

# UCLA

## UCLA Previously Published Works

### Title

Drug discovery using clinical outcome-based Connectivity Mapping: application to ovarian cancer

### Permalink

<https://escholarship.org/uc/item/01z9q0k2>

### Journal

BMC Genomics, 17(1)

### ISSN

1471-2164

### Authors

Raghavan, Rama  
Hyter, Stephen  
Pathak, Harsh B  
et al.

### Publication Date

2016-12-01

### DOI

10.1186/s12864-016-3149-5

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

RESEARCH ARTICLE

Open Access



# Drug discovery using clinical outcome-based Connectivity Mapping: application to ovarian cancer

Rama Raghavan<sup>1</sup>, Stephen Hyter<sup>2</sup>, Harsh B. Pathak<sup>2</sup>, Andrew K. Godwin<sup>2</sup>, Gottfried Konecny<sup>3</sup>, Chen Wang<sup>4</sup>, Ellen L. Goode<sup>4</sup> and Brooke L. Fridley<sup>1\*</sup>

## Abstract

**Background:** Epithelial ovarian cancer (EOC) is the fifth leading cause of cancer death among women in the United States (5 % of cancer deaths). The standard treatment for patients with advanced EOC is initial debulking surgery followed by carboplatin-paclitaxel combination chemotherapy. Unfortunately, with chemotherapy most patients relapse and die resulting in a five-year overall survival around 45 %. Thus, finding novel therapeutics for treating EOC is essential. Connectivity Mapping (CMAP) has been used widely in cancer drug discovery and generally has relied on cancer cell line gene expression and drug phenotype data. Therefore, we took a CMAP approach based on tumor information and clinical endpoints from high grade serous EOC patients.

**Methods:** We determined tumor gene expression signatures (e.g., sets of genes) associated with time to recurrence (with and without adjustment for additional clinical covariates) among patients within TCGA ( $n = 407$ ) and, separately, from the Mayo Clinic ( $n = 326$ ). Each gene signature was inputted into CMAP software (Broad Institute) to determine a set of drugs for which our signature “matches” the “reference” signature, and drugs that overlapped between the CMAP analyses and the two studies were carried forward for validation studies involving drug screens on a set of 10 EOC cell lines.

**Results:** Of the 11 drugs carried forward, five (mitoxantrone, podophyllotoxin, wortmannin, doxorubicin, and 17-AAG) were known a priori to be cytotoxics and were indeed shown to effect EOC cell viability.

**Conclusions:** Future research is needed to investigate the use of these CMAP and similar analyses for determining combination therapies that might work synergistically to kill cancer cells and to apply this *in silico* bioinformatics approach using clinical outcomes to other cancer drug screening studies.

**Keywords:** Gene expression signature, Time to recurrence, Bioinformatics, Connectivity Mapping, Drug discovery, Ovarian cancer

## Background

The American Cancer Society estimates that in 2015, approximately 14,180 women will die of epithelial ovarian cancer (EOC) in the United States, with 21,290 estimated new cases [1]. Since early ovarian cancer shows few symptoms, the vast majority of patients continue to be diagnosed with advanced stage disease, where the prognosis is poor with 5-year survival rate around 27 %. The standard

treatment protocol for patients with advanced EOC is an initial debulking surgery, followed by carboplatin-paclitaxel combination chemotherapy. Platinating agents, such as cisplatin, carboplatin, and oxaliplatin, are able to interact with DNA to form monoadducts, intra- and interstrand crosslinks, and DNA-protein crosslinks, ultimately resulting in cell death. Taxane agents are widely used chemotherapeutic drugs often in combination with platinating agents. Taxanes, such as paclitaxel and docetaxel, block cell division by binding to  $\beta$ -tubulin, stabilizing the microtubules, leading to cell death [2, 3]. Although over 70 % of ovarian cancer patients are initially sensitive to the combination

\* Correspondence: bfridley@kumc.edu

<sup>1</sup>Department of Biostatistics, University of Kansas Medical Center, 3901 Rainbow Blvd, Kansas City, KS 66160, USA

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



therapy consisting of a platinum agent and a taxane, the majority will experience relapse and subsequent resistance to the therapy [4]. Thus, finding new therapeutic options for treating patients with EOC is essential.

One approach that has been used widely in cancer drug discovery is Connectivity Mapping [5]. The Connectivity Map (also known as “CMAP”) is a collection of genome-wide transcriptional expression data from cultured human cells treated with bioactive small molecules analyzed using pattern-matching algorithms that discover relationships between the drugs, gene expression changes, and the phenotypes. This computational approach has greatly facilitated drug screening studies, as CMAP contains more than 7000 gene expression profiles for approximately 1300 compounds (<https://www.broadinstitute.org/cmap/>). In particular, it has been employed in many studies for discovering repurposing drugs against common diseases, including diabetes and Alzheimer’s disease [5], and for treating solid tumors, including those associated with colon cancer [6], breast cancer [7], and lung adenocarcinoma [8]. The basic approach for CMAP-based drug discovery studies is the identification of disease- (or phenotype) associated genomic signatures that inversely correlate with perturbation in the genomic signature associated with the administration of molecules or drugs [5]. In these studies, the essence of the protocol – the individual-gene CMAP approach – for identifying drugs for treating a specific disease is straightforward: find a set of differentially expressed genes (DEGs) obtained by comparing two sets – e.g., control and patient tissues – of gene expression microarrays, score the match between the DEG set and genomic profiles of drugs given by CMAP, and rank the drugs by score [9, 10]. The candidate drugs are those with the highest absolute scores.

However, these studies have some limitations. First, the list of DEGs used in the CMAP analysis is usually based on a relatively small number of biological replicates (i.e., a handful of cancer cell lines). Additionally, the number of cell lines in CMAP exposed to the compounds is limited to only a handful of cancer cell lines from breast, leukemia, prostate and melanoma, with each compound usually tested on only a few cell lines. Furthermore, recent studies have shown potential issues with use of cancer cell lines in terms of the lack of rigor in the estimation of drug response phenotypes in cell lines [11] and the lack of concordance between cell lines and human genomic profiles [12]. Among EOC cell lines, a recent study has found that IGROV1 is most probably not of the high grade serous subtype as it is often quoted [13]. To address the limitations of previous DEG selection based on cancer cell lines, we determined our DEGs based on two large collections of tumor gene expression data collected on high grade serous EOC patients for whom clinical endpoints were available (407 and 326 cases in TCGA and Mayo Clinic studies,

respectively). To determine the most relevant DEGs, we characterized the associations between gene expression and time to recurrence. We hypothesized that the potential therapeutic drugs for EOC are those that have a gene expression profile that are related to the gene expression signature related to clinical outcome of time to recurrence (TTR). Following CMAP analyses, we then tested key genes on ten EOC cell lines to assess the ability of the candidate drugs to effectively kill EOC cells.

## Methods

### TCGA ovarian cancer study

As part of TCGA, research collected and assessed genome-wide gene expression data on 518 samples using the Agilent Expression 244 K microarray. Gene expression and clinical data were downloaded from the TCGA Data Portal (<https://tcga-data.nci.nih.gov/tcga/dataAccessMatrix.htm>) on September 17, 2012. Gene expression data were lowess normalized with replicate probes for a gene collapsed by averaging across the probes. Of 518 serous cyst adenocarcinoma with Agilent gene expression data, 449 of the tumors were classified as high-grade. Thus, restricting to high grade serous tumors and removing the samples in the TCGA that were part of the Mayo Clinic study, we included 407 tumors in our analysis. Summary of the TCGA participants are presented in Table 1.

### Mayo clinic ovarian cancer study

Briefly, eligible cases were ascertained between 1992 and 2009 at the Mayo Clinic within 1 year of diagnosis with pathologically confirmed primary invasive high-grade serous EOC. Progression and vital status were obtained from the Mayo Clinic Tumor Registry, electronic medical records, and active patient contact. All cases provided

**Table 1** Summary of TCGA and Mayo Clinic Studies

	TCGA (n = 407)	Mayo Clinic (N = 326)
Time to recurrence/Progression free survival		
Median TTR	15.3 (months)	13.0 (months)
No recurrence/No progression	49.87 (%)	31.59 (%)
Recurrence/Progression	48.40 (%)	67.79 (%)
Stage		
I	3.19 (%)	1.53 (%)
II	4.17 (%)	3.06 (%)
III or IV	92.13 (%)	95.39 (%)
Age		
Mean, [Q1, median, Q3]	59.8 [53, 60, 68]	58.3 [50, 55.5, 66.25]
Surgical debulking		
Optimal	65.11 (%)	73.92 (%)
Sub-optimal	23.83 (%)	24.84 (%)

written informed consent for use of their tissues and medical records in research; all protocols were approved by the Mayo Clinic Institutional Review Board. RNA from fresh frozen tumors of each patient was extracted and assessed using Agilent Whole Human Genome 4 × 44 K Expression Arrays as previously described [14, 15]. The program “ComBat” was used to correct for batch-effects due to Cy5 and Cy3 labeling differences observed among experimental batches [16]. Summary of 326 Mayo Clinic participants are presented in Table 1. Data used in this study can be found at the Gene Expression Omnibus (GSE73614, GSE53963 and GSE74357).

### Statistical and CMAP analyses

Cox proportional hazard models were used to assess the association of gene expression (gene-by-gene) with time to recurrence (TTR), by study with and without adjustment for age at diagnosis, stage, and debulking status. DEG probes were selected for inclusion in the CMAP analysis if  $p < 0.01$  for the unadjusted analyses and  $p < 0.005$  for the covariate adjusted analyses; different thresholds were used as CMAP (Broad Institute) has a limit on the number of features included in any signature. These set of probes were then mapped to genes and then to the Affymetrix ID, as CMAP is based on Affymetrix probes/features. We conducted CMAP analysis on individual genes with hazard ratio (HR)  $> 1$  coded as “positively” associated and genes with HR  $< 1$  coded as “negatively” associated with TTR to determine a set of drugs for which our gene signatures matches the “reference” signature (either positively or negatively). Clustering of samples based on gene expression levels was completed using recursive partition mixture models using the R package *RPMM* [17], restricting the number of levels to a maximum of 2.

### In vitro drug screens

In vitro drug cytotoxicity assays were conducted to determine which of the drugs highlighted by CMAP analysis affected viability of EOC cells. The drugs were purchased from the following vendors: cotinine, 3-nitropropionic acid, adiphene hydrochloride, ethosuximide, and podophylotoxin (Sigma); cephalexin and mitoxantrone (Selleckchem); clemizole (BioVision); wortmannin, doxorubicin, and 17-AAG (LC Laboratories). Upon receipt, dimethyl sulfoxide (DMSO) was used to prepare 10 mM stock solutions for all of the drugs except for cephalexin, which was prepared at a 5 mM concentration due to reduced solubility. Single-use aliquots of the stock drug solutions were made and stored at  $-80^{\circ}\text{C}$ .

All cell lines used in this study were obtained or derived at the Fox Chase Cancer Center (Philadelphia, PA). Details of the origin of the EOC cell lines ( $N = 10$ : A1847, A2780, C30, CP70, OVCAR4, OVCAR5, OVCAR8, OVCAR10,

PEO4, and SKOV3) have been previously reported [18–20]. Each cell line was grown in RPMI 1640 (Corning Cellgro) containing 2 mM L-glutamine and supplemented with 10 % FBS (Gibco), 100 I.U./mL penicillin (Corning Cellgro), 100  $\mu\text{g}/\text{mL}$  streptomycin (Corning Cellgro), and 7.5  $\mu\text{g}/\text{mL}$  insulin (Gibco) and maintained at  $37^{\circ}\text{C}$  in a humidified atmosphere with 5 %  $\text{CO}_2$ . Cell lines were grown to 80 % confluency, harvested and seeded into 96-well plates at concentrations of 2000 to 4000 cells per well in a total volume of 95  $\mu\text{L}$ . Twenty-four hours after seeding, drug compounds were prepared using cell growth media and 5  $\mu\text{L}$  of each were added to the seeded cells in the 96-well plates. A Microlab Nimbus 96 pipetting robot (Hamilton) was used to prepare the serial dilutions and for addition to the cell lines. The final drug solutions consisted of eight concentrations ranging from 20 to 0.0012  $\mu\text{M}$  (serial four-fold dilutions). Vehicle-only wells were included on each plate to serve as interplate normalization controls.

Seventy-two hours following drug addition, a 1/5th volume of CellTiter Blue reagent (Promega) was added directly to each well using a Matrix WellMate (Thermo Scientific). The plates were incubated at  $37^{\circ}\text{C}$  for 150 min and the fluorescent signal was measured using an Infinite® M200 Pro microplate reader (Tecan). The ratio of the fluorescent signal in a drug treated well to that of the average fluorescent signal from the vehicle treated wells on each plate multiplied by 100 was used to calculate cell viability for each drug treated well for each cell line. A minimum of two biological replicates were performed for each cell line. The viability data were subjected to non-linear regression analysis and  $\text{IC}_{50}$  values calculated using Prism 5 software (GraphPad). All data in the viability curves are reported as mean  $\pm$  standard error of the mean (SEM).

## Results

### Genes associated with clinical outcome

To determine the gene signatures to use in CMAP analyses, we determined the set of probes (and corresponding gene) associated with TTR by fitting separate Cox proportional hazards models within each study ( $N = 407$  for TCGA,  $N = 326$  for Mayo Clinic) with and without adjustment for age at diagnosis, stage, and debulking status. Comparing the genes associated with TTR with no adjustment for covariates at the 0.05 significance level and the same direction of effect resulted in 143 genes in common between TCGA and Mayo studies; 96 genes had a hazard ratio (HR)  $< 1$  (high expression associated with better outcome) and 47 genes had a HR  $> 1$  (high expression associated with worse outcome). In contrast, the adjusted analyses resulted in 186 genes in common ( $p < 0.05$ , HR in same direction), with 128 genes having a HR  $< 1$  and 58 having a HR  $> 1$ . Table 2

**Table 2** Genes associated with epithelial ovarian cancer time to recurrence in analysis of tumors from the TCGA and Mayo Clinic ( $p < 0.01$ ) for both adjusted and unadjusted analyses

Analysis	Gene	TCGA study			Mayo Clinic study		
		HR	95 % CI	P	HR	95 % CI	P
Adjusted for covariates	<i>PTPRCAP</i>	0.71	(0.56,0.89)	0.004	0.14	(0.04,0.40)	3.6E-04
	<i>UBASH3A</i>	0.78	(0.65,0.92)	0.005	0.19	(0.06,0.50)	0.001
	<i>PPBP</i>	1.18	(1.04,1.33)	0.009	5.79	(1.94,17.21)	0.002
	<i>PVRIG</i>	0.73	(0.58,0.91)	0.005	0.28	(0.12,0.62)	0.002
	<i>IGKV3-20</i>	0.85	(0.76,0.93)	0.001	0.41	(0.23,0.72)	0.002
	<i>FCRL5</i>	0.84	(0.74,0.95)	0.008	0.30	(0.14,0.65)	0.002
	<i>ITK</i>	0.83	(0.72,0.95)	0.008	0.30	(0.13,0.66)	0.003
	<i>IGHV3OR16-12</i>	0.85	(0.77,0.94)	0.002	0.47	(0.28,0.78)	0.004
	<i>VANGL1</i>	0.67	(0.51,0.88)	0.004	0.23	(0.08,0.62)	0.004
	<i>SLC16A8</i>	1.23	(1.06,1.43)	0.006	4.55	(1.58,13.01)	0.005
	<i>IGKV3-20</i>	0.89	(0.82,0.96)	0.004	0.50	(0.30,0.83)	0.007
	<i>ACAP1</i>	0.78	(0.65,0.93)	0.006	0.23	(0.07,0.68)	0.008
	Unadjusted for covariates	<i>VANGL1</i>	0.73	(0.58,0.92)	0.008	0.26	(0.09,0.69)
<i>CD38</i>		0.88	(0.80,0.96)	0.007	0.47	(0.28,0.76)	0.003
<i>ELA1</i>		1.42	(1.10,1.82)	0.006	4.93	(1.97,12.32)	0.001
<i>PSME2</i>		0.76	(0.62,0.91)	0.004	0.16	(0.04,0.53)	0.003
<i>UBD</i>		0.92	(0.87,0.97)	0.004	0.59	(0.43,0.81)	0.001

Note, three Agilent probes did not have a current HUGO gene ID and are not presented in the table

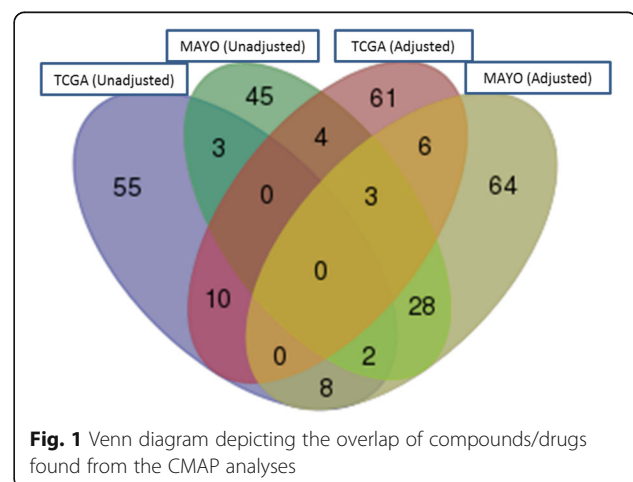
presents the genes with  $p < 0.01$  in both studies (with same direction of effect) for both the adjusted and unadjusted analyses.

To further evaluate that these sets of genes were predictive of outcome, we clustered the samples based on the expression levels of the genes in common between the two studies using recursive partitioning mixture models (RPMM). TCGA and Mayo Clinic samples were clustered separately on the set of genes in common between the two studies (with and without adjusting for covariates). The resulting cluster assignments were then assessed for association with survival using log-rank tests. The log-rank p-values for testing cluster assignment with TTR were 0.012 and 0.015 for TCGA and Mayo Clinic studies (genes in common with no adjustment for covariates) (Additional file 1: Figure S1) and 0.016 and 0.112 based on clustering of genes from model adjusting for covariates (Additional file 2: Figure S2).

### Connectivity Mapping

CMAF analysis was completed for each of the 4 DEG sets (TCGA/Unadjusted Analyses; TCGA/Adjusted Analyses; Mayo/Unadjusted Analyses; Mayo/Adjusted Analyses) using the CMAF software developed at the Broad Institute (<https://www.broadinstitute.org/cmaf/>). This resulted in ( $p < 0.05$ ): 78, 84, 85 and 111 drugs signatures were found to be either negatively or positively related with gene

signatures based on analysis of TCGA (unadjusted), TCGA (adjusted), Mayo Clinic (unadjusted) and Mayo Clinic (adjusted), respectively. We then looked at the overlap of the compounds found in both studies and found that 9 (and 5) compounds were in common between CMAF analyses based on the covariate adjusted (unadjusted) signature, as illustrated in Fig. 1. This set of compounds included the following: tanespimycin (17-AAG), ethosuximide, cotinine, clemizole, 0175029-0000





(unadjusted for covariates); wortmannin, 3-nitropropionic acid, adiphenine, cephaline, doxorubicin, podophyllotoxin, mitoxantrone, cephalixin, 5182598 (adjusted for covariates). In vitro drug screens were completed using these drugs with the exception of 0175029–0000 and 5182598 for which commercial sources were not found and cephaline which was cost prohibitive.

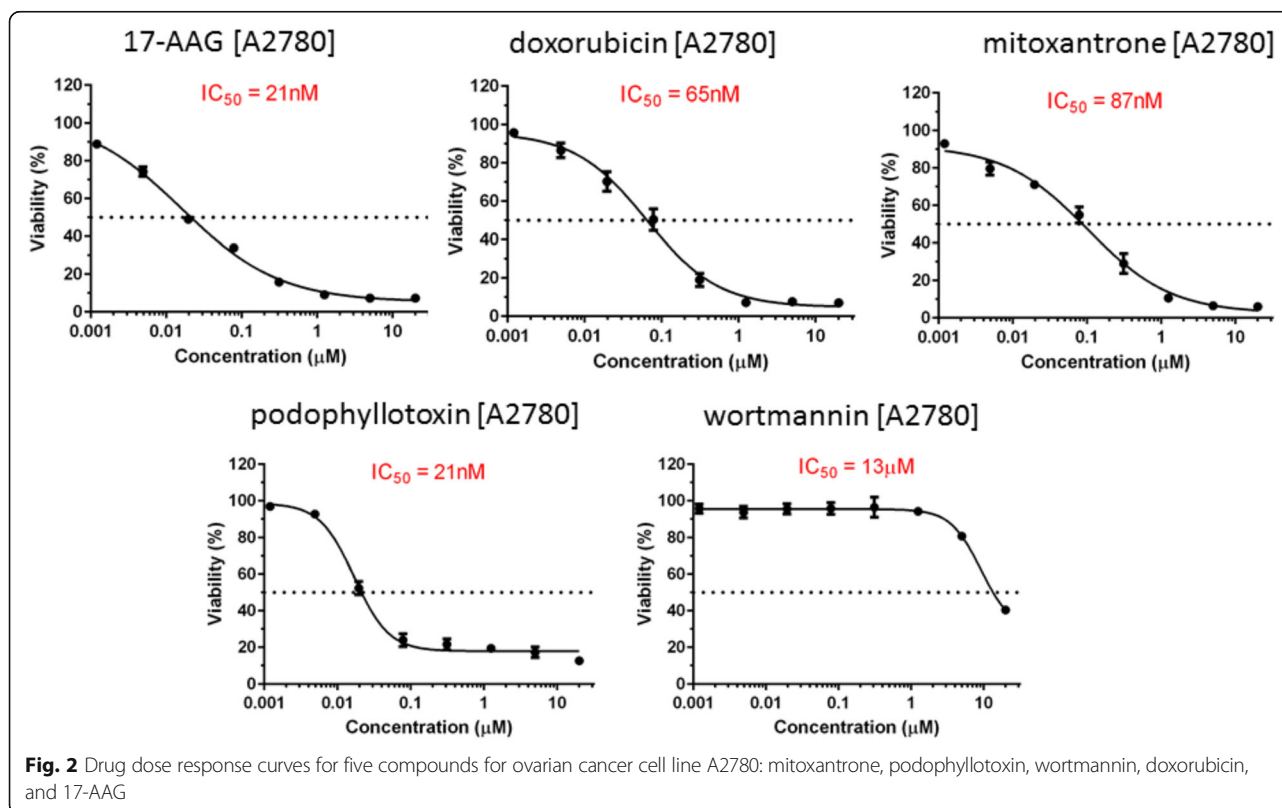
### In vitro drug screens

In order to experimentally determine the effect of the eleven CMAP-highlighted drugs on EOC cell viability, we conducted in vitro dose response studies on a set of 10 EOC cell lines. A significant reduction in cell viability following 72 h of drug treatment was observed for five of these compounds: mitoxantrone, podophyllotoxin, wortmannin, doxorubicin, and 17-AAG (Fig. 2). These 5 drugs were known to be cytotoxic and therefore were expected to affect EOC cell growth and viability based on their mechanisms of action [21–28]. The six remaining drugs failed to show in vitro anti-cancer activity, consistent with prior reports of minimal cytotoxic capacity [29–34]. The dose response data for all 11 drugs across the 10 EOC cell lines can be found in Additional file 3: Figure S3, Additional file 4: Figure S4, Additional file 5: Figure S5, Additional file 6: Figure S6, Additional file 7: Figure S7, Additional file 8:

Figure S8, Additional file 9: Figure S9, Additional file 10: Figure S10, Additional file 11: Figure S11, Additional file 12: Figure S12 and Additional file 13: Figure S13.

### Discussion

In this manuscript we have presented a bioinformatics approach for connectivity mapping based on clinical outcomes collected on a large patient population followed by functional validation of the identified drugs. The strengths of this approach are: non-reliance on a signature determined based on a small number of cancer cell lines; use clinically relevant outcomes that are directly tied to response (i.e., time to recurrence or overall survival); large clinical studies that provide the ability to look at overlap between CMAP determined drugs; functional validation to confirm the ability of the discovered drugs to kill ovarian cancer cells. The choice of applying this approach to ovarian cancer was by design, as EOC treatment is somewhat uniform where the majority of patients undergo tumor debulking or cytoreduction surgery to remove as much of the tumor as possible, followed by platinum-taxane combination chemotherapy. Application of this approach to other cancers with less standard treatments will introduce heterogeneity into clinical outcome and would need to be accounted for in the statistical analyses (if possible). Nevertheless, there are still limitations to CMAP analysis, which include: small sample size contributing to the CMAP database,



with none of the cell lines having been derived from EOC; “batch effects” in CMAP cell line cultures [35]; and the signatures are based on only gene expression measured using microarrays.

This study is a proof of principle that clinical outcomes from large studies (of which one is publically available for research) have the ability to be leveraged for drug discovery. Of the 11 drugs carried forward, we a priori hypothesized that 5 of these drugs would affect EOC cell viability (mitoxantrone, podophyllotoxin, wortmannin, doxorubicin, and 17-AAG), for which all 5 showed an ability to kill EOC cells *in vitro* (cell lines were treated using serial dilutions of the drugs (0 to 20  $\mu$ M) for 72 h followed by measuring cell viability using the CellTiter-Blue assay). Their anti-cancer activities have been studied extensively and mitoxantrone and doxorubicin have been used in the treatment of a variety of cancers [36–43]. However, the CMAP analyses of the TCGA and Mayo Clinic studies did not identify the two most commonly used therapies for EOC, with none of the 11 drugs identified having similar structure or mechanism of action to carboplatin or paclitaxel. However, three of the 11 compounds are topoisomerase II inhibitors (doxorubicin, podophyllotoxin, mitoxantrone), a class of drugs used often in the treatment of breast cancer, lung cancer, testicular cancer, lymphomas and sarcomas [44].

The natural products podophyllotoxin, wortmannin, and 17-AAG have proven highly toxic to human subjects but subsequent analogues such as etoposide, PX-866, and ganetespi, respectively, have decreased side effects and are currently being investigated in the treatment of EOC [13, 45, 46]. The six compounds that failed to exhibit significant *in vitro* cytotoxicity across the EOC cell line panel belong to classes of drugs related to central nervous system pathways, exhibit antimicrobial activity or are required in high micromolar concentrations which may not be physiologically achievable. Interestingly, adiphenine was also identified using CMAP analysis as an adjuvant therapy to treat the psychological distress associated with EOC diagnosis [33]. While novel drugs were not identified, a filtered list of six was obtained for directed *in vitro* testing.

## Conclusions

Future research is needed to investigate the use of these CMAP-like analyses for determining combination therapies that might work synergistically to kill cancer cells and to apply this *in silico* bioinformatics approach using clinical outcomes to other cancer drug screening studies. Last of all, CMAP analyses only determines candidate drugs that can be tested in future studies with no information provided on the optimal dose in

humans; research into the optimal therapeutic dosage needs to be considered in the planning of future drug studies [47].

## Additional files

**Additional file 1: Figure S1.** Results from clustering tumors from the TCGA and Mayo Clinic studies. Analyses were completed with genes found to be associated with clinical outcome in both studies (no adjustment for clinical covariates). (PNG 1066 kb)

**Additional file 2: Figure S2.** Results from clustering tumors from the TCGA and Mayo Clinic studies. Analyses were completed with genes found to be associated with clinical outcome in both studies (adjustment for clinical covariates). (PNG 1017 kb)

**Additional file 3: Figure S3.** The dose response data for 3-nitropropionic acid across the 10 EOC cell lines. (PNG 1781 kb)

**Additional file 4: Figure S4.** The dose response data for 17-AAG across the 10 EOC cell lines. (PNG 1963 kb)

**Additional file 5: Figure S5.** The dose response data for adiphenine hydrochloride across the 10 EOC cell lines. (PNG 1763 kb)

**Additional file 6: Figure S6.** The dose response data for cephalexin across the 10 EOC cell lines. (PNG 1780 kb)

**Additional file 7: Figure S7.** The dose response data for clemizole across the 10 EOC cell lines. (PNG 1768 kb)

**Additional file 8: Figure S8.** The dose response data for cotinine across the 10 EOC cell lines. (PNG 1793 kb)

**Additional file 9: Figure S9.** The dose response data for doxorubicin across the 10 EOC cell lines. (PNG 2018 kb)

**Additional file 10: Figure S10.** The dose response data for ethosuximide across the 10 EOC cell lines. (PNG 1762 kb)

**Additional file 11: Figure S11.** The dose response data for mitoxantrone across the 10 EOC cell lines. (PNG 2056 kb)

**Additional file 12: Figure S12.** The dose response data for podophyllotoxin across the 10 EOC cell lines. (PNG 1927 kb)

**Additional file 13: Figure S13.** The dose response data for wortmannin across the 10 EOC cell lines. (PNG 1835 kb)

**Additional file 14: Table S1.** DGE and CMAP results. (XLSX 489 kb)

## Abbreviations

CMAP: Connectivity Mapping; DEGs: Differentially expressed genes; DMSO: Dimethyl sulfoxide; EOC: Epithelial ovarian cancer; HR: Hazard ratio; RPM: Recursive partition mixture model; SEM: Standard error of the mean; TCGA: The Cancer Genome Atlas; TTR: Time to recurrence

## Acknowledgements

We would like to thank TCGA for making their data available for community research use. We would also like to thank the BROAD institute for developing CMAP and creating interface for users.

## Funding

Funding for this research was provided in part by National Institute of Health (R01 CA122443, P50 CA136393, P30 CA15083, P30 CA168524, P20 GM103418, R21 CA182715, K12 HD052027). The agencies had no role in the design of study, analysis or interpretation of results.

## Availability of data and materials

Data from the ovarian tumor samples within TCGA were downloaded from the TCGA Data Portal (<https://tcga-data.nci.nih.gov/tcga/dataAccessMatrix.htm>). Gene expression data for the Mayo Clinic study can be found at the Gene Expression Omnibus (GSE73614, GSE53963 and GSE74357). Clinical information on patients in the Mayo Clinic Study can be requested from Dr. Goode or via Dr. Fridley (corresponding author). The drug screen data collected on the 10 EOC cell lines for the 11 drugs can be found in the Additional files 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 and 14.

**Authors' contributions**

Conceived and designed the study: BLF, RR. Performed gene expression studies in tumors: ELG, CW, GK. Performed the in vitro validation experiments: HP, SH, AG. Statistical Analyses: RR, CW. Wrote the paper: RR, BLF, SH, HP. Reviewed Manuscript: ALL. All authors read and approved the final manuscript.

**Competing interests**

The authors declare that they have no competing interests.

**Consent for publication**

Not applicable.

**Ethics approval and consent to participate**

All cases included in the Mayo Clinic study provided written informed consent for use of their tissues and medical records in research; all protocols were approved by the Mayo Clinic Institutional Review Board.

**Author details**

<sup>1</sup>Department of Biostatistics, University of Kansas Medical Center, 3901 Rainbow Blvd, Kansas City, KS 66160, USA. <sup>2</sup>Department of Pathology and Laboratory Medicine, University of Kansas Medical Center, Kansas City, KS 66160, USA. <sup>3</sup>Department of Medicine, Hematology & Oncology, University of California – Los Angeles, Los Angeles, CA 90095, USA. <sup>4</sup>Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55901, USA.

Received: 3 March 2016 Accepted: 7 October 2016

Published online: 19 October 2016

**References**

- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin.* 2015; 65(1):5–29.
- Jordan MA, Wilson L. Microtubules as a target for anticancer drugs. *Nat Rev Cancer.* 2004;4(4):253–65.
- Huizing MT, Misser VH, Pieters RC, ten Bokkel Huinink WW, Veenhof CH, Vermorken JB, Pinedo HM, Beijnen JH. Taxanes: a new class of antitumor agents. *Cancer Invest.* 1995;13(4):381–404.
- Berkenblit A, Cannistra SA. Advances in the management of epithelial ovarian cancer. *J Reprod Med.* 2005;50(6):426–38.
- Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet J-P, Subramanian A, Ross KN, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science.* 2005;313(5795):1929–35.
- Garman KS, Acharya CR, Edelman E, Grade M, Gaedcke J, Sud S, Barry W, Diehl AM, Provenzale D, Ginsburg GS, et al. A genomic approach to colon cancer risk stratification yields biologic insights into therapeutic opportunities. *Proc Natl Acad Sci U S A.* 2008;105(49):19432–7.
- Huang L, Zhao S, Frasar JM, Dai Y. An integrated bioinformatics approach identifies elevated cyclin E2 expression and E2F activity as distinct features of tamoxifen resistant breast tumors. *PLoS One.* 2011;6(7):e22274.
- Wang G, Ye Y, Yang X, Liao H, Zhao C, Liang S. Expression-based in silico screening of candidate therapeutic compounds for lung adenocarcinoma. *PLoS One.* 2011;6(11):e24573.
- Sirota M, Dudley JT, Kim J, Chiang AP, Morgan AA, Sweet-Cordero A, Sage J, Butte AJ. Discovery and preclinical validation of drug indications using compendia of public gene expression data. *Sci Transl Med.* 2011;3(96):96ra77.
- Dudley JT, Sirota M, Shenoy M, Pai RK, Roedder S, Chiang AP, Morgan AA, Sarwal MM, Pasricha PJ, Butte AJ. Computational repositioning of the anticonvulsant topiramate for inflammatory bowel disease. *Sci Transl Med.* 2011;3(96):96ra76.
- Haibe-Kains B, El-Hachem N, Birkbak NJ, Jin AC, Beck AH, Aerts HJ, Quackenbush J. Inconsistency in large pharmacogenomic studies. *Nature.* 2013;504(7480):389–93.
- 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.
- Ihle NT, Williams R, Chow S, Chew W, Berggren MI, Paine-Murrieta G, Minion DJ, Halter RJ, Wipf P, Abraham R, et al. Molecular pharmacology and antitumor activity of PX-866, a novel inhibitor of phosphoinositide-3-kinase signaling. *Mol Cancer Ther.* 2004;3(7):763–72.
- Konecny GE, Wang C, Hamidi H, Winterhoff B, Kalli KR, Dering J, Ginther C, Chen HW, Dowdy S, Cliby W, et al. Prognostic and therapeutic relevance of molecular subtypes in high-grade serous ovarian cancer. *J Natl Cancer Inst.* 2014;106:10.
- Goode EL, DeRycke M, Kalli KR, Oberg AL, Cunningham JM, Maurer MJ, Fridley BL, Armasu SM, Serie DJ, Ramar P, et al. Inherited variants in regulatory T cell genes and outcome of ovarian cancer. *PLoS One.* 2013;8(1):e53903.
- Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics.* 2007;8(1):118–27.
- Houseman EA, Christensen BC, Yeh RF, Marsit CJ, Karagas MR, Wrensch M, Nelson HH, Wiemels J, Zheng S, Wiencke JK, et al. Model-based clustering of DNA methylation array data: a recursive-partitioning algorithm for high-dimensional data arising as a mixture of beta distributions. *BMC Bioinformatics.* 2008;9:365.
- Shoemaker RH, Monks A, Alley MC, Scudiero DA, Fine DL, McLemore TL, Abbott BJ, Paull KD, Mayo JG, Boyd MR. Development of human tumor cell line panels for use in disease-oriented drug screening. *Prog Clin Biol Res.* 1988;276:265–86.
- Eva A, Robbins KC, Andersen PR, Srinivasan A, Tronick SR, Reddy EP, Ellmore NW, Galen AT, Lautenberger JA, Papas TS, et al. Cellular genes analogous to retroviral onc genes are transcribed in human tumour cells. *Nature.* 1982;295(5845):116–9.
- Godwin AK, Meister A, O'Dwyer PJ, Huang CS, Hamilton TC, Anderson ME. High resistance to cisplatin in human ovarian cancer cell lines is associated with marked increase of glutathione synthesis. *Proc Natl Acad Sci U S A.* 1992;89(7):3070–4.
- Solit DB, Zheng FF, Drobnjak M, Munster PN, Higgins B, Verbel D, Heller G, Tong W, Cordón-Cardo C, Agus DB, et al. 17-Allylamino-17-demethoxygeldanamycin induces the degradation of androgen receptor and HER-2/neu and inhibits the growth of prostate cancer xenografts. *Clin Cancer Res.* 2002;8(5):986–93.
- Kamal A, Thao L, Sensintaffar J, Zhang L, Boehm MF, Fritz LC, Burrows FJ. A high-affinity conformation of Hsp90 confers tumour selectivity on Hsp90 inhibitors. *Nature.* 2003;425(6956):407–10.
- Sun W, Kalen AL, Smith BJ, Cullen JJ, Oberley LW. Enhancing the antitumor activity of adriamycin and ionizing radiation. *Cancer Res.* 2009;69(10):4294–300.
- Ammar H, Closser JL. Clusterin activates survival through the phosphatidylinositol 3-kinase/Akt pathway. *J Biol Chem.* 2008;283(19):12851–61.
- Shang H, Chen H, Zhao D, Tang X, Liu Y, Pan L, Cheng M. Synthesis and biological evaluation of 4alpha/4beta-imidazolyl podophyllotoxin analogues as antitumor agents. *Arch Pharm.* 2012;345(1):43–8.
- Gordaliza M. Natural products as leads to anticancer drugs. *Clin Transl Oncol.* 2007;9(12):767–76.
- Burkhardt RA, Peng Y, Norris ZA, Tholey RM, Talbott VA, Liang Q, Ai Y, Miller K, Lal S, Cozzitorto JA, et al. Mitoxantrone targets human ubiquitin-specific peptidase 11 (USP11) and is a potent inhibitor of pancreatic cancer cell survival. *Mol Cancer Res.* 2013;11(8):901–11.
- Bellosillo B, Colomer D, Pons G, Gil J. Mitoxantrone, a topoisomerase II inhibitor, induces apoptosis of B-chronic lymphocytic leukaemia cells. *Br J Haematol.* 1998;100(1):142–6.
- Mathivet P, Bernasconi R, De Barry J, Marescaux C, Bittiger H. Binding characteristics of gamma-hydroxybutyric acid as a weak but selective GABAB receptor agonist. *Eur J Pharmacol.* 1997;321(1):67–75.
- Richter JM, Schaefer M, Hill K. Clemizole hydrochloride is a novel and potent inhibitor of transient receptor potential channel TRPC5. *Mol Pharmacol.* 2014;86(5):514–21.
- Watanabe S, Tsuda M, Terada T, Katsura T, Inui K. Reduced renal clearance of a zwitterionic substrate cephalixin in MATE1-deficient mice. *J Pharmacol Exp Ther.* 2010;334(2):651–6.
- Bizat N, Hermel JM, Humbert S, Jacquard C, Creminon C, Escartin C, Saudou F, Krajewski S, Hantraye P, Brouillet E. In vivo calpain/caspase cross-talk during 3-nitropropionic acid-induced striatal degeneration: implication of a calpain-mediated cleavage of active caspase-3. *J Biol Chem.* 2003;278(44):43245–53.
- Sun N, Zang W, Li W. Bioinformatics analysis reveals potential candidate drugs for psychological stress in ovarian cancer. *Eur Rev Med Pharmacol Sci.* 2012;16(10):1362–6.
- Terry Jr AV, Hernandez CM, Hohnadel EJ, Bouchard KP, Buccafusco JJ. Cotinine, a neuroactive metabolite of nicotine: potential for treating disorders of impaired cognition. *CNS Drug Rev.* 2005;11(3):229–52.
- Wang K, Sun J, Zhou S, Wan C, Qin S, Li C, He L, Yang L. Prediction of drug-target interactions for drug repositioning only based on genomic expression similarity. *PLoS Comput Biol.* 2013;9(11):e1003315.
- Pujade-Lauraine E, Wagner U, Aavall-Lundqvist E, Gebusi V, Heywood M, Vasey PA, Volgger B, Vergote I, Pignata S, Ferrero A, et al. Pegylated



- liposomal Doxorubicin and Carboplatin compared with Paclitaxel and Carboplatin for patients with platinum-sensitive ovarian cancer in late relapse. *J Clin Oncol.* 2010;28(20):3323–9.
37. Monk BJ, Herzog TJ, Kaye SB, Krasner CN, Vermorken JB, Muggia FM, Pujade-Lauraine E, Lisyanskaya AS, Makhson AN, Rolski J, et al. Trabectedin plus pegylated liposomal Doxorubicin in recurrent ovarian cancer. *J Clin Oncol.* 2010;28(19):3107–14.
  38. Ferrandina G, Ludovisi M, Lorusso D, Pignata S, Breda E, Savarese A, Del Medico P, Scaltriti L, Katsaros D, Priolo D, et al. Phase III trial of gemcitabine compared with pegylated liposomal doxorubicin in progressive or recurrent ovarian cancer. *J Clin Oncol.* 2008;26(6):890–6.
  39. Alba E, Ruiz-Borrego M, Margeli M, Rodríguez-Lescure A, Sanchez-Rovira P, Ruiz A, Mel-Lorenzo JR, Ramos-Vazquez M, Ribelles N, Calvo E, et al. Maintenance treatment with pegylated liposomal doxorubicin versus observation following induction chemotherapy for metastatic breast cancer: GEICAM 2001-01 study. *Breast Cancer Res Treat.* 2010;122(1):169–76.
  40. O'Brien ME, Wigler N, Inbar M, Rosso R, Grischke E, Santoro A, Catane R, Kieback DG, Tomczak P, Ackland SP, et al. Reduced cardiotoxicity and comparable efficacy in a phase III trial of pegylated liposomal doxorubicin HCl (CAELYX/Doxil) versus conventional doxorubicin for first-line treatment of metastatic breast cancer. *Ann Oncol.* 2004;15(3):440–9.
  41. Bennett JM, Muss HB, Doroshow JH, Wolff S, Kremenz ET, Cartwright K, Dukart G, Reisman A, Schoch I. A randomized multicenter trial comparing mitoxantrone, cyclophosphamide, and fluorouracil with doxorubicin, cyclophosphamide, and fluorouracil in the therapy of metastatic breast carcinoma. *J Clin Oncol.* 1988;6(10):1611–20.
  42. Hagemeister F, Cabanillas F, Coleman M, Gregory SA, Zinzani PL. The role of mitoxantrone in the treatment of indolent lymphomas. *Oncologist.* 2005;10(2):150–9.
  43. Wiseman LR, Spencer CM. Mitoxantrone. A review of its pharmacology and clinical efficacy in the management of hormone-resistant advanced prostate cancer. *Drugs Aging.* 1997;10(6):473–85.
  44. Hande KR. Topoisomerase II inhibitors. Update on Cancer Therapeutics. 2008;3(1):13–26.
  45. Liu H, Xiao F, Serebriiskii IG, O'Brien SW, Maglaty MA, Astsaturov I, Litwin S, Martin LP, Proia DA, Golemis EA, et al. Network analysis identifies an HSP90-central hub susceptible in ovarian cancer. *Clin Cancer Res.* 2013;19(18):5053–67.
  46. Rose PG, Blessing JA, Mayer AR, Homesley HD. Prolonged oral etoposide as second-line therapy for platinum-resistant and platinum-sensitive ovarian carcinoma: a Gynecologic Oncology Group study. *J Clin Oncol.* 1998;16(2):405–10.
  47. Weng Z, Wang K, Li H, Shi Q. A comprehensive study of the association between drug hepatotoxicity and daily dose, liver metabolism, and lipophilicity using 975 oral medications. *Oncotarget.* 2015;6(19):17031–8.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

