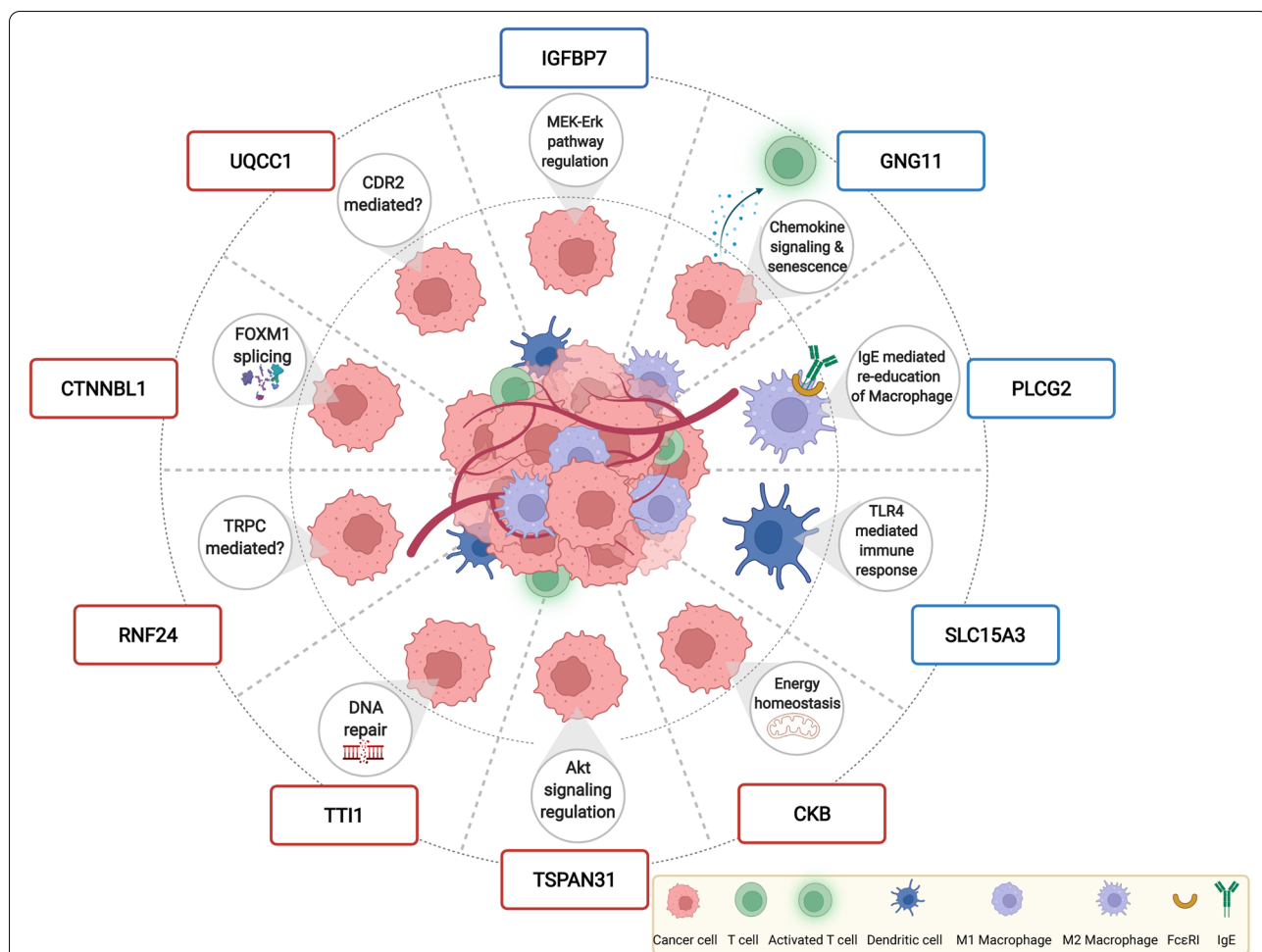


Fig. 4 Picture showing the potential role of the identified genes in the development of cell-autonomous and non-cell-autonomous resistance to antineoplastic agents, according to literature data. Created with BioRender.com

the immediate microenvironment (revised in 6, [39]). In this context it is worthy to note that the pre-existing state of the tumor immune microenvironment has been demonstrated to play a key role in inherent-drug resistance in HGSOs, infiltrated by a massive amounts of innate and adaptive immune cells [40].

Among the pathways we identified as significantly dysregulated in resistant patients, the Fc epsilon RI signaling (FcεRI) is of particular interest with regard to the TME-driven non-cell-autonomous resistance to antineoplastic agents. Indeed, recent fascinating hypotheses give support to the notion that, in tumor, IgE via interaction with its receptors (FcεRI and CD23) can engage and re-educate alternatively-activated macrophages (TAMs) towards pro-inflammatory phenotypes and prime all subsets to mediate anti-tumor functions, finally pointing to IgE as a novel anti-cancer modality [41–43]. Noteworthy, TAMs, the most abundant population of tumor-infiltrating immune cells in TME, with high density in HGSO [44, 45], are critical in fostering tumor progression and chemoresistance in ovarian cancer [46]. KEGG analysis revealed that the FcεRI pathway was enriched in downregulated genes including PLCG2, ALOX5AP and INPP5D. In particular, PLCG2, one of the gene in our signature, is a critical signaling molecule in both the innate and adaptive immune systems [47], being activated in monocytes, macrophages, and mast cells through crosslinking of the FcεR and FcγR [48]. Since macrophages actually represent the major FcεR-expressing effector cells, we could speculate that PLCG2 deficiency, by impairing the FcεR signaling in macrophages, will finally reduce the proinflammatory and anti-tumoral states induced by IgE, with a strong impact on response to therapy. Notably, bioinformatic analysis of publicly available ovarian cancer transcriptome datasets corroborates our hypothesis, by showing a favorable prognostic role of PLCG2 expression in advanced-stage HGSO (data not shown).

Consistent with a role of the immune system in driving chemosensitivity are also our findings showing a downregulation of SLC15A3 in R vs S patients. Indeed, SLC15A3 is a peptide/histidine transporter preferentially expressed in the immune cells, especially those of the innate immune system [49]. Despite the limited information available, SLC15A3 has been reported to be regulated by various Toll-like receptors (TLR), playing an important role in regulating TLR4-mediated inflammatory responses [50]. Relevant to our results, previous studies have demonstrated that TLR4-mediated immune response to dying tumor cells is essential for the efficacy of anticancer chemotherapy and radiotherapy; specifically, it has been reported that in the treatment of

established tumors with chemotherapy or radiotherapy the presence of TLR4 dictates the therapeutic outcome [51].

A dual role in mediating chemotherapy resistance is suggested by literature and current data for GNG11, a member of the heterotrimeric guanine nucleotide-binding protein (G protein) gamma family complex, playing a vital role in the transmembrane signaling system. GNG11 was the top down regulated gene in our analysis, and although very little is known about this gene, it has been reported to work as a novel type of regulator in cellular senescence following oxidative stress [52]. Crucially to our findings, incomplete and heterogeneous senescence response of cancer cells contributes to chemotherapy resistance [53]. Besides, in the KEGG pathway analysis, GNG11 was found enriched in chemokine signaling pathway (along with PLCG2, PLCB1 and CXCL16), and studies have uncovered the importance of G protein-coupled receptor (GPCR) for the maintenance of T cell trophic state, growth, and proliferation [54]. This is particularly significant when considering evidence indicating that the antitumor activity of cisplatin is not limited to its ability to inhibit mitosis, but it also has important immunomodulatory effects, including recruitment and proliferation of CD8+ T cells [55]. On the whole, therefore, findings available suggest that GNG11 could be involved in different signaling pathways within cancer cells and TME.

Finally, among the downregulated genes, IGFBP7 (also known as Mac25 or IGFBP-related protein 1) encodes a secreted IGFBP-related protein, with binding affinity to insulin growth factor (IGF). IGFBP7 also possesses an IGF-independent activity [56]. Notably, it has been proposed as a candidate tumor suppressor gene in several cancer types, including ovarian cancer [57], although conflicting results in different tumors suggest that it IGFBP7 functions may be a “double-edged sword” in cancer proliferation, progression, and prognosis. Relevant to our results, loss of IGFBP7 has been associated with chemo-resistance to cisplatin in human cancer cell lines and lung xenografts, possibly by the extracellular signal-regulated kinase (ERK) mitogen-activated protein kinase (MEK-ERK) pathway regulation [58].

Although limited information is available with regard to their specific function, the six upregulated genes in the signature (CKB, CTNBL1, TSPAN31, TTI1, RNF24 and UQCC1) have all been previously reported to be involved in the development of cancer and/or in mediating resistance to drug-induced cytotoxicity. In detail, an interesting correlate of drug response captured by our study was represented by KIAA0406/TTI1. Indeed, this protein has been identified as a critical component of DNA damage response (DDR), required for DNA damage resistance and whose depletion leads to cellular sensitivity to

radiation, as well as other genotoxic agents [59]. Results from the present study are fully coherent with this proposed function.

With regard to CTNNB1, UQCC1 and CKB, all of them have been previously associated to ovarian cancer. Specifically, the spliceosome-associated factor CTNNB1, has been related to proliferation and invasion in HGSOc, at least partially through regulation of FOXM1 splicing [60]. Notably, the FOXM1 signaling pathway is activated in 84% high-grade serous ovarian cancer [61] and linked to platinum resistance [62, 63]. Likewise, UQCC1, a protein involved in cytochrome b translation and/or stability, has been recently identified as a candidate ovarian cancer susceptibility gene [64]. Interestingly, it has been reported that UQCC1 physically interacts with CDR2 [65], a target antigen of naturally occurring human tumor immunity, widely expressed in the majority of ovarian cancer. Relevant to our results, recent studies have underscored the importance of CDR2 as target of an effective immune response in patients with paraneoplastic cerebellar degeneration (usually observed in gynecological cancer), pointing to its role in fueling immunogenic cell death (ICD) initiated by some chemotherapeutic agents [33, 66]. The inverse correlation between UQCC1 and CDR2 expression in ovarian cancer (as highlighted by analysis of TCGA data), gives support to our results. Finally, CKB gene expression has been reported to be up-regulated in ovarian cancer cells *in vitro* and *in vivo* and CKB enzyme activity to be significantly elevated in sera from ovarian cancer patients, including those with stage I disease [67]. Mechanistic studies have suggested that CKB is involved in the regulation of energy homeostasis and metabolic state in ovarian cancer [68].

Very little is known about the remaining two up-regulated genes in our analysis, i.e. TSPAN31 and RNF24. TSPAN31, recently discovered to be linked to cancer, as natural antisense transcript of cyclin dependent kinase 4 (CDK4), regulates the expression of CDK4 mRNA and protein [69]. Notably, TSPAN31 is also involved in the activation of Akt signaling pathway, a key regulator of survival in ovarian cancer, also associated with chemoresistance [70]. Lastly, RNF24 is a membrane protein which interacts with TRPC (transient receptor potential channel) proteins [71] and, although literature data have suggested that it may act as a cancer-promoting factor in different tumors [72, 73], its mechanistic role is still to be uncovered.

Besides patients' stratification, findings from the present study could also aid to define potential new treatment strategies. Here we used primary HGSOc cell lines as a faithful experimental model of disease to identify, among the panel of the ten predictive biomarker

genes, those candidates that we could investigate in *in vitro* experimental system, due to their possible cell-autonomous mechanisms of resistance. Interestingly, we showed that CTNNB1, RNF24 and TTI1 are associated to resistance to therapy in both primary HGSOc cell lines and patients' samples, in line with available literature data on the role of these proteins as tumor cell-derived resistance factors. On the other hand, established HGSOc cell lines, although easier to work with than primary cells, being genetically modified/transformed, may not fully represent the *in vivo* state and, therefore, may not be recommended for the identification of clinically relevant mechanisms of resistance to cytotoxic and targeted anti-cancer drugs.

Among the peculiar strengths of this study is that patients were treated in a single tertiary-care center, with accurately completed clinical records. In addition, all samples were managed and analyzed with standard protocols. Finally, we employed rigorous standardized normalization and statistical selection to overcome confounding factors and discrepancies resulting from differences in the experimental groups. Besides the above-mentioned strengths, few limitations in our study need to be acknowledged. First, the retrospective nature of this study makes it susceptible to bias. In addition, sample size was not large enough for detecting definitive associations between biomarkers expression and response to therapy. Therefore, validation of the gene signature in prospective trials is warranted. Besides, future studies are needed to pinpoint the mechanisms underpinning the role of the identified genes in drug-resistance and to investigate their potential druggability.

Conclusions

Predicting platinum sensitivity still remains a challenge in women affected by HGSOc. Our results show that the Random Forest Classification model based on the expression levels of a limited number of marker genes could accurately predict response to first-line chemotherapy in BRCAwt patients. Further validation of this profile will contribute to establishing auxiliary predictive markers for the therapy of HGSOc, defining better and more effective clinical treatments, finally helping the field moving toward a more personalized treatment approach.

Abbreviations

ATCC: American Type Culture Collection; BH: Benjamini-Hochberg; BRCA1: Breast Cancer gene 1; BRCA2: Breast Cancer gene 2; BRCAwt: Breast Cancer gene 1/2 wild type; CDK4: Cyclin dependent kinase 4; cDNA: Complementary DNA; CDR2: Cerebellar degeneration-related protein 2; CKB: Creatine kinase B; C_{ss}/C_{max}: Concentration steady state/maximum concentration; Ct: Cycle threshold; CTNNB1: Catenin beta like 1; CV: Cross-validation; DAB: 3,3'-diaminobenzidine substrate; DAPI: 4',6-diamidino-2-phenylindole; DDR: DNA damage response; DEGs: Differentially expressed genes; DMEM: Dulbecco's modified Eagle's medium; DNA: Deoxyribonucleic acid; ECACC

: European Collection of Cell Cultures; ECM: Extracellular matrix; ERK: Extracellular signal-regulated kinase; FBS: Fetal bovine serum; FC: Fold changes; FDR: False Discovery Rate; FNR: False Negative Rate; FOXM1: Forkhead box M1; FPR: False Positive rate; G protein: Guanine nucleotide-binding protein; GAPDH: Glyceraldehyde-3-phosphate dehydrogenase; GCG: Gynecologic Cancer Intergroup; GNG11: G protein subunit gamma 11; GPCR: G protein-coupled receptor; HGSOC: High-grade serous ovarian cancer; HRD: Homologous recombination deficiency; HT: High-throughput; IC50: Half maximal inhibitory concentration; ICD: Immunogenic cell death; IDS: Interval debulking surgery; IGFBP7: Insulin like growth factor binding protein 7; KEGG: Kyoto Encyclopedia of Genes and Genomes; MDA: Mean decrease accuracy; MDG: Mean decrease Gini; MEK-ERK: Mitogen-activated protein kinase; MEM: Minimum Essential Medium; NACT: Neoadjuvant chemotherapy; OD: Optical density; ORA: Over Representation Analysis; OV.GEM: Anonymous numerical codes for primary cell lines; padj: Adjusted *P* value; PARP: Poly (ADP-ribose) polymerase; PARPi: PARP inhibitors; PBS: Phosphate buffered saline; PDS: Primary cytoreductive surgery; PFI: Platinum-free interval; PLCG2: Phospholipase C gamma 2; PPV: Positive Predictive Value; R: Resistant; RECIST: Response Evaluation Criteria in Solid Tumors; RFC: Random Forest Classifier; RIFC: Residual Iterative Conditional Fitting; RNA: Ribonucleic acid; RNA-Seq: RNA sequencing; RNF24: Ring finger protein 24; RPL13A: Ribosomal protein L13a; RPL4: Ribosomal protein L4; RPMI: Roswell Park Memorial Institute Medium; RT: Room temperature; RT-qPCR: Quantitative real-time Polymerase chain reaction; S: Sensitive; SD: Standard deviation; SLC15A3: Solute carrier family 15 member 3; STR: Short Tandem Repeat; TAMs: Tumor-associated macrophages; TGCA: The Cancer Genome Atlas; TLR: Toll-like receptors; TMAP: Torrent Mapping Alignment Program; TME: Tumor Microenvironment; TRPC: Transient receptor potential channel; TSPAN31: Tetraspanin 31; TTI1: TELO2 interacting protein 1; UQC1: Ubiquinol-cytochrome c reductase complex assembly factor 1; WT: Wild type.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13046-022-02265-w>.

Additional file 1: Figure S1. Pathology-based estimates of the percentage of tumor cells were evaluated from hematoxylin-eosin-stained slides. Only samples with higher than 70% of tumor purity were analyzed. **Figure S2.** a) Representative pictures of primary HGSOC cells stained using cytokeratin 7 (CK7) and TE-7 antibodies. The negligible staining for TE-7 (arrowhead) confirmed the absence of stromal fibroblast overgrowth in the primary cell lines generated. b) Morphology of established HGSOC cell lines and representative images of CK7 immunohistochemical staining. **Table S1.** Clinicopathological features of HGSOCs used for primary cell line isolation. **Table S2.** List of primers DeltaGene (Fluidigm) used in RT-qPCR. **Table S4.** Top DEGs in resistant vs sensitive HGSOC patients (downregulated genes). **Table S5.** Top DEGs in resistant vs sensitive HGSOC patients (upregulated genes). **Table S5.** IC50 values of anticancer drugs against primary and commercially available HGSOC cell lines.

Additional file 2: Table S3. Containing RNA-seq raw data.

Acknowledgements

We thank Flavia Filippetti, Cristiana Gaggini and Enrica Martinelli for their skilful technical assistance. We are grateful, for his support in bioinformatics analysis, to our colleague Gabriele Babini, who prematurely passed away.

Authors' contributions

MB and AC designed the study. CM, EP, and AF recruited patients, acquired clinical samples and clinical data. MB, AC and GR mainly performed the experiments. FP and LG performed bioinformatics analysis. MB and AM analysed the data. MB and FP draft the manuscript. GS and AF supervised the study. GS was responsible for funding acquisition. DG conceived, designed and supervised the study, wrote and revised the manuscript. All authors read and approved the final manuscript.

Funding

This work, as part of the activities of the Unit of Translational Medicine for Woman and Child Health, was partially supported by the "Associazione OPPO e le sue stanze" Onlus.

Availability of data and materials

All data generated during this study are included in this published article [and its supplementary information files]. The datasets analysed during the current study are available in the CBioPortal repository (<https://www.cbioportal.org/>) (accessed on 15 September 2021) and GeneMANIA repository (<https://genemania.org>).

Declarations

Ethics approval and consent to participate

Clinical studies were approved by the local Ethics Committee and Institutional Review Board (Protocol 2552/20 ID 2694 and Protocol 19402/18 ID 2045) and were conducted in accordance with the European Good Clinical Practice requirements (Declaration of Helsinki). Written informed consent was obtained as appropriate.

Consent for publication

Not applicable.

Competing interests

The authors declare no conflicts of interest.

Author details

¹Unità di Medicina Traslazionale per la Salute della Donna e del Bambino, Dipartimento Scienze della Salute della Donna, del Bambino e di Sanità Pubblica, Fondazione Policlinico Universitario A. Gemelli, IRCCS, Largo A. Gemelli 8, 00168 Rome, Italy. ²Dipartimento Universitario Scienze della Vita e Sanità pubblica – Sezione di Ginecologia ed Ostetricia - Università Cattolica del Sacro Cuore, Largo A. Gemelli 8, 00168 Rome, Italy. ³Bioinformatics Facility Core Research, Gemelli Science and Technology Park (GSTeP) Fondazione Policlinico Universitario A. Gemelli, IRCCS, Roma, Italy. ⁴Dipartimento Scienze della Salute della Donna, del Bambino e di Sanità Pubblica, Fondazione Policlinico Universitario A. Gemelli, IRCCS, Roma, Italy. ⁵Molecular and Genomic Diagnostics Unit (MGDUnit), Fondazione Policlinico Universitario A. Gemelli IRCCS, Roma, Italy.

Received: 19 November 2021 Accepted: 18 January 2022

Published online: 04 February 2022

References

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71:209–49.
- Kim J, Park EY, Kim O, Schilder JM, Coey DM, Cho CH, et al. Cell origins of high-grade serous ovarian cancer. *Cancers*. 2018;10:433.
- Lheureux S, Gourley C, Vergote I, Oza AM. Epithelial ovarian cancer. *Lancet*. 2019;393:1240–53.
- Chiappa M, Guffanti F, Bertoni F, Colombo I, Damia G. Overcoming PARPi resistance: Preclinical and clinical evidence in ovarian cancer. *Drug Resist Updat*. 2021;55:100744.
- Konstantinopoulos PA, Lheureux S, Moore KN. PARP inhibitors for ovarian Cancer: current indications, future combinations, and novel assets in development to target DNA damage repair. *Am Soc Clin Oncol Educ Book*. 2020;40:1–16.
- Marchetti C, De Felice F, Romito A, Iacobelli V, Sassu CM, Corrado G, et al. Chemotherapy resistance in epithelial ovarian cancer: mechanisms and emerging treatments. *Semin Cancer Biol*. 2021;77:144–66.
- Kim SI, Lee M, Kim HS, Chung HH, Kim JW, Park NH, et al. Effect of BRCA mutational status on survival outcome in advanced-stage high-grade serous ovarian cancer. *J Ovarian Res*. 2019;12(1):40.
- Funingana IG, Reinius MAV, Petrillo A, Ang JE, Brenton JD. Can integrative biomarker approaches improve prediction of platinum and PARP inhibitor response in ovarian cancer? *Semin Cancer Biol*. 2021;77:67–82.
- Ivy SP, Kunos CA, Arnaldez FI, Kohn EC. Defining and targeting wild-type BRCA high-grade serous ovarian cancer: DNA repair and cell cycle checkpoints. *Expert Opin Investig Drugs*. 2019;28:771–85.
- Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, et al. New guidelines to evaluate the response to treatment in solid tumors. *J Natl Cancer Inst*. 2000;92:205–16.

11. Rustin GJS, Marples M, Nelstrop AE, Mahmoudi M, Meyer T. Use of CA-125 to define progression of ovarian cancer in patients with persistently elevated levels. *J Clin Oncol*. 2001;19:4054–7.
12. van den Berg-Bakker CA, Hagemeyer A, Franken-Postma EM, Smit VT, Kuppen PJ, van Ravenswaay Claassen HH, et al. Establishment and characterization of 7 ovarian carcinoma cell lines and one granulosa tumor cell line: growth features and cytogenetics. *Int J Cancer*. 1993;53:613–20.
13. Hamilton TC, Young RC, McKoy WM, Grotzinger KR, Green JA, Chu EW, et al. Characterization of a human ovarian carcinoma cell line (NIH:OVCA-3) with androgen and estrogen receptors. *Cancer Res*. 1983;43:5379–89.
14. Provencher DM, Lounis H, Champoux L, Tétrault M, Manderson EN, Wang JC, et al. Characterization of four novel epithelial ovarian cancer cell lines. *In Vitro Cell Dev Biol Anim*. 2000;36:357–61.
15. Buick RN, Pullano R, Trent JM. Comparative properties of five human ovarian adenocarcinoma cell lines. *Cancer Res*. 1985;45:3668–76.
16. Li W, Turner A, Aggarwal P, Matter A, Storvick E, Arnett DK, et al. Comprehensive evaluation of AmpliSeq transcriptome, a novel targeted whole transcriptome RNA sequencing methodology for global gene expression analysis. *BMC Genomics*. 2015;16:1069.
17. Buttarelli M, Babini G, Raspaglio G, Filippetti F, Battaglia A, Ciucci A, et al. A combined ANXA2-NDRG1-STAT1 gene signature predicts response to chemoradiotherapy in cervical cancer. *J Exp Clin Cancer Res*. 2019;38:279.
18. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paeppe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*. 2002;3(7):RESEARCH0034.
19. Andersen CL, Jensen JL, Ørntoft TF. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and Colon Cancer data sets. *Cancer Res*. 2004;64:5245–50.
20. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods*. 2001;25:402–8.
21. Kucukural A, Yukselen O, Ozata DM, Moore MJ, Garber M. DEBrowser: interactive differential expression analysis and visualization tool for count data. *BMC Genomics*. 2019;20:6.
22. Leek JT, Johnson WE, Parker HS, Fertig EJ, Jaffe AE, Zhang Y, et al. Sva: surrogate variable analysis. R package version 3.40.0; 2021.
23. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550.
24. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple hypothesis testing. *J Royal Stat Soc Ser B*. 1995;57:289–300.
25. R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2021. <https://www.R-project.org/>
26. Liaw A, Wiener M. Classification and regression by randomForest. *R News*. 2002;2:18–22.
27. Slawski M, Boulesteix A, Bernau C (2021). CMA: synthesis of microarray-based classification. R package version 1.50.0.
28. Drton M, Eichler M, Richardson TS. Computing maximum likelihood estimates in recursive linear models with correlated errors. *J Mach Learn Res*. 2009;10:2329–48.
29. Palluzzi F, Grassi M. SEMgraph: An R Package for Causal Network Analysis of High-Throughput Data with Structural Equation Models. 2021; arXiv:2103.08332.
30. Xie Z, Bailey A, Kuleshov MV, Clarke DJB, Evangelista JE, Jenkins SL, et al. Gene set knowledge discovery with Enrichr. *Curr Protoc*. 2021;1(3):e90.
31. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag; 2016.
32. Montojo J, Zuberi K, Rodriguez H, Kazi F, Wright G, Donaldson SL, et al. GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. *Bioinformatics*. 2010;26:2927–8.
33. Darnell JC, Albert ML, Darnell RB. Cdr2, a target antigen of naturally occurring human tumor immunity, is widely expressed in gynecological tumors. *Cancer Res*. 2000;60:2136–9.
34. Junttila MR, de Sauvage FJ. Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature*. 2013;501:346–54.
35. Nagai N, Kinoshita M, Ogata H, Tsujino D, Wada Y, Someya K, et al. Relationship between pharmacokinetics of unchanged cisplatin and nephrotoxicity after intravenous infusions of cisplatin to cancer patients. *Cancer Chemother Pharmacol*. 1996;39:131–7.
36. de Witte CJ, Espejo Valle-Inclan J, Hami N, Löhmußaar K, Kopper O, Vreuls CPH, et al. Patient-derived ovarian Cancer Organoids mimic clinical response and exhibit heterogeneous inter- and Intrapatient drug responses. *Cell Rep*. 2020;31(11):107762.
37. Domcke S, Sinha R, Levine DA, Sander C, Schultz N. Evaluating cell lines as tumour models by comparison of genomic profiles. *Nat Commun*. 2013;4:2126.
38. Gillet JP, Calcagno AM, Varma S, Marino M, Green LJ, Vora MI, et al. Redefining the relevance of established cancer cell lines to the study of mechanisms of clinical anti-cancer drug resistance. *Proc Natl Acad Sci U S A*. 2011;108:18708–13.
39. Yang Y, Yang Y, Yang J, Zhao X, Wei X. Tumor microenvironment in ovarian Cancer: function and therapeutic strategy. *Front Cell Dev Biol*. 2020;8:758.
40. Koti M, Siu A, Clément I, Bidarimath M, Turashvili G, Edwards A, et al. A distinct pre-existing inflammatory tumour microenvironment is associated with chemotherapy resistance in high-grade serous epithelial ovarian cancer. *Br J Cancer*. 2015;112:1215–22.
41. Josephs DH, Spicer JF, Karagiannis P, Gould HJ, Karagiannis SN. IgE immunotherapy: a novel concept with promise for the treatment of cancer. *MAbs*. 2014;6:54–72.
42. Pellizzari G, Bax HJ, Josephs DH, Gotovina J, Jensen-Jarolim E, Spicer JF, et al. Harnessing therapeutic IgE antibodies to re-educate macrophages against Cancer. *Trends Mol Med*. 2020;26:615–26.
43. Pellizzari G, Hoskin C, Crescioli S, Mele S, Gotovina J, Chiaruttini G, et al. IgE re-programs alternatively-activated human macrophages towards pro-inflammatory anti-tumoural states. *EBioMedicine*. 2019;43:67–81.
44. Ciucci A, Zannoni GF, Buttarelli M, Martinelli E, Mascilini F, Petrillo M, et al. Ovarian low and high grade serous carcinomas: hidden divergent features in the tumor microenvironment. *Oncotarget*. 2016a;7:68033–43.
45. Ciucci A, Zannoni GF, Buttarelli M, Lisi L, Travaglia D, Martinelli E, et al. Multiple direct and indirect mechanisms drive estrogen-induced tumor growth in high grade serous ovarian cancers. *Oncotarget*. 2016b;7:8155–71.
46. Nowak M, Klink M. The role of tumor-associated macrophages in the progression and Chemoresistance of ovarian Cancer. *Cells*. 2020;9:1299.
47. Jackson JT, Mulazzani E, Nutt SL, Masters SL. The role of PLC γ 2 in immunological disorders, cancer, and neurodegeneration. *J Biol Chem*. 2021;297:100905.
48. Wen R, Jou ST, Chen Y, Hoffmeyer A, Wang D. Phospholipase C gamma 2 is essential for specific functions of fc epsilon R and fc gamma R. *J Immunol*. 2002;169:6743–52.
49. Toyama-Sorimachi N, Kobayashi T. Lysosomal amino acid transporters as key players in inflammatory diseases. *Int Immunol*:dxab069. 33:853–58.
50. Song F, Yi Y, Li C, Hu Y, Wang J, Smith DE, et al. Regulation and biological role of the peptide/histidine transporter SLC15A3 in toll-like receptor-mediated inflammatory responses in macrophage. *Cell Death Dis*. 2018;9:770.
51. Apetoh L, Ghiringhelli F, Tesniere A, Obeid M, Ortiz C, Criollo A, et al. Toll-like receptor 4-dependent contribution of the immune system to anticancer chemotherapy and radiotherapy. *Nat Med*. 2007;13:1050–9.
52. Hossain MN, Sakemura R, Fujii M, Ayusawa D. G-protein gamma subunit GNG11 strongly regulates cellular senescence. *Biochem Biophys Res Commun*. 2006;351:645–50.
53. Guillon J, Petit C, Toutain B, Guette C, Lelièvre E, Coqueret O. Chemotherapy-induced senescence, an adaptive mechanism driving resistance and tumor heterogeneity. *Cell Cycle*. 2019;18:2385–97.
54. Cinali RM, Herman CE, Lew BO, Wieman HL, Thompson CB, Rathmell JC. T cell homeostasis requires G protein-coupled receptor-mediated access to trophic signals that promote growth and inhibit chemotaxis. *Eur J Immunol*. 2005;35:786–95.
55. de Biasi AR, Villena-Vargas J, Adusumilli PS. Cisplatin-induced antitumor immunomodulation: a review of preclinical and clinical evidence. *Clin Cancer Res*. 2014;20:5384–91.
56. Jin L, Shen F, Weinfeld M, Sergi C. Insulin growth factor binding protein 7 (IGFBP7)-related Cancer and IGFBP3 and IGFBP7 crosstalk. *Front Oncol*. 2020;10:727.
57. Gambaro K, Quinn MC, Cáceres-Gorriti KY, Shapiro RS, Provencher D, Rahimi K, et al. Low levels of IGFBP7 expression in high-grade serous ovarian carcinoma is associated with patient outcome. *BMC Cancer*. 2015;15:135.

58. Okamura J, Huang Y, Moon D, Brait M, Chang X, Kim MS. Downregulation of insulin-like growth factor-binding protein 7 in cisplatin-resistant non-small cell lung cancer. *Cancer Biol Ther*. 2012;13:148–55.
59. Hurov KE, Cotta-Ramusino C, Elledge SJ. A genetic screen identifies the triple T complex required for DNA damage signaling and ATM and ATR stability. *Genes Dev*. 2010;24:1939–50.
60. Li Y, Guo H, Jin C, Qiu C, Gao M, Zhang L, et al. Spliceosome-associated factor CTNNB1 promotes proliferation and invasion in ovarian cancer. *Exp Cell Res*. 2017;357:124–34.
61. Cancer Genome Atlas Research. Integrated genomic analyses of ovarian carcinoma. *Nature*. 2011;474:609–15.
62. Zhao F, Siu MK, Jiang L, Tam KF, Ngan HY, Le XF, et al. Overexpression of forkhead box protein M1 (FOXM1) in ovarian cancer correlates with poor patient survival and contributes to paclitaxel resistance. *PLoS One*. 2014;9:e113478.
63. Tassi RA, Todeschini P, Siegel ER, Calza S, Cappella P, Ardighieri L, et al. FOXM1 expression is significantly associated with chemotherapy resistance and adverse prognosis in non-serous epithelial ovarian cancer patients. *J Exp Clin Cancer Res*. 2017;36:63.
64. Kar SP, Considine DPC, Tyrer JP, Plummer JT, Chen S, Dezem FS, et al. Pleiotropy-guided transcriptome imputation from normal and tumor tissues identifies candidate susceptibility genes for breast and ovarian cancer. *HGG Adv*. 2021;2:100042.
65. Boldt K, van Reeuwijk J, Lu Q, Koutroumpas K, Nguyen TM, Texier Y, et al. An organelle-specific protein landscape identifies novel diseases and molecular mechanisms. *Nat Commun*. 2016;13(7):11491.
66. Lontos M, Fiste O, Drakopoulou D, Thomakos N, Goula K, Zagouri F, et al. Paraneoplastic cerebellar degeneration in platinum-responsive endometrial cancer: a case report and review of literature. *Gynecol Oncol Rep*. 2021;37:100826.
67. Huddleston HG, Wong KK, Welch WR, Berkowitz RS, Mok SC. Clinical applications of microarray technology: creatine kinase B is an up-regulated gene in epithelial ovarian cancer and shows promise as a serum marker. *Gynecol Oncol*. 2005;96:77–83.
68. Li XH, Chen XJ, Ou WB, Zhang Q, Lv ZR, Zhan Y, et al. Knockdown of creatine kinase B inhibits ovarian cancer progression by decreasing glycolysis. *Int J Biochem Cell Biol*. 2013;45:979–86.
69. Wang J, Zhou Y, Li D, Sun X, Deng Y, Zhao Q. TSPAN31 is a critical regulator on transduction of survival and apoptotic signals in hepatocellular carcinoma cells. *FEBS Lett*. 2017;591:2905–18.
70. Arend RC, Londoño-Joshi AI, Straughn JM Jr, Buchsbaum DJ. The Wnt/ β -catenin pathway in ovarian cancer: a review. *Gynecol Oncol*. 2013;131(3):772–9.
71. Lussier MP, Lepage PK, Bousquet SM, Boulay G. RNF24, a new TRPC interacting protein, causes the intracellular retention of TRPC. *Cell Calcium*. 2008;43:432–43.
72. Dressman HK, Hans C, Bild A, Olson JA, Rosen E, Marcom PK, et al. Gene expression profiles of multiple breast cancer phenotypes and response to neoadjuvant chemotherapy. *Clin Cancer Res*. 2006;12:819–26.
73. Lin T, Gu J, Qu K, Zhang X, Ma X, Miao R, et al. A new risk score based on twelve hepatocellular carcinoma-specific gene expression can predict the patients' prognosis. *Aging (Albany N. Y.)*. 2018;10:2480.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

