Network Analysis of Differential Gene Expression to Identify Hub Genes in Ovarian Cancer

Akram Siavoshi[1], Mahdieh Taghizadeh[2], Elahe Dookhe[3], Mehran Piran[4], Mahsa Saliani[5], Shahla Mohammad Ganji[6*]

[1] *Department of Biology, Sinabioelixir Group, Alborz Health Technology Development Center, Tehran, Iran*
[2] *Department of Medical Genetic, Tarbiat Modares University, Tehran, Iran*
[3] *Department of Biology, Research and Science Branch, Islamic Azad University, Tehran, Iran*
[4] *Department of Medical Biotechnology, Drug Design and Bioinformatics Unit, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran*
[5] *Department of Chemistry, Faculty of Science, Ferdowsi University of Mashhad, Mashhad, Iran*
[6] *Institute of Medical Biotechnology, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran*

Abstract

Epithelial ovarian cancer (EOC), as a challenging disease among women with poor prognosis and unclear molecular pathogenesis, each year is responsible for 140000 deaths globally. Recent progress in the field revealed the importance of proteins as key players of different biological events. Considering the complicated protein interactions, taking a deeper look at protein-protein interactions (PPIs) could be considered as a superior strategy to unravel complex mechanisms encountered with regulatory cell signaling pathways of ovarian cancer. Hence, PPI network analysis was performed on differentially expressed genes (DEGs) of ovarian cancer to discover hub genes which have the potential to be introduced as biomarkers with clinical utility. A PPI network with 600 DEGs was constructed. Network topology analysis determined UBC, FN1, SPP1, ACTB, GAPDH, JUN, and RPL13A, with the highest Degree (K) and betweenness centrality (BC), as shortcuts of the network. KEGG pathway analysis showed that these genes are commonly enriched in ribosome and ECM-receptor interaction pathways. These pivotal hub genes, mainly UBC, FN1, RPL13A, SPP1, and JUN have been reported previously as potential prognostic biomarkers of different types of cancer. However, further experimental molecular studies and computational processes are required to confirm the function and association of the identified hub genes with epithelial ovarian cancer prognosis.

Introduction

Epithelial ovarian cancer (EOC) as a challenging disease is diagnosed in nearly a quarter of a million women, and it is responsible for 140000 deaths worldwide per annum (Krzystyniak et al., 2016; Torre et al., 2018). Lack of early diagnosis and empirically-validated treatments were considered as the most common causes of mortality (Cho et al., 2015). During the past decade, extensive research has been conducted to identify methods to predict and evaluate cancer progression (Li et al., 2015; Loghmani et al., 2014). Currently, the use of biomarkers such as serum cancer antigen 125 (CA125) and human epididymis protein 4 (HE4) is very common among all methods used to diagnose, prognose, and management of ovarian cancer (Archana et al., 2013). However, the potential of these biomarkers for efficient prediction of outcome remains a significant challenge regardless of the different stages and complexity of the disease. Thus, the survival rate of EOC is still low, and there has not been any remarkable success in treatment, especially in patients with advanced epithelial ovarian cancer.

Recently, some researchers proposed some target genes with specific coverage of a determined stage of the disease (Arnedos et al., 2019; Li et al., 2018; Zhang et al., 2019). In this regard, differentially expressed gene analysis (DEGA) (Anders et al., 2010) as the most important application of RNA-Seq experiments, can be used to compare different genes expression levels between normal and cancerous cells. The results of such analyses reveal a list of differentially expressed genes (DEGs). Although, as most human cancers are very complex, and are

---

[*]Corresponding author's e-mail address: *Shahla@nigeb.ac.ir*

encountered with sets of genes and their complicated interactions, identification of the exact molecular mechanism is very difficult, especially, in human cases.

Investigating protein-protein interactions (PPIs), with key roles in the biological function of the cells, is one of the beneficial methods to discover complex molecular mechanisms which are responsible for cell signaling and cell to cell communications (Huang et al., 2016). In a previous study, PPI networks were created based on the DEGs analysis to discover hub genes that have the potential to be introduced as biomarkers of esophageal squamous cell carcinoma (Wu et al., 2014).

In the present study, RNA-seq data obtained from normal and cancerous cells of the ovarian tissue were compared statistically to discover DEGs. Then, Systems biology analyses such as gene ontology (GO) and pathways analysis (KEGG) (Ashburner et al., 2000; Kanehisa et al., 2000) were performed to provide insights into the key cellular processes which are responsible for normal/diseased condition. Furthermore, to explore new biomarkers, the PPI network was constructed by mapping all determined DEGs to the network data. Functional enrichment analysis was performed to assign functional categories to the subnetworks of genes.

By introducing several hub genes, results of the present study may facilitate our vision regarding the molecular mechanisms involved in ovarian cancer pathogenesis. These experiments could be effective for defining proper treatment strategies in the clinical settings. However, further confirmatory studies are required for validation of data and announcement of novel panels of genes.

Materials and Methods

RNA sequencing data processing and differential gene expression analysis

Three separate Fastq data files for untreated ovarian tumor cell line SKOV-3, and normal cell line FT194 (De Cristofaro et al., 2016) were retrieved from the sequence read archive (SRA; http://www.ncbi.nlm.nih.gov/geo/). In order to provide clean data for downstream analyses, quality filtration was conducted to omit low quality sequence reads (more than 30% of reads) and adaptors (the first 15bp of Illumina reads) by the Trimmomatic program (Bolger et al., 2014). Then, using the HISAT2 alignment program (Kim et al., 2015) all clean reads were mapped to the Homo sapiens (human) genome assembly GRCh37 (hg19). Counting of transcripts (mapping efficiencies

(95%)) was performed with HTSeq (Anders et al., 2014). Count data normalization was performed to determine statistically significant DEGs across two conditions. The DESeq2-Bioconductor package (version 1.6.3) was applied to improve stability and interpretability of estimates. Adjusted P value<0.01 and a |log2FC|>2 were defined as the cut-off criteria. Biological significance of DEGs was explored through GO term enrichment analysis including biological processes (BP), cellular components (CC), and molecular functions (MF), and then KEGG pathway enrichment analysis was performed using enricheR-Bioconductor package (version 2.1).

PPI Network construction

The search tool for the retrieval of interacting genes (STRING, https://string-db.org/cgi/input.pl; version: 11.0), was used for obtaining direct (physical) and indirect (functional) human PPI networks (PPIN). The attribute that we applied to construct network was based on the highest confidence score of 0.07. Then, the constructed PPI network was analyzed using Cytoscape (version 3.7.0). The topological analysis of the PPI network was performed with the Network Analyzer. Betweenness centrality (BC), closeness centrality (CC), and degree (K) were considered as fundamental parameters during our experiments to determine node properties.

Identification of modules and functional annotation analysis

The Molecular complex detection (MCODE) plugin was applied to visualize the significant gene modules in EOC with default parameters and the maximum depth of 100. Selection criteria for top 3 significant modules were set as follows: MCODE scores ≥ 6, and number of nodes ≥ 10. Functional enrichment analysis for each module was performed using g:Profiler (URL: http://biit.cs.ut.ee/gprofiler/).

Results

Differential gene expression analysis results

Distribution of expression values across samples before and after normalization was evaluated to ensure that expression values were similar across normalized counts.

Complete plot of raw counts using log2 transformation (log2 (Non-normalized counts+1)) and then a plot of normalized counts using the DESeq2 are shown in Figure 1 (A) and (B), respectively.
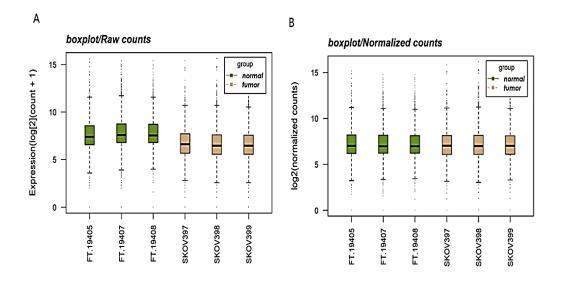
2

Figure 1. The gene count distributions. Box plots of non-normalized counts (log2 (counts+1)) per sample (A), and normalized counts (log2 (normalized counts)) per sample (B) are shown. The x-axis represents samples and the Y-axis represents log2 (counts +1).

After DE analysis between ovarian tumor and normal groups, 1000 DEGs (padj <0.01) were obtained with 232 upregulated genes (padj <0.01, log2 FC> 2), and 324 downregulated genes (padj <0.01, log2 FC <-2) (Table S1).

Functional analysis of DEGs

Gene ontology (GO) and pathway analysis of differentially expressed genes

To investigate activated and suppressed DEGs in different functional categories, GO and KEGG pathway analysis using enrichR were performed. Based on the results from these experiments genes were classified into different functional categories according to the GO term for biological processes (BP) (Figure. S1), molecular functions (MF) (Figure. S2) and cellular components (CC) (Figure. S3). Totally 572 out of 1000 profiled DEGs assigned to 930 GO terms (padj <0.01). The top 1 significantly upregulated and downregulated GO categories are shown in Table. 1.

Overall, 156 upregulated genes (padj<0.01, log2 FC >2), and 82 downregulated genes (padj<0.01, log2 FC <-2) were mapped to 283 KEGG pathways. The top 14 enriched pathways are shown in Figure. 2. The upregulated genes were highly clustered in signaling pathways including glycolysis, pyruvate metabolism, tryptophan metabolism, and fatty acid degradation; while, the most downregulated genes were highly clustered into ribosome, salmonella infection, focal adhesion, and apoptosis.

PPI Network construction

After DEGA, the significant result of String analyses was based on confidence score (0.007), the average degree of nodes (5.11), and average local clustering coefficient (0.406), and a PPI network with 797 interactions between 600 DEGs was performed (Figure. 3). In order to detect the key parameters of the network, interaction pairs of the PPI network were visualized by Network-Analyzer Cytoscape plugin (cut off values: BC> 0.02, and K> 10)(Table. 2).

Identification of modules and functional annotation analysis

The module analysis of PPI network using MCODE resulted in 13 modules. According to the Table S2 and Figure 4, four significant modules were identified with MCODE (score ≥ 5 and nodes ≥ 6). Among which UBC (Ubiquitin C) as the main hub was clustered in module 3. Ribosomal Protein Small (RPS) subunit genes and Ribosomal Protein Large (RPL) genes were clustered in module 1, and other hub genes including Secreted Phosphoprotein 1 (SPP1), calumenin (CALU), complement C3 (C3), and Fibronectin 1 (FN1) were clustered in module 2.

3

Table 1. The top 1 enriched gene ontology term of up- and down regulated genes involved in biological processes (BP), cellular components (CC), and molecular functions (MF)

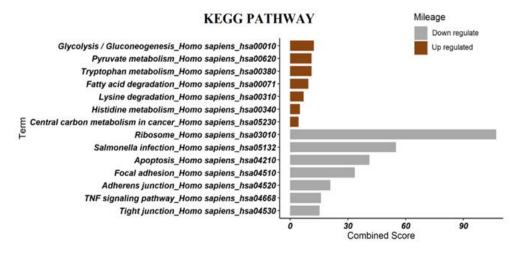| Gene Ontology (GO) terms and ID	Source	Adjusted p-value	Gene symbol
Up-regulated genes			
drug transport (GO:0015893)	BP	0.0197973	SLC47A2; SLC19A1
solute:sodium symporter activity (GO:0015370)	MF	0.06705	SLC5A9; SLC25A22
intrinsic component of the cytoplasmic side of the plasma membrane (GO:0031235)	CC	0.0109396	MIEN1; SPTB
Down-regulated genes			
SRP-dependent cotranslational protein targeting to membrane (GO:0006614)	BP	1.78E-10	RPL41;RPL3;RPL32;RPL13A;RPS25;RPS19;RPL36; RPL14;RPL13;RPL37;RPL26;RPL29;RPS24;RPL19
RNA binding (GO:0003723)	MF	1.56E-13	RBM25;RPL3;RPL32;HMGB2;PSIP1;YBX1;IFIT3;R PS19;RPL36;HIST1H1D;KIF1C;RPL37;HMGN2;HIS T1H1B;HIST1H1C;CAST;DDX58;ACTN1;DNTTIP2 ;RPL13A;PPHLN1;GNL2;GTF2F1;SMC1A;RANGA P1;EEF1D;MYH9;LUC7L3;RPL26;SREK1;RPL29;E ZR;PLEC;DHX8;SRRT;DDX21;PDCD11;TERT;PES 1;UBC;RPL14;RPL13;FLNA;FLNB;SRSF11;RPL19; RBM39;PRPF38B;RPL41;JUN;KRR1;PRRC2C;DEK; EEF2;RPS25;H1F0;MYBBP1A;ACO1;VIM;CALR;R PS24;WRAP53
cytosolic large ribosomal subunit (GO:0022625)	CC	5.33E-10	RPL41;RPL3;RPL32;RPL36;RPL14;RPL13A;RPL13; RPL37;RPL26;RPL29;RPL19

Figure 2. Pathway enrichment analysis of up- and downregulated genes in ovarian cancer samples in comparison to normal cases. The x-axis represents pathways, and KEGG IDs and the Y-axis represents combined score. As shown, the most downregulated genes are enriched in the ribosome pathway and the most upregulated genes are enriched in the Glycolysis/Gluconeogenesis pathway.
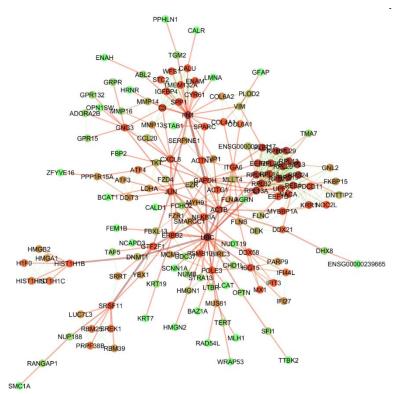
Figure 3. Overview of the PPI network constructed using Cytoscape. The network includes 547 edges (interactions) among 63 nodes. The nodes with dark brown, light brown, and green colors represent key genes in the network. Among key genes, nodes with dark brown color represent the super hubs with the highest BC and K.

Table 2. The main topological parameters including, Betweenness centrality (BC), closeness centrality (CC), and Degree (K) of the PPI network. The hub genes in the network based on cut off values of BC > 0.02 and degree >10 were demonstrated with light gray.

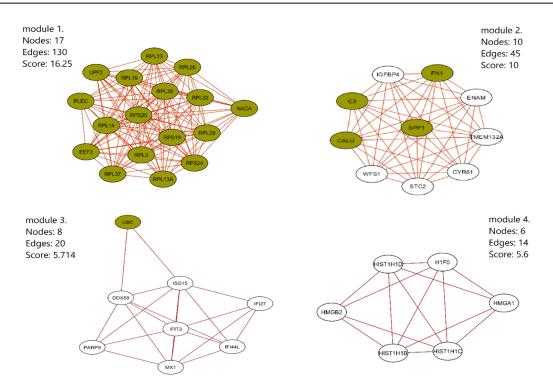| Gene	K	BC	CC
UBC	27	0.39930027	0.40957447
FN1	23	0.22956964	0.35240275
ACTB	17	0.21966548	0.39896373
GAPDH	13	0.19844427	0.40633245
JUN	12	0.12427487	0.36842105
RPL13A	23	0.1023342	0.38118812
CXCL8	11	0.09960491	0.34684685
RPL19	21	0.05464691	0.36150235
RPL32	20	0.04926171	0.35981308
C3	13	0.02902674	0.28308824
PLEC	16	0.02815066	0.30985915
PES1	17	0.02631353	0.29222011

5

Figure 4. Subnetworks identified from the PPI network (Module 1, Module 2, Module 3, and Module 4, respectively). The light green nodes in each of the clusters represent hub genes which were extracted from the PPI network by Network-Analyzer plugin. The white nodes represent genes which are involved in modules. The lines represent node's interactions.

Functional annotation analysis was applied to the hub genes of the PPI Network and each module separately. The top three functional annotation categories (BB, MF, CC, and KEGG) for module 1 are shown in Table. 3. Pathway analysis mainly involved KEGG pathway revealed that genes were commonly enriched in the ribosome pathway. The constituent structures of the ribosome, nuclear-transcribed mRNA catabolic process, nonsense-mediated mRNA decay, and cytosolic ribosome were the most related terms to MF, BP, and CC of the module 1 with the most enriched gene, respectively.

Table 3. The top three functional annotation categories (BB, MF, CC, and KEGG) for module 1 with the most enriched gene.

Source	GO/Pathway ID	Term name	Adjusted p-value	Intersections
MF	GO:0003735	structural constituent of ribosome	4.98E-19	RPL32,RPS24,RPL26,RPL36,RPS19,RPL3,RPL14,RPL37,RPL19,RPL13A,RPL29,RPL13
MF	GO:0005198	RNA binding	4.26E-13	RPL32,RPS24,RPL26,RPL36,RPS19,RPL3,RPL14,RPL37,RPL19,PLEC,RPL13A,RPL29,RPL13
MF	GO:0003723	structural molecule activity	7.63E-13	RPL32,RPS24,RPL26,RPL36,RPS19,RPL3,RPL14,RPL37,RPL19,RPS25,PLEC,RPL13A,RPL29,UPF2,RPL13,EEF2
BP	GO:0000184	nuclear-transcribed mRNA catabolic process, nonsense-mediated mRNA decay	2.49E-25	RPL32,RPS24,RPL26,RPL36,RPS19,RPL3,RPL14,RPL37,RPL19,RPS25,RPL13A,RPL29,UPF2,RPL13

6

BP	GO:00066 14	SRP-dependent cotranslational protein targeting to the membrane	5.23E-24	RPL32,RPS24,RPL26,RPL36,RPS19,RPL3,RPL14,RPL37,RPL19,RPS25,RPL13A,RPL2 9,RPL13
BP	GO:00066 13	cotranslational protein targeting to the membrane	1.06E-23	RPL32,RPS24,RPL26,RPL36,RPS19,RPL3,RPL14,RPL37,RPL19,RPS25,RPL13A,RPL2 9,RPL13
CC	GO:00226 26	cytosolic ribosome	3.23E-15	NACA,RPL32,RPS24,RPL26,RPL36,RPS19,RPL3,RPL14,RPL37,RPL19,RPS25,RPL13 A,RPL29,RPL13,EEF2
CC	GO:00058 40	Ribosome	9.54E-14	RPL32,RPS24,RPL26,RPL36,RPS19,RPL3,RPL14,RPL37,RPL19,RPS25,RPL13A,RPL2 9,RPL13
CC	GO:00443 91	ribosomal subunit	7.20E-13	RPL32,RPS24,RPL26,RPL36,RPS19,RPL3,RPL14,RPL37,RPL19,RPS25,RPL13A,RPL2 9,RPL13,EEF2
KEG G	KEGG:03 010	Ribosome	2.85E-19	RPL32,RPS24,RPL26,RPL36,RPS19,RPL3,RPL14,RPL37,RPL19,RPS25,RPL13A,RPL2 9,RPL13

Discussion

Epithelial ovarian cancer (EOC) has the highest mortality rate among different types of women's cancers due to the poor diagnosis (Hao et al., 2010). Studies have shown that in order to achieve effective methods for early diagnosis and prevention of metastasis, it is important to study the molecular mechanisms of the carcinogenesis.

The purpose of the current study is the analysis of existing RNA-seq data and their comparative interpretation between normal and diseased conditions to investigate novel DEGs involved in PPI Networks and regulatory pathways of EOC. Many pivotal genes and pathways which are associated with ovarian cancer were identified in the present study. Totally, among 1000 DEGs (232 upregulated and 324 downregulated genes), migration and invasion enhancer 1 (MIEN1) and AP001610.5 were the most up- and down-regulated genes, respectively. MIEN1 is an intrinsic component of the cytoplasmic side of the plasma membrane, which plays a pivotal role in the regulation of apoptosis. It was previously proposed by other studies as an important target to be considered in molecular cancer therapy procedures (Evans et al., 2006). The Ribosome pathway and Glycolysis/Gluconeogenesis were also identified as the most significantly enriched pathways in KEGG analysis. ALDH3A2 (Marcato et al, 2011), a member of the aldehyde dehydrogenase (ALDH) gene family (Warburg, 1956), is the most significant gene in Glycolysis/Gluconeogenesis pathway. Whereas, high glycolysis in tumor cells correlates with the degree of tumor malignancy, an argument to justify the significance of the glycolysis pathway in this study is the potential need of chronic cell proliferation to provide energy in order to fuel rapid cell growth and division (Board et al., 1992). In the present study, ribosomal protein (Rps) genes and large ribosomal proteins (RPL) including RPL41, RPL3, RPL32, RPL13A, RPS25, RPS19, RPL14 and RPL36 were the most significant downregulated DEGs. These genes related to the signal-recognition particle (SRP)-dependent cotranslational protein-membrane targeting, RNA binding, and cytosolic large ribosomal subunit.

After analyzing the topology of the PPI network totally 28 nodes with BC> 0.02, and K> 10 were extracted as hub genes and among them some nodes such as UBC, FN1, ACTB, GAPDH, JUN, and RPL13A with high K, BC, and CC were shortlisted. These hub genes were downregulated in EOC samples in comparison to adjacent normal samples. Clustering was performed to investigate the relationship between hub genes with other genes of the network using MCODE and previously identified 13 modules. Furthermore, functional annotation was performed on 4 modules to determine the top affected functions in EOC. Functional annotation of main hubs clustered in these modules showed that Ubiquitin C (UBC), with the highest degree of connectivity, was clustered in module 3 along with other genes, including IFI44L, IFI27, DDX58, IFIT3, PARP9, MX1, and ISG15. This module mainly enriched with protein tag, defense response, and RIG-I-like receptor signaling pathways.

The ubiquitin is encoded by the ubiquitin C (UBC) and ubiquitin B (UBB) in humans. These two genes are essential for maintenance of cellular ubiquitin

7

levels under stress conditions (Castello et al., 2017). Moreover, they play key roles as tumor suppressors in a variety of cancers, DNA damage repair and regulation of protein turnover through the ubiquitin-proteasome system (UPS) (Kimura et al., 2016). Recent studies indicated that the transcriptional repression of UBB is a cancer-subtype-specific event which occurs in approximately 30% of high-grade serous ovarian cancer (HGSOC) cases. Silencing of UBB reduces cellular ubiquitin levels which is resulted in the overexpression of UBC to compensate the lost function of UBB. These changes may have prognostic value (Dasgupta et al., 2009). Fibronectin 1 (FN1) was clustered with the Secreted Phosphoprotein 1 (SPP1), Transmembrane Protein 132A (TMEM132A), Stanniocalcin 2 (STC2), cysteine-rich angiogenic inducer 61 (CYR61), Wolfram syndrome type 1 (WFS1), Insulin-like Growth Factor Binding Protein 4 (IGFBP4), and Enamelin (ENAM) in the module 2 which is enriched with extracellular matrix structural constituent. The FN1 has numerous functional properties and is involved in cell adhesion, growth, migration, and differentiation procedures. Previous studies reported morphological alterations in tumors and tumor-derived cell lines that have been attributed to the decrease fibronectin expression, increased fibronectin degradation, and/or decreased expression of fibronectin-binding receptors, such as α5β1 integrin (Zhuo et al., 2016). The main functions of module 1 were correlated with RPs and RPL genes and structural constituent of ribosomes and nuclear-transcribed mRNA catabolic process.

Secreted phosphoprotein 1 (SPP1), also known as Osteopontin (OPN), as an upregulated gene in the present study, was found to be overexpressed in numerous tumors, including lung, colon, breast, and ovarian cancers (Wang et al., 2014; Zeng et al., 2018). Many recent studies demonstrated that the existence of SPP1 in cancerous tissue samples and sera of women with ovarian cancer promotes ovarian cancer progression via Integrin β1/FAK/AKT signaling pathway (Shevde et al., 2014). The SPP1 along with TMEM132A, CALU, C3, STC2, CYR61, WFS1, IGFBP4, FN1, and ENAM were correlated with the most upregulated gene-enriched signaling pathways including post-translational protein modifications, signaling receptor bindings, and ECM-receptor interactions in the module 1. To data, among all mentioned pathways, the ECM-receptor interactions pathway has been highlighted in cancer studies and also the interaction of this pathway with DEGs has been introduced as a diagnostic marker (Bao et al., 2019). The main cancer-related activity of this pathway is related to

adhesion, migration, differentiation, proliferation, and apoptosis. Therefore, Increasing the expression of SPP1 as an inflammatory, fibrotic, and carcinogenic gene has been well justified in the ECM-receptor interactions pathway.

Conclusion

The current study demonstrates that, the hub genes derived from the PPI network, including UBC, FN1, ACTB, SPP1, JUN, and RPL13A tend to be present in different cancer-related pathways and Go functions. After following the function of these genes in causing cancer we suggested that these genes may be have potential to become biomarker panel related to the EOC. Yet, more molecular biology experiments, computational method analysis on big data is needed to support this suggestion.

Acknowledgments

Conflict of interest
The authors declare no conflict of interest.

Reference

Anders S. and Huber W. (2010) Differential expression analysis for sequence count data. Genome Biol, 11:R106.

Anders S., Pyl P. T. and Huber W. (2014) HTSeq-a Python framework to work with high-throughput sequencing data. Bioinformatics, 31: 166-9.

Archana R., Simmons K. B., Robert C. and Bast J. R. (2013) The Emerging Role of HE4 in the Evaluation of Advanced Epithelial Ovarian and Endometrial Carcinomas. Oncology, 27: 548-556.

Arnedos M., Roulleaux Degage M., Perez-Garcia J. and Cortes J. (2019) Window of Opportunity trials for biomarker discovery in breast cancer. Curr. Opin. Oncol, 31: 486-492.

Ashburner M., Ball C. A., Blake J. A., Botstein D. and Butler H. (2000) Gene Ontology: tool for the unification of biology. Nat. Genet, 25: 25-9.

Bao Y., Wang L., Shi L., Yun F., Liu X., Chen Y., Chen C., Ren Y. and Jia Y. (2019) Transcriptome profiling revealed multiple genes and ECM-receptor interaction pathways that may be associated with breast cancer. Cell Mol Biol Lett, 24(38): 2-20.

Board P. G, Coggan M., Baker R. T., Vuust J. and Webb G. C. (1992) Localization of the human UBC polyubiquitin gene to chromosome band 12q24.3. Genomics, 12: 639-42.

Bolger A. M., Lohse M. and Usadel B. (2014) Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics, 30: 2114-20.

Castello L. M., Raineri D., Salmi L., Clemente N. and Vaschetto R. (2017) Osteopontin at the Crossroads of Inflammation and Tumor Progression. Mediators Inflamm, 22.

Cho A., Howell V. M. and Colvin E. K. (2015) The Extracellular Matrix in Epithelial Ovarian Cancer. A Piece of a Puzzle. Front. Oncol, 5: 245.

Dasgupta S., Wasson L. M., Rauniyar N., Prokai L. and Borejdo J. (2009) Novel gene C17orf37 in 17q12 amplicon promotes migration and invasion of prostate cancer cells. Oncogene, 28: 2860-2872.

De Cristofaro T., Di Palma T., Soriano A., Monticelli A. and Affinito O. (2016) Candidate genes and pathways downstream of PAX8 involved in ovarian high-grade serous carcinoma. J. Oncotarget, 7: 41929-41947.

Evans E. E., Henn A. D., Jonason A., Paris M. J. and Schiffhauer L. M. (2006) C35 (C17orf37) is a novel tumor biomarker abundantly expressed in breast cancer. Mol Cancer Ther, 11: 2919-30S.

Hao J. M, Chen J. Z., Sui H. M., Si-Ma X. Q. and Li GQ. (2010) A five-gene signature as a potential predictor of metastasis and survival in colorectal cancer. J. Pathol, 220: 475-89.

Huang Y. A., You Z. H., Chen X., Chan K. and Luo X. (2016) Sequence-based prediction of protein-protein interactions using weighted sparse representation model combined with global encoding. BMC Bioinformatics, 17: 184.

Kanehisa M. and Goto S. (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res, 28: 27-30.

Kim D., Langmead B. and Salzberg S. L. (2015) HISAT: a fast spliced aligner with low memory requirements. Nat Methods, 12: 357-60.

Kimura Y. and Tanaka K. (2016) Regulatory mechanisms involved in the control of ubiquitin homeostasis. J. Biochem, 147: 793-8.

Krzystyniak J., Ceppi L., Dizon D. S. and Birrer M. J. (2016) Epithelial ovarian cancer: the molecular genetics of epithelial. Ann. Oncol, 27: i4-i10.

Li M. X., Jin L. T., Wang T. J., Feng Y. J., Pan C. P. and Zhao D. M. (2018) Identification of potential core genes in triple negative breast cancer using bioinformatics analysis. Onco. Targets Ther, 11: 4105-4112.

Li Y., Xiao X., Ji X., Liu B. and Amos C. I. (2015) RNA-seq analysis of lung adenocarcinomas reveals different gene expression profiles between smoking and nonsmoking patients. Tumour. Biol, 36: 8993-9003.

Loghmani H., Noruzinia M., Abdul-Tehrani H., Taghizadeh M. and Karbassiane M. H. (2014) Association of estrogen receptors' promoter methylation and clinicopathological characteristics in Iranian patients with breast cancer. Molecular and Biochemical diagnosis (MBD), 1: 21-33.

Love M. I., Huber W. and Anders S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol, 15: 550.

Marcato P., Dean C. A., Araslanova R., Gillis M. and Joshi M. (2011) Aldehyde dehydrogenase activity of breast cancer stem cells is primarily due to isoform ALDH1A3 and its expression is predictive of metastasis. J. Stem. Cells, 29: 32-45.

Shevde L. A. and Samant R. S. (2014) Role of osteopontin in the pathophysiology of cancer. Matrix Biol, 37: 131–141.

Thor A. D., Young R. H. and Clement P. B. (1991) Pathology of the fallopian tube, broad ligament, peritoneum, and pelvic soft tissues. Hum. Pathol, 22: 856-867.

Torre L. A., Trabert B., DeSantis C. E., Miller K. D. and Samimi G. (2018) Ovarian cancer statistics. CA. Cancer J. Clin, 68: 284-296.

Wang S., Zhou L., Han L. and Yuan Y. (2014) Expression and purification of non-tagged recombinant mouse SPP1 in E. coli and its biological significance. Bioengineered, 5: 405-408.

Warburg O. (1956) On respiratory impairment in cancer cells. Science, 124: 269-70.

Wu B., Xie J., Du Z., Wu J. and Zhang P. (2014) PPI network analysis of mRNA expression profile of ezrin knockdown in esophageal squamous cell carcinoma. Biomed Res Int, 3: 651954.

Zeng B., Min Z., Wu H. and Zhengai X. (2018) SPP1 promotes ovarian cancer progression via Integrin β1/FAK/AKT signaling pathway. Onco Targets Ther, 11: 1333-1343.

Zhang X, Wang Y. (2019) Identification of hub genes and key pathways associated with the progression of gynecological cancer. Oncol Lett, 18: 6516-6524.

Zhuo C., Li X., Zhuang H., Tian S. and Cui H. (2016) Elevated THBS2, COL1A2, and SPP1 expression levels as predictors of gastric cancer prognosis. Cell Physiol Biochem, 40: 1316-1324.

Open Access Statement:

Supplementary Materials:
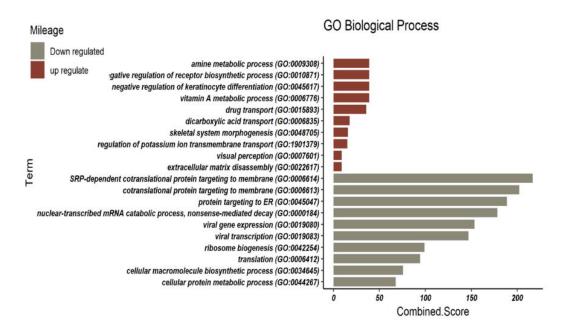
Supplementary Figures

Figure S1. Comparative gene ontology enrichment analysis of biological processes (BP) for up- and down-regulated genes of normal and cancerous ovarian samples. As shown, the most down-regulated genes are enriched in the (SRP)-dependent cotranslational protein-membrane targeting and the most up- regulated genes are enriched in the amino metabolic process.
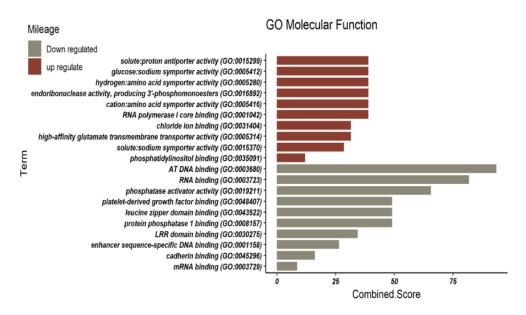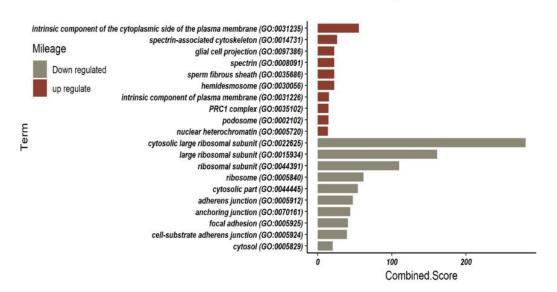

Figure S2. Comparative gene ontology enrichment analysis of molecular functions (MF) for up- and down-regulated genes of normal and cancerous ovarian samples. As shown, the most downregulated genes are enriched in the AT DNA binding and the most upregulated genes are enriched in the Solute: proton antiporter activity.

GO Cellular Component

Figure S3. Comparative gene ontology enrichment analysis of cellular components (CC) for up- and down- regulated genes of normal and cancerous ovarian samples. As shown, the most downregulated genes are enriched in cytosolic large ribosomal subunit and the most upregulated genes are enriched in the intrinsic component of the cytoplasmic side of the plasma membrane.

Supplementary Table

Table S1. The list of upregulated and downregulated genes (DEGs). This table is supplied as an excel file.