

## ORIGINAL ARTICLES

# Evidence and potential clinical significance of changes in gene network interactions in ovarian cancer

Christopher G. Hill, John F. McDonald

Integrated Cancer Research Center, School of Biology and Parker H. Petit Institute of Bioengineering and Biosciences, Georgia Institute of Technology, Atlanta, USA.

**Correspondence:** John F. McDonald. Address: Georgia Institute of Technology, 315 Ferst Dr., Atlanta, USA. Email: john.mcdonald@biology.gatech.edu

**Received:** July 27, 2015

**Accepted:** September 5, 2015

**Online Published:** September 8, 2015

**DOI:** 10.5430/jbei.v2n1p1

**URL:** <http://dx.doi.org/10.5430/jbei.v2n1p1>

## Abstract

The molecular basis of cancer is not merely the consequence of structural and/or regulatory mutations in genes, but additionally to disruptions in networks of regulatory interactions existing among these genes and other components of the genome. Disruptions in network relationships may manifest as the loss, gain or reversal of functionally significant interactive gene relationships in cancer cells. In this study, we first employ an unsupervised (Pearson correlation) approach to quantitatively estimate the overall change in network relationships between precursor (control) ovarian surface epithelial cells and ovarian cancer epithelial cells. We find that ovarian cancer cells display a significant overall reduction in correlated gene network interactions relative to normal precursor cells reflective of an overall loss of regulatory control. We next focus on gene relationships that qualitatively change between normal and cancer samples. We find that biological processes significantly over represented among differentially expressed genes are substantially different from those associated with genes involved in qualitatively disrupted network interactions. Our findings provide novel insights into the processes underlying ovarian cancer and identify a potential new class of genes for targeted therapy.

## Key words

Gene network, Ovarian cancer, Targeted gene therapy

## 1 Introduction

It has become increasingly clear in recent years that the molecular basis of cancer and other complex human diseases is not merely attributable to structural and/or regulatory mutations in one or a few genes, but additionally to disruptions in networks of regulatory interactions existing among these and other components of the genome <sup>[1-4]</sup>. Although a determination of the myriad of potential genetic interactions that characterize human cells is only beginning to be established experimentally <sup>[2, 5]</sup>, a number of computational approaches have recently been developed that can directly or indirectly infer interactive network relationships among genes in normal and cancer cells <sup>[6-9]</sup>.

A standard input for gene regulatory network algorithms is gene expression (RNA-seq and/or microarray) datasets. Interactions (edges) between genes (nodes) can be inferred in an unsupervised fashion based upon consistent and highly

correlated changes (positive or negative) in the expression patterns that exist among genes in multiple biological samples [4, 10, 11]. In this study, we begin by employing such an unsupervised (Pearson correlation) approach to estimate overall gene network relationships between precursor (control) ovarian surface epithelial cells (OSE) [12] and serous papillary ovarian cancer epithelial cells (CEPI) isolated from human patient samples by laser capture microdissection (LCM) and analyzed by gene expression microarray (Affymetrix, U133 Plus 2) as previously described [13]. The results indicate that ovarian cancers display a significant overall reduction in correlated gene interactions relative to normal precursor cells.

We next focus on gene interactions that display a qualitative change between normal and cancer cells. For example, instances where the expression levels of genes positively correlate in normal cells but negatively correlate in cancer cells or *vice versa*. Such reversal of regulatory relationships between genes has been experimentally shown to be of particular functional significance in cancer [5], but it is a phenomenon that has not been extensively examined on the computational level. Our results indicate that nearly 45,000 correlated gene pairs display qualitative changes in their relationships in our cancer samples. Genes most frequently involved in these qualitative changes were found to be associated with biological processes distinct from those associated with genes most significantly differentially expressed in cancer. Collectively, our results indicate that the identification of genes involved in disrupted network interactions in ovarian and other cancers may not only provide new insights into processes underlying the disease but may also identify a new class of genes for targeted therapy.

## 2 Methods

### 2.1 Gene expression profiles

Expression profiles were generated in our lab [14] on ten normal OSE and ten serous CEPI samples (GEO accession: GSE52037) and analyzed for changes in mRNA expression using the Affymetrix Gene Chip Operating System (GCOS HG-U133 Plus 2.0). CEL files generated by GCOS were converted to expression values using GCRMA normalization on the ArrayAnalysis.org [15] website, that also included quality control metrics and cluster dendrograms.

Present/absent calls were generated from the MAS 5.0 statistical algorithm as implemented in Affymetrix Expression Console. Probe sets with > 60% present calls in either of the two groups (OSE and CEPI) were selected for further analysis. After log<sub>2</sub> transformation, the signal values of those probe sets were submitted to Statistical Analysis of Microarrays (SAM) for multiple testing correction where a 5.5% FDR was applied resulting in 7,461 probe sets representing 5,910 unique genes; annotations for probe sets were obtained from Affymetrix. Genes were further filtered for absolute fold change > 1.5 between normal and cancer samples, resulting in 5,144 differentially expressed (DE) genes. Further analysis of DE genes that were detected by exactly two probe sets revealed that ~2.5% of the probe sets reported contradictory results for a given gene; genes with contradictory probe sets were filtered out prior to network creation, leaving 5,070 DE genes.

### 2.2 Correlation calculation

Pearson correlations between DE genes were calculated using the Mathematica [16] correlation function. Since it is possible that the gene-to-gene correlation could be randomly generated, a significance test was conducted using expression signals from normal samples. For each Affymetrix probe set, expression signals from the ten normal samples were shuffled among those ten samples. That is, expression values were randomly reassigned to different samples, thus preserving the mean and standard deviation. By graphing the size of the largest connected subgraph for different values of “*r*” ranging from 0.70 to 0.99 (see Figure 1), it was found that networks of random signals could appear to be connected for values of  $r < 0.85$ . Thus, to minimize false positives, the absolute value of “*r*” was limited to values > 0.85. Baseline relationships

between pairs of DE genes were established by correlating the expression of 12.8 million pairs of genes across ten normal samples, and selecting pairs satisfying  $r > 0.85$ ;  $df = 8$ ,  $P < .0018$ . This resulted in a total of 682,194 highly correlated gene pairs in normal samples. Using this same method for cancer samples resulted in 193,276 highly correlated gene pairs.

### 2.3 Gene ontology enrichment analysis

Lists of 350 inconsistent (see results) and 350 highly DE genes were subjected to genome ontology enrichment analysis by using the Database for Annotation, Visualization and Integrated Discovery (DAVID)<sup>[17]</sup>. The lists were analyzed using Functional Annotation Clustering. Biological process terms having  $P < .05$  were considered significant.

### 2.4 Gene ontology selection

Cell cycle genes ( $n = 577$ ) (GO:0007049) and genes involved in the Regulation of Programmed Cell Death ( $n = 892$ ) (GO: 0043067) were downloaded from QuickGO<sup>[18]</sup> after filtering for Taxon (human: 9606), Evidence (Manual all), GO Identifier (GO: 0007049 or GO: 0043067 and their descendants), and Aspect (Biological Process). These genes formed the basis for Figures 2A and 2B that were built using Cytoscape<sup>[19]</sup>.

### 2.5 Graph construction and visualization

Cytoscape<sup>[19]</sup> was used to visualize the correlation networks consisting of DE genes downloaded from QuickGO, by using a Prefuse Force Directed layout. The top six most highly correlated neighbors for each gene were connected to form a graph. Genes with only one neighbor are not shown.

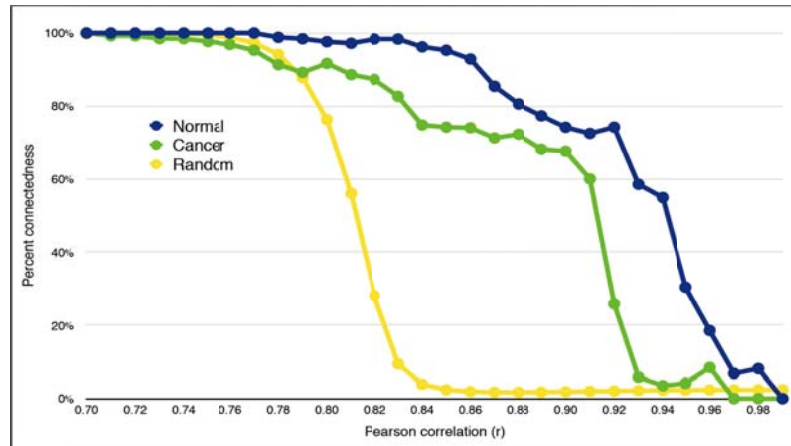
## 3 Results

### 3.1 Ovarian cancer is associated with a significant quantitative reduction in gene network interactions relative to normal precursor cells

We employed microarray gene expression analysis to compare differences in gene expression levels between precursor OSE cells and CEPI isolated from patient tissue samples by laser capture microdissection (LCM). Gene expression profiling identified 5,070 significantly differentially expressed genes (mRNAs) between the OSE and CEPI samples (Supplementary Table 1). Of these, 1,436 (28%) were significantly up regulated and 3,634 (72%) significantly down regulated in the cancer samples relative to normal controls.

Baseline relationships (Pearson correlations) between the significantly differentiated genes (5,070) were established in the control samples (OSE) by comparing the expression levels of 12.8 million pairs of genes ( $5,070 \times 5,069 / 2$ ) across ten normal patient samples and selecting gene pairs displaying highly correlated (positive or negative) expression levels across samples. This resulted in a total of 682,194 (~5% of all possible pairings) significantly correlated pairs of differentially expressed genes in the normal samples ( $r > 0.85$ ;  $P < .0018$ ; Note: we determined that  $r < 0.85$  can occur purely by chance, see Figure 1). Of these 682,194 correlated gene pairs, 567,886 were positively correlated with one another while 114,308 were negatively correlated. Conducting the same analysis across the ten cancer (CEPI) samples resulted in 193,276 highly correlated (164,283 positively correlated; 28,993 negatively correlated) gene pairs ( $r > 0.85$ ,  $P < .0018$ ).

Only 35,000 of the 682,194 significantly correlated gene pairs observed in normal (~5%) were found to maintain the same high level of correlation between one another in cancer samples. Thus, consistent with previous studies, our results indicate that the transition from normal to cancer cells involves a substantial loss of regulatory controls<sup>[20-22]</sup>.



**Figure 1.** Comparison of observed network connectivity in normal and cancer samples relative to that observed in randomly selected pairs of genes. Each point on the three lines of this graph represents the size of the largest connected sub-network as a percentage of the total number of correlated genes (percent connectedness) across ten samples, at that  $r$ -value ( $0.70 < r < 0.99$ ). The graph compares connectedness among correlated DE genes among normal samples (blue line), cancer samples (green line) and 500 randomly generated (yellow) expression values across different values of  $r$ . The results demonstrate that expression values associated with randomly selected pairs of genes can result in connected networks at  $r < 0.85$ .

### 3.2 Nearly 45,000 gene pair relationships identified in normal samples qualitatively changes in cancer samples

The above results indicate that a large number of gene pair interactions are significantly altered in ovarian cancer. We were next interested in focusing on a particular sub-class of altered interactions, *i.e.*, those instances where the correlated relationship between gene pairs in normal cells is qualitatively reversed in cancer cells. Molecular studies have previously shown that such reversals in regulatory interactions are often functionally significant in cancer [23-25] but this sub-class of qualitative changes has not been extensively analyzed on the computational level.

In order to quantify the number of interaction reversals in cancer, we categorized gene pair relationships as being either consistent or inconsistent. Gene pairs displaying the same qualitative relationships in normal and cancer samples, *i.e.*, gene pairs with expression patterns that are consistently positively or negatively correlated in both normal and cancer samples, are classified as “consistent”. Gene pairs with expression patterns that display qualitatively different relationships in the normal and cancer samples, *i.e.*, gene pairs that display positively correlated relationships in normal tissues but negatively correlated relationships in cancer tissues and *vice versa*, are classified as “inconsistent”. Of the 567,886 positively correlated gene pairs detected in the normal samples, 534,011 (94%) were found to maintain this same qualitative relationship in the cancer samples (consistent). Similarly, of the 114,308 negatively correlated gene pairs detected in normal samples, 103,277 (90%) were found to maintain this same qualitative relationship in cancer (consistent). Thus, a total of only 44,906 (6.6%) gene pairs displayed a reversal in gene pair relationships in cancer (“inconsistent”).

### 3.3 Genes associated with inconsistent gene pair relationships in cancer are significantly enriched for cell cycle and apoptotic related functions

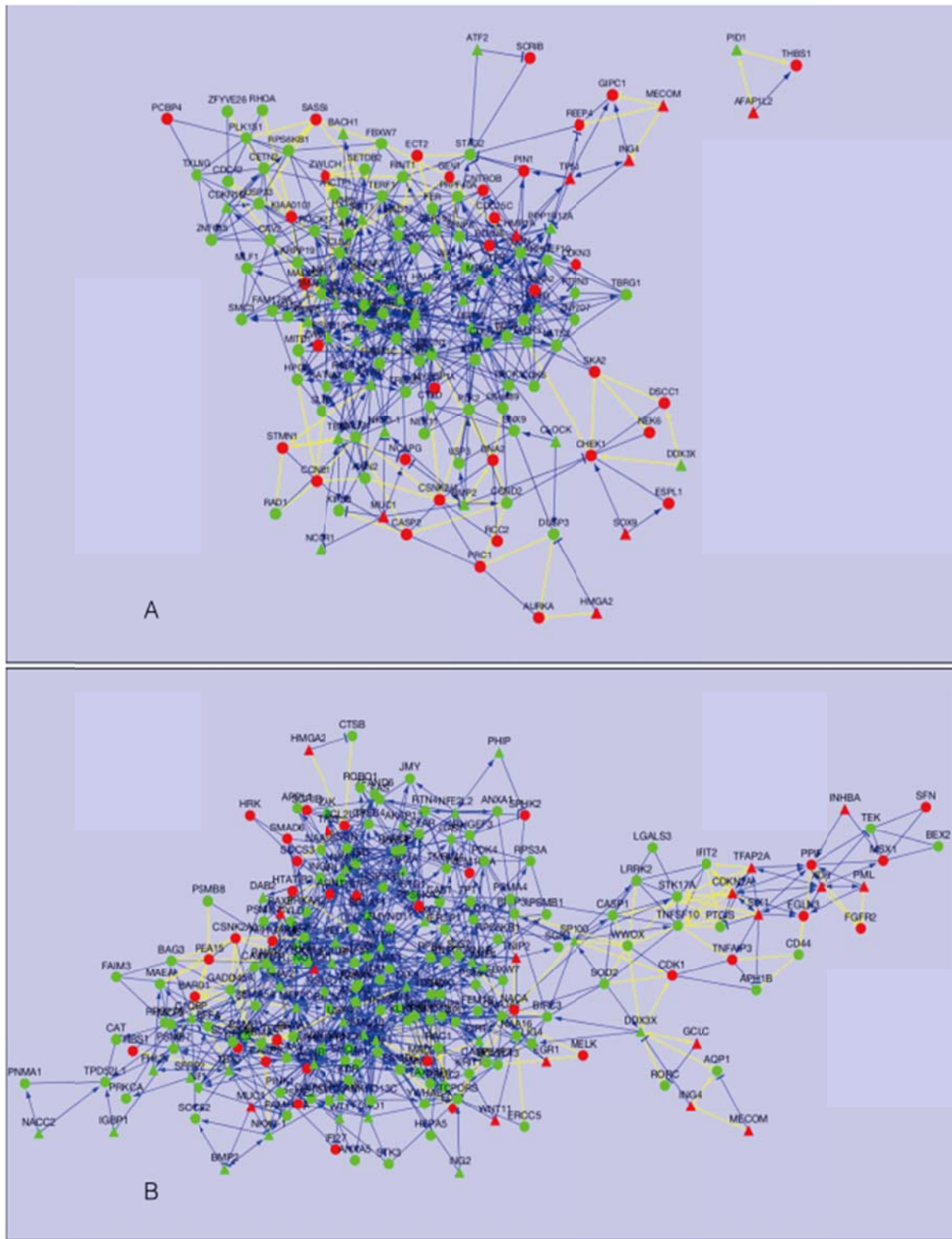
To determine if genes associated with inconsistent interactive relationships in our cancer samples were enriched for specific biological functions, we selected from the 44,906 pairs of genes involved in at least one inconsistent interaction, 350 genes associated with the greatest number of inconsistent interactions (Supplementary Table 2) and submitted them to DAVID [17] for pathway enrichment analysis. The results identified cell cycle related functions as the most significantly

enriched collection of biological processes in this group of genes followed closely by functions associated with programmed cell death (see Table 1, Figure 2 panel A and panel B).

**Table 1.** Genes involved in inconsistent gene pair interactions in ovarian cancer are most significantly functionally enriched for cell cycle and cell death related processes

Term	Gene Count	%	P-value
GO:0000279 M phase	28	8.12	9.38E-10
GO:0000280 nuclear division	21	6.09	3.15E-08
GO:0007067 mitosis	21	6.09	3.15E-08
GO:0022403 cell cycle phase	29	8.41	3.50E-08
GO:0000087 M phase of mitotic cell cycle	21	6.09	4.28E-08
GO:0007049 cell cycle	40	11.59	1.79E-07
GO:0022402 cell cycle process	32	9.28	6.83E-07
GO:0000278 mitotic cell cycle	25	7.25	7.66E-07
GO:0051301 cell division	21	6.09	3.26E-06
GO:0051276 chromosome organization	24	6.96	1.26E-04
GO:0007059 chromosome segregation	9	2.61	2.64E-04
GO:0000070 mitotic sister chromatid segregation	6	1.74	8.63E-04
GO:0000819 sister chromatid segregation	6	1.74	9.80E-04
GO:0051726 regulation of cell cycle	17	4.93	1.42E-03
GO:0033554 cellular response to stress	24	6.96	1.51E-03
GO:0006974 response to DNA damage stimulus	18	5.22	1.91E-03
GO:0007093 mitotic cell cycle checkpoint	6	1.74	1.96E-03
GO:0000075 cell cycle checkpoint	8	2.32	2.46E-03
GO:0007346 regulation of mitotic cell cycle	10	2.90	4.53E-03
GO:0010564 regulation of cell cycle process	8	2.32	9.63E-03
GO:0006916 anti-apoptosis	11	3.19	1.06E-02
GO:0007051 spindle organization	5	1.45	1.43E-02
GO:0006260 DNA replication	10	2.90	1.50E-02
GO:0006917 induction of apoptosis	14	4.06	1.59E-02
GO:0012502 induction of programmed cell death	14	4.06	1.63E-02
GO:0043068 positive regulation of programmed cell death	17	4.93	1.83E-02
GO:0010942 positive regulation of cell death	17	4.93	1.92E-02
GO:0007088 regulation of mitosis	5	1.45	2.79E-02
GO:0043067 regulation of programmed cell death	26	7.54	3.02E-02
GO:0010941 regulation of cell death	26	7.54	3.20E-02
GO:0043065 positive regulation of apoptosis	16	4.64	3.42E-02
GO:0006915 apoptosis	20	5.80	4.17E-02
GO:0012501 programmed cell death	20	5.80	4.70E-02

*Note.* As shown in Table 1, the 350 genes involved in the greatest number of inconsistent gene pair interactions were analyzed for biological process enrichment using DAVID<sup>[17]</sup>. Biological functions related to cell cycle (no fill) were found to be the most significantly enriched processes. Apoptosis and 11 other cell-death related functions (gray fill) were the second most enriched processes.



**Figure 2.** Networks of biological processes most significantly enriched among genes displaying qualitative changes in gene-pair relationships in ovarian cancer. Lines (edges) represent consistent (blue) or inconsistent (yellow) interactions between genes (nodes): (A) 577 cell cycle associated genes (GO category 0007049) with highly correlated ( $r > 0.85$ ) changes in interactions in ovarian cancer; (B) 892 cell death associated genes (GO category: 0043067) with highly correlated ( $r > 0.85$ ) changes in ovarian cancer. Red nodes are genes significantly up regulated in ovarian cancer; green nodes are genes significantly down regulated in ovarian cancer (triangles are transcription factors).

### 3.4 Genes associated with qualitatively disrupted network interactions in ovarian cancer identify a potential new class of candidates for targeted gene therapy

One molecular standard commonly used to identify functionally significant alterations in cancer is changes in gene expression<sup>[13, 26, 27]</sup>. To determine the extent to which processes identified as significantly overrepresented among genes involved in disrupted network interactions (see above) were similar to or different from processes overrepresented in genes most significantly differentially expressed, we selected the 350 (out of 5,070) most significantly differentially expressed genes between our normal and cancer samples (Supplementary Table 3) and subjected them to the same gene ontology analysis. As shown in Table 2, the biological processes most significantly overrepresented among the top 350 most significantly differentially expressed genes were substantially different from those most significantly overrepresented among genes involved with inconsistent interactions. Indeed, of the 350 genes most significantly differentially expressed between our normal and cancer samples, only 22 overlapped with the 350 genes involved in the most frequently inconsistent interactions (Supplementary Table 4). These results suggest that the identification and characterization of genes involved in inconsistent gene interactions in ovarian and perhaps other cancers may constitute a source of functionally significant information unavailable through gene expression analyses alone.

**Table 2.** Genes most significantly differentially expressed in ovarian cancer are most significantly functionally enriched for cell adhesion associated processes

Term	Gene Count	%	P-value
GO:0007155 cell adhesion	36	10.50	9.25E-07
GO:0022610 biological adhesion	36	10.50	9.60E-07
GO:0048732 gland development	13	3.79	2.44E-05
GO:0060429 epithelium development	15	4.37	1.22E-04
GO:0006928 cell motion	24	7.00	1.24E-04
GO:0043627 response to estrogen stimulus	10	2.92	3.49E-04
GO:0030182 neuron differentiation	21	6.12	7.17E-04
GO:0030855 epithelial cell differentiation	10	2.92	8.26E-04
GO:0010033 response to organic substance	29	8.45	1.19E-03
GO:0007411 axon guidance	9	2.62	1.41E-03
GO:0048812 neuron projection morphogenesis	13	3.79	1.44E-03
GO:0032989 cellular component morphogenesis	19	5.54	1.47E-03
GO:0009611 response to wounding	23	6.71	1.49E-03
GO:0009725 response to hormone stimulus	18	5.25	2.02E-03
GO:0001655 urogenital system development	9	2.62	2.28E-03
GO:0009719 response to endogenous stimulus	19	5.54	2.35E-03
GO:0048545 response to steroid hormone stimulus	12	3.50	2.45E-03
GO:0031175 neuron projection development	14	4.08	2.46E-03
GO:0050878 regulation of body fluid levels	10	2.92	2.75E-03
GO:0030204 chondroitin sulfate metabolic process	4	1.17	2.85E-03

*Note.* As shown in Table 2, the 350 genes with the greatest fold change between normal and cancer samples ( $FC < -9.21$  and  $9.27 < FC$ ) were analyzed for biological process enrichment using DAVID<sup>[17]</sup>. Cell adhesion and a variety of other biological processes were found to be the most significantly enriched.

### 3.5 Many of the genes associated with inconsistent interactions have been previously implicated in cancer

We found that many of the genes involved in inconsistent interactions in ovarian cancer have been previously associated with cancer (see Table 3). For example, *CHEK1* (checkpoint kinase 1), a regulator of DNA damage-induced cell cycle (G2/M) arrest<sup>[28]</sup> and associated with 187 inconsistent interactions in our cancer samples, has been previously identified as a potential target for ovarian and other cancer therapy<sup>[29]</sup>. Likewise, *MECOM* (MDS1 and EVI1 complex locus) has been shown to be commonly amplified in OC and has also been identified as a potential target for gene therapy<sup>[30,31]</sup>.

A number of other genes involved in inconsistent interactions in our OC samples, although not yet implicated in ovarian cancer, have been identified as important contributors to other types of cancer. For example, the gene involved in the greatest number of inconsistent interactions in OC (650), *VRK1* (vaccinia-related kinase-1), has been previously associated with cell cycle and cell death functions<sup>[32]</sup> and has recently been implicated in breast cancer<sup>[33]</sup>. Likewise, the *TET1* (ten-eleven translocation 1) gene that we found to be involved in 574 inconsistent interactions in OC, has been recently identified as a tumor suppressor gene associated with breast and prostate cancer<sup>[34]</sup>. Another tumor suppressor gene, *HELLS* (helicase lymphoid-specific) found to be involved in 576 inconsistent interactions in OC has been previously associated with epigenetic deregulation leading to lung cancer onset and/or progression<sup>[35]</sup>.

A number of the genes involved in inconsistent interactions in our OC samples, although not yet explicitly implicated in any human cancer, are nevertheless known to be involved in important cellular functions and thus may represent a currently unrecognized class of cancer associated genes. For example, *XPR1* (xenotropic and polytropic retrovirus receptor-1) is an atypical trans-membrane signaling receptor associated with G-protein coupled receptor activity<sup>[36]</sup>. Although *XPR1* is currently functionally recognized as a receptor for xenotropic and polytropic retroviruses, this is clearly not its only cellular function. The fact that *XPR1* is associated with 645 inconsistent interactions in our ovarian cancer samples suggests that it may play a significant, albeit as-yet-unrecognized role in ovarian cancer onset and/or progression.

**Table 3.** Examples of genes involved in inconsistent gene pair interactions identified in this study that have been previously associated with ovarian or other cancers (\* Based on previous results, the gene has been identified as a drug target in the Cancer Resource<sup>[53]</sup>)

Druggable*	Gene symbol	Fold change between control and OC samples	Number of inconsistent interactions	Cancer Type
Yes	VRK1	1.9	650	breast
Yes	XPR1	2.3	645	unknown
Yes	HELLS	2.2	576	NSCLC
Yes	TET1	1.6	574	prostate
Yes	CHEK1	4.0	187	ovarian
Yes	MECOM	18.6	20	ovarian

*Note.* VRK1= vaccinia-related kinase-1; XPR1= xenotropic and polytropic retrovirus receptor-1; HELLS= helicase lymphoid -specific; TET1= ten-eleven translocation 1; CHEK1= checkpoint kinase 1; MECOM= MDS1 and EVI1 complex locus; NSCLC=non-small cell lung carcinoma.

## 4 Discussion

Modern high-throughput sequencing and gene expression technologies are providing unprecedented insights into molecular genetic processes disrupted in cancer<sup>[37]</sup> and have led to the possibility of a personalized, targeted gene approach to cancer therapy<sup>[38]</sup>. However, since not all variants detected by high-throughput technologies may be of



functional significance, high-throughput sequencing and gene expression data must be coupled with the capability to distinguish functionally important or “driver” variants from “passenger” variants of little or no relevance. A major barrier to the attainment of this goal is the fact that the eukaryotic genome is a highly integrated network<sup>[39-41]</sup>. As a consequence, the biological significance of a genetic variant in any particular cancer can reasonably be expected to depend, to a greater or lesser extent, upon the functional status of other genes with which it interacts. The growing realization of the potential importance of gene interactions in a variety of diseases<sup>[9, 39, 40]</sup>, including cancer<sup>[5, 41]</sup> is spurring the development of new experimental and computational tools to identify and evaluate the significance of changes in gene network relationships in eukaryotic cells<sup>[14, 42-45]</sup>.

Interactions between genes are typically computationally inferred in an unsupervised fashion based upon consistent and highly correlated changes (positive or negative) in the expression patterns between genes in multiple biological samples<sup>[46-48]</sup>. Although, such unsupervised approaches typically infer network relationships exclusively from minimally processed gene expression data, more supervised methods (*e.g.*, Bayesian<sup>[49]</sup>, partial correlation<sup>[47]</sup>) incorporate additional supplemental information, such as, current understanding of specific gene regulatory relationships. While supervised methods may have advantages when inferring network changes in previously established, well-defined pathways, they can be less effective in detecting previously unrecognized interactions and/or in detecting important yet previously undefined system-wide changes in network structure. For example, supervised algorithms using Bayesian networks typically operate on directed acyclic graphs that do not include loops in the network. Such procedures can result in the loss of biologically relevant data since feedback loops are widely recognized as biologically significant in eukaryotic cells in general<sup>[50]</sup> and specifically in cancer<sup>[51, 52]</sup>.

In this study, we were interested in exploring the possible significance of changes in gene interactions in OC by comparing differences in correlated patterns of gene expression between precursor OSE cells and CEPI isolated from ten normal and ten cancer patient samples. Using an unsupervised computational approach (Pearson correlation), we first established evidence of a significant overall change in the number of gene-gene interactions in our cancer samples. These findings are consistent with earlier computational and experimental studies suggesting that cancers are generally associated with significant changes in regulatory control<sup>[20-22]</sup>.

While whole transcriptome analyses can determine if significant changes in overall gene-gene interactions occur during cancer development, more focused studies are required to uncover specific subclasses of gene interactions that may be of functional importance. For example, in a recent study of transcription factor gene interactions, a number of significant changes were detected between control and OC samples analyzed<sup>[47]</sup>. In our study, we were interested in focusing on a different sub-class of altered interactions, *i.e.*, those where the correlated relationship between gene pairs in normal cells is qualitatively reversed in cancer cells. The motivation for our analysis was the fact that recent molecular studies suggest that reversals in regulatory interactions may be of particular functional significance in cancer<sup>[23-25]</sup> and the fact that this sub-class of regulatory changes in cancer has not previously been extensively analyzed on the computational level.

We found that nearly 45,000 gene pairs displayed a reversal in gene pair relationships in cancer (“inconsistent”). Although a number of genes previously identified as being involved in disrupted transcription factor-target gene interactions in OC were identified in our study as being significantly differentially expressed in OC (*e.g.*, *Sp3*, *NFKB1*; Supplementary Table 1), none of these genes were among the genes most frequently involved in inconsistent gene pair relationships. This suggests that the majority of genes involved in qualitative reversals in interactive relationships in OC are not encoding transcription factors. Among the genes most frequently involved in a reversal in gene pair relationships in OC were several previously implicated in the development of other cancers (see Table 3). For example, the gene involved in the largest number (650) of reversed or inconsistent relationships, *VRK1*, is a member of the serine/threonine family of protein kinases and is a well-known “druggable” gene target<sup>[53]</sup>.

Finally, in an effort to identify biological functions potentially disrupted by reversals in gene pair relationships in OC, we selected the 350 genes most frequently involved in inconsistent gene pair relationships and subjected them to gene ontology analysis. The results indicated a significant enrichment for genes involved in cell cycle regulation and cell death. For comparison, we conducted a similar analysis for the 350 genes most significantly differentially expressed in our cancer samples relative to controls. There was little overlap among biological processes most significantly over-represented in these two datasets indicating that the identification of genes involved in inconsistent gene pair interactions in cancer may represent a clinically significant class of genes that can go undetected in standard DNA sequence and gene expression analyses. Indeed, a number of these genes identified to be involved in inconsistent interactions in OC have been previously identified as being of functional significance in other cancers and may represent promising new candidates for targeted gene therapy in OC.

## 5 Conclusions

The long-term hope of converting cancer from a lethal to a manageable chronic disease rests not only upon the availability of technologies to accurately detect structural and expression genetic variants associated with individual tumors, but analytical methods that can reliably identify variants that may be causally responsible for the cancer phenotype in individual patients. Our results are consistent with previous gene network analyses indicating that the simple identification of genes mutated and/or differentially expressed in cancer tissues may not be sufficient to identify all genes of functional significance. The development and implementation of novel methods to identify genes involved in disrupted network interactions in cancer promises to provide important new insights into the processes underlying cancer and the potential identification of a new class of genes for targeted gene therapy.

## Acknowledgements

Funding for this project was provided by the Ovarian Cancer Institute, Northside Hospital (Atlanta), The Deborah Nash Endowment Fund, The Josephine Robinson Family, and The J.D. Rhodes Trust.

## Conflict of Interest

The authors declare that they have no competing interests.

## Supplementary Information

Supplemental data are available at this McDonald Lab website:  
<http://www.mcdonaldlab.biology.gatech.edu/bioinformatics.htm>

## References

- [1] Ideker T, Galitski T, Hood L. A new approach to decoding life: systems biology. *Annu Rev Genomics Hum Genet.* 2001; 2: 343-72. PMID:11701654. <http://dx.doi.org/10.1146/annurev.genom.2.1.343>
- [2] Vidal M, Cusick ME, Barabasi AL. Interactome networks and human disease. *Cell.* 2011; 144: 986-98. PMID:21414488. <http://dx.doi.org/10.1016/j.cell.2011.02.016>
- [3] Wu C, Zhu J, Zhang X. Integrating gene expression and protein-protein interaction network to prioritize cancer-associated genes. *BMC Bioinformatics.* 2012; 13: 182. PMID:22838965. <http://dx.doi.org/10.1186/1471-2105-13-182>
- [4] Ying H, Lv J, Ying T, *et al.* Gene-gene interaction network analysis of ovarian cancer using TCGA data. *J Ovarian Res.* 2013; 6: 88. PMID:24314048. <http://dx.doi.org/10.1186/1757-2215-6-88>
- [5] Ashworth A, Lord CJ, Reis-Filho JS. Genetic interactions in cancer progression and treatment. *Cell.* 2011; 145: 30-8. PMID:21458666. <http://dx.doi.org/10.1016/j.cell.2011.03.020>
- [6] Laubenbacher R, Stigler B. A computational algebra approach to the reverse engineering of gene regulatory networks. *J Theor Biol.* 2004; 229: 523-37. PMID:15246788. <http://dx.doi.org/10.1016/j.jtbi.2004.04.037>

- [7] Schafer J, Strimmer K. An empirical Bayes approach to inferring large-scale gene association networks. *Bioinformatic*. 2005; 21: 754-64. PMID:15479708. <http://dx.doi.org/10.1093/bioinformatics/bti062>
- [8] Kar G, Gursoy A, Keskin O. Human cancer protein-protein interaction network: a structural perspective. *PLoS Comput Biol*. 2009; 5: e1000601. PMID:20011507. <http://dx.doi.org/10.1371/journal.pcbi.1000601>
- [9] Barabasi AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genetics*. 2010; 12: 56-68. PMID:21164525. <http://dx.doi.org/10.1038/nrg2918>
- [10] Roth RB, Hevezi P, Lee J, *et al*. Gene expression analyses reveal molecular relationships among 20 regions of the human CNS. *Neurogenetics*. 2006; 7: 67-80. PMID:16572319. <http://dx.doi.org/10.1007/s10048-006-0032-6>
- [11] Amar D, Safer H, Shamir R. Dissection of regulatory networks that are altered in disease via differential co-expression. *PLoS Comput Biol*. 2013; 9: e1002955. PMID:23505361. <http://dx.doi.org/10.1371/journal.pcbi.1002955>
- [12] Auersperg N. Ovarian surface epithelium as a source of ovarian cancers: unwarranted speculation or evidence-based hypothesis? *Gynecol Oncol*. 2013; 130: 246-51. PMID:23558054. <http://dx.doi.org/10.1016/j.ygyno.2013.03.021>
- [13] Bowen NJ, Walker LD, Matyunina LV, *et al*. Gene expression profiling supports the hypothesis that human ovarian surface epithelia are multipotent and capable of serving as ovarian cancer initiating cells. *BMC Med Genomics*. 2009; 2: 71. PMID:20040092. <http://dx.doi.org/10.1186/1755-8794-2-71>
- [14] Hill CG, Matyunina LV, Walker D, *et al*. Transcriptional override: a regulatory network model of indirect responses to modulations in microRNA expression. *BMC Syst Biol*. 2014; 8: 36. PMID:24666724. <http://dx.doi.org/10.1186/1752-0509-8-36>
- [15] Eijssen LM, Jaillard M, Adriaens ME, *et al*. User-friendly solutions for microarray quality control and pre-processing on ArrayAnalysis.org. *Nucleic Acids Res*. 2013; 41: W71-6. PMID:23620278. <http://dx.doi.org/10.1093/nar/gkt293>
- [16] Wolfram Research, Inc. Mathematica, Version 8.0, Champaign, IL: Wolfram Research, Inc. 2010.
- [17] Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2009; 37: 1-13. PMID:19033363. <http://dx.doi.org/10.1093/nar/gkn923>
- [18] Binns D, Dimmer E, Huntley R, *et al*. QuickGO: a web-based tool for Gene Ontology searching. *Bioinformatics*. 2009; 25: 3045-6. PMID:19744993. <http://dx.doi.org/10.1093/bioinformatics/btp536>
- [19] Shannon P, Markiel A, Ozier O, *et al*. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003; 13: 2498-504. PMID:14597658. <http://dx.doi.org/10.1101/gr.1239303>
- [20] Zhao M, Sun J, Zhao Z. Distinct and competitive regulatory patterns of tumor suppressor genes and oncogenes in ovarian cancer. *PLoS One*. 2012; 7: e44175. PMID:22952919. <http://dx.doi.org/10.1371/journal.pone.0044175>
- [21] Li L, Lorzadeh A, Hirst M. Regulatory variation: an emerging vantage point for cancer biology. *Wiley Interdisip Rev Syst Biol Med*. 2013; 6: 37-59. PMID:24254976. <http://dx.doi.org/10.1002/wsbm.1250>
- [22] Cordero D, Sole X, Crous-Bou M, *et al*. Large differences in global transcriptional regulatory programs of normal and tumor colon cells. *BMC Cancer*. 2014; 14: 708. PMID:25253512. <http://dx.doi.org/10.1186/1471-2407-14-708>
- [23] Remenyi A, Scholer HR, Wilmanns M. Combinatorial control of gene expression. *Nat Struct Mol Biol*. 2004; 11: 812-5. PMID:15332082. <http://dx.doi.org/10.1038/nsmb820>
- [24] Gordon S, Akopyan G, Garban H, *et al*. Transcription factor YY1: structure, function, and therapeutic implications in cancer biology. *Oncogene*. 2006; 25: 1125-42. PMID:16314846. <http://dx.doi.org/10.1038/sj.onc.1209080>
- [25] Thompson MR, Xu D, Williams BR. ATF3 transcription factor and its emerging roles in immunity and cancer. *J Mol Med*. 2009; 87: 1053-60. PMID:19705082. <http://dx.doi.org/10.1007/s00109-009-0520-x>
- [26] Moreno CS, Matyunina L, Dickerson EB, *et al*. Evidence that p53-mediated cell-cycle-arrest inhibits chemotherapeutic treatment of ovarian carcinomas. *PLoS One*. 2007; 2: e441. PMID:17505532. <http://dx.doi.org/10.1371/journal.pone.0000441>
- [27] Lili L, Matyunina L, Walker LD, *et al*. Molecular profiling supports the role of epithelial-to-mesenchymal transition (EMT) in ovarian cancer metastasis. *J Ovar Res*. 2013; 6: 49. PMID:23837907. <http://dx.doi.org/10.1186/1757-2215-6-49>
- [28] Bartek J, Lukas J. Chk1 and Chk2 kinases in checkpoint control and cancer. *Can Cell*. 2003; 3: 421-9. [http://dx.doi.org/10.1016/S1535-6108\(03\)00110-7](http://dx.doi.org/10.1016/S1535-6108(03)00110-7)
- [29] Bryant C, Rawlinson R, Massey AJ. Chk1 inhibition as a novel therapeutic strategy for treating triple-negative breast and ovarian cancers. *BMC Cancer*. 2014; 14: 570. PMID:25104095. <http://dx.doi.org/10.1186/1471-2407-14-570>
- [30] Ramakrishna M, Williams LH, Boyle SE, *et al*. Identification of candidate growth promoting genes in ovarian cancer through integrated copy number and expression analysis. *PLoS One*. 2010; 5: e9983. PMID:20386695. <http://dx.doi.org/10.1371/journal.pone.0009983>
- [31] Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature*. 2011; 474: 609-15. PMID:21720365. <http://dx.doi.org/10.1038/nature10166>
- [32] Valbuena A, Castro-Obregon S, Lazo PA. Downregulation of VRK1 by p53 in response to DNA damage is mediated by the autophagic pathway. *PLoS One*. 2011; 6: e17320. PMID:21386980. <http://dx.doi.org/10.1371/journal.pone.0017320>

- [33] Molitor TP, Traktman P. Molecular genetic analysis of VRK1 in mammary epithelial cells: depletion slows proliferation in vitro and tumor growth and metastasis in vivo. *Oncogenesis*. 2013; 2: e48. PMID:23732708. <http://dx.doi.org/10.1038/oncsis.2013.11>
- [34] Hsu CH, Peng KL, Kang ML, *et al.* TET1 suppresses cancer invasion by activating the tissue inhibitors of metalloproteinases. *Cell Reports*. 2012; 2: 568-79. PMID:22999938. <http://dx.doi.org/10.1016/j.celrep.2012.08.030>
- [35] Yano M, Ouchida M, Shigematsu H, *et al.* Tumor-specific exon creation of the HELLS/SMARCA6 gene in non-small cell lung cancer. *Int J Cancer*. 2004; 112: 8-13. PMID:15305370. <http://dx.doi.org/10.1002/ijc.20407>
- [36] Vaughan AE, Mendoza R, Aranda R, *et al.* Xpr1 is an atypical G-protein-coupled receptor that mediates xenotropic and polytropic murine retrovirus neurotoxicity. *J Virol*. 2012; 86: 1661-9. PMID:22090134. <http://dx.doi.org/10.1128/JVI.06073-11>
- [37] Yates LR, Campbell PJ. Evolution of the cancer genome. *Nature Rev Genetics*. 2012; 13: 795-806. PMID:23044827. <http://dx.doi.org/10.1038/nrg3317>
- [38] Dachs GU, Dougherty GJ, Stratford IJ, *et al.* Targeting gene therapy to cancer: a review. *Oncol Res*. 1997; 9: 313-25. PMID:9406237.
- [39] Barabasi AL, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nature Revs Genetics*. 2004; 5: 101-13. PMID:14735121. <http://dx.doi.org/10.1038/nrg1272>
- [40] Barabasi AL. Network medicine--from obesity to the "diseasome". *N Engl J Med*. 2007; 357: 404-7. PMID:17652657. <http://dx.doi.org/10.1056/NEJMe078114>
- [41] Kim D, Li R, Dudek SM, *et al.* Knowledge-driven genomic interactions: an application in ovarian cancer. *BioData Mining*. 2014; 7: 20. PMID:25214892. <http://dx.doi.org/10.1186/1756-0381-7-20>
- [42] Djebbari A, Quackenbush J. Seeded Bayesian Networks: constructing genetic networks from microarray data. *BMC Syst Biol*. 2008; 2: 57.
- [43] Laubenbacher R, Jarrah AS. Algebraic models of biochemical networks. *Methods Enzymol*. 2009; 467: 163-96. [http://dx.doi.org/10.1016/S0076-6879\(09\)67007-5](http://dx.doi.org/10.1016/S0076-6879(09)67007-5)
- [44] Nagaraj S, Reverter A. A Boolean-based systems biology approach to predict novel genes associated with cancer: Application to colorectal cancer. *BMC Syst Biol*. 2011; 5: 35. PMID:21352556. <http://dx.doi.org/10.1186/1752-0509-5-35>
- [45] Moreau Y, Tranchevent LC. Computational tools for prioritizing candidate genes: boosting disease gene discovery. *Nature Rev Genetics*. 2012; 13: 523-36. PMID:22751426. <http://dx.doi.org/10.1038/nrg3253>
- [46] Sonesson C, Lilljebjorn H, Fioretos T, *et al.* Integrative analysis of gene expression and copy number alterations using canonical correlation analysis. *BMC Bioinformatics*. 2010; 11: 191. PMID:20398334. <http://dx.doi.org/10.1186/1471-2105-11-191>
- [47] Madhamshettiwar PB, Maetschke SR, Davis MJ, *et al.* Gene regulatory network inference: evaluation and application to ovarian cancer allows the prioritization of drug targets. *Genome Med*. 2012; 4: 41. PMID:22548828. <http://dx.doi.org/10.1186/gm340>
- [48] Stephens PJ, Tarpey PS, Davies H, *et al.* The landscape of cancer genes and mutational processes in breast cancer. *Nature*. 2012; 486: 400-4. <http://dx.doi.org/10.1038/nature11017>
- [49] Kang J, Zheng W, Li L, *et al.* Use of Bayesian networks to dissect the complexity of genetic disease: application to the Genetic Analysis Workshop 17 simulated data. *BMC Proc*. 2011; 9: S37. PMID:22373110. <http://dx.doi.org/10.1186/1753-6561-5-S9-S37>
- [50] Pomerening JR. Positive-feedback loops in cell cycle progression. *FEBS Lett*. 2009; 583: 3388-96. PMID:19818353. <http://dx.doi.org/10.1016/j.febslet.2009.10.001>
- [51] Harris SL, Levine AJ. The p53 pathway: positive and negative feedback loops. *Oncogene*. 2005; 24: 2899-908. PMID:15838523. <http://dx.doi.org/10.1038/sj.onc.1208615>
- [52] Su S, Liu Q, Chen J, *et al.* A positive feedback loop between mesenchymal-like cancer cells and macrophages is essential to breast cancer metastasis. *Cancer Cell*. 2014; 25: 605-20. PMID:24823638. <http://dx.doi.org/10.1016/j.ccr.2014.03.021>
- [53] Ahmed J, Meinel T, Dunkel M, *et al.* Cancer Resource: a comprehensive database of cancer-relevant proteins and compound interactions supported by experimental knowledge. *Nucleic Acids Res*. 2011; 39: D960-7. PMID:20952398. <http://dx.doi.org/10.1093/nar/gkq910>