Identification of Hub Genes in High-Grade Serous Ovarian Cancer Using Weighted Gene Co-Expression Network Analysis

Corresponding Author: Guoyan Liu, e-mail: liuguoyan211@126.com
Source of support: This work was supported by a grant from the National Natural Science Foundation of China (grant number 81472761 to GL)

Background: High-grade serous ovarian cancer (HGSOC) is the most malignant gynecologic tumor. This study reveals biomarkers related to HGSOC incidence and progression using the bioinformatics method.

Material/Methods: Five gene expression profiles were downloaded from GEO. Differentially-expressed genes (DEGs) in HGSOC and normal ovarian tissue samples were screened using limma and the function of DEGs was annotated by KEGG and GO analysis using clusterProfiler. A co-expression network utilizing the WGCNA package was established to define several hub genes from the key module. Furthermore, survival analysis was performed, followed by expression validation with datasets from TCGA and GTEx. Finally, we used single-gene GSEA to detect the function of prognostic hub genes.

Results: Out of the 1874 DEGs detected from 114 HGSOC versus 49 normal tissue samples, 956 were upregulated and 919 were downregulated. The functional annotation indicated that upregulated DEGs were mostly enriched in cell cycle, whereas the downregulated DEGs were enriched in the MAPK or Ras signaling pathway. Two modules significantly associated with HGSOC were excavated through WGCNA. After survival analysis and expression validation of hub genes, we found that 2 upregulated genes (MAD2L1 and PKD2) and 3 downregulated genes (DOCK5, FANCD2, and TBRG1) were positively correlated with HGSOC prognosis. GSEA for single-hub genes revealed that MAD2L1 and PKD2 were associated with proliferation, while DOCK5, FANCD2, and TBRG1 were associated with immune response.

Conclusions: We found that FANCD2, PKD2, TBRG1, and DOCK5 had prognostic value and could be used as potential biomarkers for HGSOC treatment.

MeSH Keywords: Computational Biology • Ovarian Neoplasms • Survival Analysis

Full-text PDF: https://www.medscimonit.com/abstract/index/idArt/922107
Background

Ovarian cancer has a high mortality rate, which ranks first among gynecologic malignant tumors. Most deaths (70%) of patients presented with advanced-staged, high-grade serous ovarian cancer (HGSOC) due to the lack of specific symptoms at the early stage [1]. Therefore, it is of great significance to study the potential prognostic biomarkers related to the development of HGSOC.

In recent years, bioinformatics-assisted analyses of expression profile have been widely used to detect the biomarkers of human diseases [2]. Weighted gene co-expression network analysis (WGCNA) is a biological approach to determine highly synergistic gene sets and to identify the association between gene modules and phenotype of samples [3]. WGCNA has been comprehensively utilized in multiple cancer-associated studies to determine hub genes that could be associated with respective traits, such as pancreatic carcinoma [4], colon cancer [5], and ovarian cancer [6–10]. However, few previous studies have focused on HGSOC.

To identify potential biomarkers for specific diagnosis and therapy targets in HGSOC, WGCNA was performed to discover the hub genes that play an essential role in the development of HGSOC.

Material and Methods

Data collection

Five gene expression profiles were downloaded from the Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/). Datasets GSE18520 [11], GSE27651 [12], GSE54388 [13], GSE10971 [14] and GSE14001 [15] are listed in Supplementary Table 1, with a sample size of 114 for HGSOC and 49 normal tissue samples. All samples were processed using the Affymetrix human genome U133 plus 2.0 array. Genomic and clinical data were obtained from The Cancer Genome Atlas (TCGA) (https://cancergenome.nih.gov/) and GTEx (https://gtexportal.org/home/) [16] using the TCGAbiolinks package (Version 2.14.0; https://github.com/BioinformaticsFMRP/TCGAbiolinks [17]. The RNA expression profiles were sampled from 363 high-grade serous ovarian cancer and 108 normal tissues.

Figure 1. Flow diagram of this study.

Research design and data preprocessing

The research design is shown in a flowchart (Figure 1). The raw data from 5 datasets were chosen for integrated analysis using the affy package (Version 3.8; http://bioconductor.org/packages/release/bioc/html/affy.html) [18]. The batch effect of datasets was removed using the SVA package (Version 3.8; http://bioconductor.org/packages/release/bioc/html/sva.html) with its combat function (Supplementary Figure 1) [19].

Differential gene expression analysis

We detected the DEGs between HGSOC and normal ovarian tissue samples using the limma package (Version 3.30.0; http://www.bioconductor.org/packages/release/bioc/html/limma.html) [20]. A false discovery rate (FDR) <0.05 and |log2FC|>1 were set as the criteria value. The expression intensity and direction of DEGs were represented using the pheatmap package (Version 1.0.12, https://cran.r-project.org/web/packages/pheatmap).
Figure 2. The KEGG pathway enrichment analysis of differently-expressed genes. (A) KEGG pathways enrichment for upregulated genes. (B) KEGG pathways enrichment for downregulated genes.

Figure 3. The GO enrichment analysis of differently-expressed genes. (A) GO enrichment analysis of upregulated genes. (B) GO enrichment analysis of downregulated genes.

**Function enrichment analyses**

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses of DEGs were conducted using the clusterProfiler package (Version 6.8; http://www.bioconductor.org/packages/release/bioc/html/clusterProfiler.html) [21] to predict their underlying molecular functions. A p value of <0.05 was considered statistically significant.

**Weighted gene co-expression network analysis (WGCNA)**

We utilized the WGCNA package (Version 1.67; https://cran.r-project.org/web/packages/WGCNA/index.html) to construct a co-expression network for DEGs. To identify key modules, soft-thresholding power was set as $\beta=9$ (scale-free R$^2=0.84$), and cut height was set as 0.25 (Supplementary Figure 2). We then explored the biological function of the modules that had the highest correlation with traits through GO and KEGG pathway analyses, and hub genes in a module were selected with |MM|>0.85 and |GS|>0.3.

**Survival analysis and expression validation of hub genes**

Survival analysis was performed for hub genes using the survival package (Version 2.43.3; https://cran.r-project.org/package=survival) and survminer package (Version 0.4.3; https://cran.r-project.org/package=survminer). The Kaplan-Meier curves were plotted by the expression profiles from TCGA,

This work is licensed under Creative Common Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)
which were divided into 2 groups based on a certain gene's cutoff value as determined by survminer. The hub gene expression levels between HGSOC and normal tissue samples were also validated.

Figure 4. Co-expression modules construction and selection. (A) Samples clustering and trait heatmap of datasets from GEO according to the DEGs expression between HGSOC and normal tissue samples. (B) Dendrogram of all DEGs were clustered with dissimilarity according to topological overlap (1-TOM). (C) associations between modules and traits. In each cell, the upper number is the correlation coefficient of the module in the trait, and the lower number is the p value. Among them, the turquoise and blue modules were the most correlated with normal and cancer traits. (D) Distribution of average gene significance in the modules correlated with HGSOC. TOM – topological overlap matrix.

Gene set enrichment analysis (GSEA)

GSEA analysis of each hub gene with the TCGA-OV dataset was performed. The HGSOC samples (n=363) were divided into
2 groups according to the median expression value of each hub gene (high vs. low). A p value of <0.05 was considered as statistically significant. The “h.all.v6.2.entrez.gmt” were selected as reference gene sets, which were downloaded from the Molecular Signature Database (MSigDB, http://software.broadinstitute.org/gsea/msigdb).

Results

Differential expression analysis

We screened 1874 DEGs, including 919 downregulated and 956 upregulated genes, between HGSOC samples and normal tissue samples. The expression changes of DEGs were represented by a heatmap (Supplementary Figure 3), which showed that the samples were divided into 2 clusters.

Function enrichment for DEGs

The potential biological functions of the upregulated and downregulated differentially-expressed genes were annotated by clusterProfiler R package. The KEGG pathway analysis revealed that the upregulated genes were mainly involved in regulation of cell cycle, DNA replication, and biosynthesis of amino acids, while the downregulated genes were mainly linked to the Ras signaling pathway, complement and coagulation cascades, and the MAPK and JAK-STAT signaling pathways (Figure 2 and Supplementary Table 2). Furthermore, we performed GO enrichment analysis, the biological progress of which was in line with the KEGG enrichment results. Chromosome segregation and mitotic nuclear division were indicated for the upregulated genes, while morphogenesis of an epithelium and collagen-containing extracellular matrix were indicated for the downregulated genes (Figure 3 and Supplementary Table 3).
Co-expression modules construction

To construct co-expression modules and find the key modules related to HGSOC, the expression profiles of 1874 DEGs were assessed with the WGCNA package. Hierarchical clustering analysis is presented in Figure 4A. Then, the highly related genes were put into modules. The MED threshold was set as 0.25, and 6 modules were excavated (Figure 4B). The genes that

Figure 6. Protein–Protein Interaction (PPI) network of genes in 2 modules. (A) The genes in blue module. (B) The genes in turquoise module. The color presents the fold change (upregulated genes are red, downregulated genes are green).

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]
To quantify co-expression similarity of the 5 modules, we calculated the connectivity of eigengenes. Positively correlated eigengenes were grouped together, with 2 of 5 modules were classified into one cluster and 3 into another. The heatmap of the adjacencies is also presented (Figure 5C).

There were 76 hub genes from the turquoise module and 76 hub genes from the blue module selected, with a threshold module membership (MM) >0.85 and gene significance (GS) >0.3 (Supplementary Table 4).

The turquoise and blue modules were analyzed by STRING database, with a combined score >0.7 and were visualized by Cytoscape software (Figure 6).
To investigate the potential functions of the genes within the 2 modules (turquoise and blue), we performed GO and KEGG pathway analyses, and showed the most significant GO terms and KEGG pathways in Figure 7. This analysis revealed that genes in the blue module were mainly enriched in cell cycle and DNA replication, while genes in the turquoise module played their roles in different signal pathways.

Validation of hub genes

Analyzing the results of WGCNA, we found that the turquoise and blue modules had the highest association with HGSOC. Accordingly, we hypothesized that the genes in the turquoise module might act as tumor suppressors and genes in the blue module might act as tumor promoters. Survival analyses were
performed among the 152 hub genes selected from the 2 modules. We found that MAD2L1 and FANCD2 in the blue module and PKD2, TBRG1, and DOCK5 in the turquoise module were consistent with our speculation. Survival curves showed that higher expression of MAD2L1 and FANCD2 was significantly associated with poor prognosis of patients, as was the lower expression of PKD2, TBRG1, and DOCK5 (Figure 8). Finally, we used gene profiles downloaded from TCGA and GTEx to validate the expression of these genes, and the results were similar to the expression exhibited by GEO (Figure 9).

**Potential function of hub genes through GSEA**

To better understand the potential biological functions of MAD2L1, FANCD2, PKD2, TBRG1, and DOCK5 in HGSOC, we performed GSEA based on the TCGA-OV dataset. As shown in Figure 10, genes in higher-expression groups of MAD2L1 and FANCD2 were all involved in “E2F TARGETS” and “G2M CHECKPOINT” of the cell cycle, which indicated that these 2 upregulated genes are closely associated with tumor proliferation, whereas “TNFA SIGNALING VIA NFKB”, “interferon gamma RESPONSE” and “inflammatory response” were enriched in the PKD2, TBRG1, and DOCK5 high-expression groups, which indicated these downregulated genes are involved in immune response (Supplementary Table 5).

**Discussion**

With the purpose of identifying the molecular mechanism of HGSOC and to investigate potential biomarkers for better detection and therapy, we integrated the gene expression profiles of GSE54388, GSE27651, GSE10971, GSE18520, and GSE14001, which contained 114 samples of HGSOC tissue and 49 samples of normal tissue.

We identified 1874 DEGs that were correlated with HGSOC, and the cutoff criteria were $p<0.05$ and $|\log FC| \geq 1$. In KEGG analysis, the upregulated genes were predominantly involved in cell cycle and DNA replication, while the downregulated genes...
were highly involved in Ras signaling, complement and coagulation cascades, and MAPK signaling pathways. The GO analysis supported the previous enrichment analysis, which both help to understand the role of DEGs in HGSOC.

WGCNA analysis was used to select co-expression modules related to the development of HGSOC, and 2 modules (blue and turquoise) were found to have the highest correlation with HGSOC. We showed the Protein–Protein Interactions (PPI) network and also performed GO and KEGG analyses for genes in the 2 modules. The results indicated that genes in the blue module were enriched in cell cycle and DNA replication, while genes in the turquoise module were involved in different signaling pathways. After filtering with MM and GS value, we detected 152 hub genes from the 2 modules. Five genes – MAD2L1 and FANCD2 in the blue module and PKD2, TBRG1, and DOCK5 in the turquoise module – were excavated after survival analysis and expression validation with datasets downloaded from TCGA, and were found to have prognostic value for HGSOC. Among these 5 hub genes, MAD2L1 and FANCD2 are associated with ovarian cancer.

Figure 10. Gene set enrichment analysis (GSEA) of hub genes in the TCGA-OV dataset. Three gene sets enriched in the high-expressed group of single-hub genes. (A) MAD2L1, (B) FANCD2, (C) PKD2, (D) TBRG1, and (E) DOCK5.
As a component of the mitotic checkpoint, high levels of MAD2L1 are related to increased cellular proliferation, migration, and metastasis, which can lead to shorter survival in various cancers [22–26]. However, in ovarian cancer, the role of MAD2L1 did not agree with previous findings that patients with lower MAD2L1 levels were less sensitive to paclitaxel and had shorter progression-free survival (PFS) and overall survival (OS) [27,28]. This discrepancy might have been caused by our analysis, ignoring the mutations of p53 and BRCA1, which are known regulators of MAD2L1 and are commonly mutated in HGSOC [29,30].

High FANCD2 levels have been shown to be associated with poor prognosis in many types of cancer [31–35], as well as in ovarian cancer [36]. FANCD2 overexpression can stabilize the replication fork, and create BRCA1/2 mutant tumor resistance towards PARP1/2 inhibitor treatments [37]. The results indicated that FANCD2 expression can influence cancer sensitivity to PARP1/2 inhibitors and thus could be used as a potential target of therapy.

To further explore the biological functions of the 5 selected hub genes, we conducted single-gene GSEA. “E2F TARGETS” and “G2M CHECKPOINT” were enriched in the high-expression groups of MAD2L1 and FANCD2, indicating their contribution to HGSOC proliferation. In the high-expression groups of PKD2, TBRG1, and DOCK5, immune-related signals, such as “TNFA SIGNALING VIA NFKB”, “INTERFERON GAMMA RESPONSE” and “INFLAMMATORY RESPONSE” were enriched, indicating the activity of immune response.

Conclusions

We identified several DEGs and meaningful gene modules in HGSOC. Four valuable hub genes (FANCD2, PKD2, TBRG1, and DOCK5) were strongly dysregulated in HGSOC tissues. GSEA further suggested that FANCD2 is associated with tumor proliferation, while PKD2, TBRG1, and DOCK5 influence immune response. More work is needed to fully reveal their individual contributions towards the pathogenesis of HGSOC and to validate their value as prognostic biomarkers.

Limitations of this study include the lack of analysis for detailed clinical classification of HGSOC, such as grade, stage, lymph node metastasis, and prognosis. In future research, we will explore hub genes and their potential function based on this clinical information in detail.

Conflict of interest

None.

Supplementary Data

Supplementary Table 1. Characteristics of the included datasets.

| Dataset ID | GPL ID | High-grade serous ovarian carcinoma | Normal ovarian surface epithelium |
|------------|--------|------------------------------------|----------------------------------|
| GSE18520   | GPL570 | 53                                 | 10                               |
| GSE27651   | GPL570 | 22                                 | 6                                |
| GSE54388   | GPL570 | 16                                 | 6                                |
| GSE10971   | GPL570 | 13                                 | 24                               |
| GSE14001   | GPL570 | 10                                 | 3                                |

Wu M. et al.: Identification of hub genes in high-grade serous ovarian cancer… © Med Sci Monit, 2020; 26: e922107

This work is licensed under Creative Common Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)
Supplementary Figure 1. Samples clustering of 5 datasets after removing the batch effects.

Supplementary Figure 2. Soft-thresholding power determination in WGCNA. (A) Analysis of the scale-free fit index for different soft-thresholding powers. (B) Mean connectivity for various soft-thresholding powers. (C) Histogram of connectivity distribution when $\beta=9$. (D) Check scale-free topology when $\beta=9$. 

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]
Supplementary Figure 3. Heatmap of the top 200 DEGs based on the value of |logFC|. High or low expression is shown as a red or blue strip, respectively. The experimental group was labelled HGSOC, while the control group was named Nor.

Supplementary Table 2. The KEGG enrichment analysis of genes.

| ID      | Description                              | p. adjust | Count | Regulation |
|---------|------------------------------------------|-----------|-------|------------|
| hsa04110 | Cell cycle                               | 0.000633  | 35    | up         |
| hsa03030 | DNA replication                           | 0.012093  | 14    | up         |
| hsa01230 | Biosynthesis of amino acids               | 0.028478  | 12    | up         |
| hsa04014 | Ras signaling pathway                     | 0.007503  | 25    | down       |
| hsa04610 | Complement and coagulation cascades       | 0.007503  | 14    | down       |
| hsa04010 | MAPK signaling pathway                    | 0.007503  | 29    | down       |
| hsa04630 | JAK-STAT signaling pathway                | 0.007503  | 13    | down       |
| hsa04728 | Dopaminergic synapse                      | 0.007503  | 13    | down       |
| hsa05032 | Morphine addiction                        | 0.007503  | 13    | down       |
| hsa04261 | Adrenergic signaling in cardiomyocytes    | 0.013251  | 12    | down       |
| hsa05146 | Amoebiasis                                | 0.03213   | 14    | down       |
### Supplementary Table 3. The GO enrichment analysis of genes.

| Ontology | ID       | Description                              | p. adjust  | Count | Regulation |
|----------|----------|------------------------------------------|------------|-------|------------|
| BP       | GO: 0007059 | Chromosome segregation                    | 7.51E-30   | 76    | up         |
| BP       | GO: 0140014 | Mitotic nuclear division                   | 8.88E-28   | 69    | up         |
| BP       | GO: 0000280 | Nuclear division                           | 1.11E-26   | 83    | up         |
| BP       | GO: 0048285 | Organelle fission                          | 3.71E-25   | 85    | up         |
| BP       | GO: 0098813 | Nuclear chromosome segregation             | 5.13E-25   | 63    | up         |
| CC       | GO: 0098687 | Chromosomal region                         | 1.26E-26   | 74    | up         |
| CC       | GO: 0000775 | Chromosome, centromeric region             | 2.73E-24   | 53    | up         |
| CC       | GO: 0000793 | Condensed chromosome                       | 1.10E-20   | 52    | up         |
| CC       | GO: 0000779 | Condensed chromosome, centromeric region   | 2.30E-19   | 37    | up         |
| CC       | GO: 0005819 | Spindle                                   | 2.30E-19   | 62    | up         |
| MF       | GO: 0140097 | Catalytic activity, acting on DNA          | 2.32E-08   | 34    | up         |
| MF       | GO: 0008094 | DNA-dependent ATPase activity              | 1.57E-05   | 19    | up         |
| MF       | GO: 0016887 | ATPase activity                           | 0.000226   | 47    | up         |
| MF       | GO: 0043142 | Single-stranded DNA-dependent ATPase activity | 0.000325   | 7     | up         |
| MF       | GO: 0001077 | Proximal promoter DNA-binding transcription activator activity, RNA polymerase II-specific | 0.000325 | 35 | up |
| BP       | GO: 002009  | Morphogenesis of an epithelium             | 9.98E-07   | 53    | down       |
| BP       | GO: 0016049 | Cell growth                                | 2.29E-06   | 52    | down       |
| BP       | GO: 0001822 | Kidney development                         | 6.68E-06   | 35    | down       |
| BP       | GO: 0072001 | Renal system development                   | 6.68E-06   | 36    | down       |
| BP       | GO: 0003002 | Regionalization                            | 6.68E-06   | 40    | down       |
| CC       | GO: 0062023 | Collagen-containing extracellular matrix    | 1.22E-11   | 49    | down       |
| CC       | GO: 0031012 | Extracellular matrix                       | 1.29E-10   | 59    | down       |
| CC       | GO: 0042383 | Sarcolemma                                 | 3.36E-05   | 21    | down       |
| CC       | GO: 0005604 | Basement membrane                          | 3.64E-05   | 17    | down       |
| CC       | GO: 0031252 | Cell leading edge                          | 0.006422   | 34    | down       |
| MF       | GO: 005201  | Extracellular matrix structural constituent | 4.99E-05   | 24    | down       |
| MF       | GO: 0005339 | Glycosaminoglycan binding                  | 5.13E-05   | 29    | down       |
| MF       | GO: 1901681 | Sulfur compound binding                    | 6.67E-05   | 30    | down       |
| MF       | GO: 0008201 | Heparin binding                            | 0.000129   | 23    | down       |
| MF       | GO: 0005518 | Collagen binding                           | 0.000836   | 13    | down       |
### Supplementary Table 4. Hub genes in blue and turquoise module (|MM|>0.85 and |GS|>0.3).  

| Gene | Blue module | Turquoise module |
|------|-------------|------------------|
|      | MM          | GS               | MM          | GS               |
| CSE1L | 0.885182    | 0.654936         | LRRN4       | 0.863969         | -0.76695 |
| PCNA  | 0.853268    | 0.634531         | LINC01105   | 0.85413          | -0.83749 |
| MNB1C | 0.816407    | 0.649059         | LMX2        | 0.828442         | -0.6425  |
| TOP2A | 0.891679    | 0.822887         | CELF2       | 0.888927         | -0.77641 |
| SMC4  | 0.917093    | 0.806692         | LAMA4       | 0.850672         | -0.68463 |
| MTHFD2| 0.877458    | 0.806098         | SPOCK1      | 0.867584         | -0.80181 |
| PSRC1 | 0.871528    | 0.788497         | PAPSS2      | 0.879786         | -0.72486 |
| CKS1B | 0.934063    | 0.787083         | DAPK1       | 0.899065         | -0.7982  |
| MCM2  | 0.922548    | 0.841694         | PROC4       | 0.898493         | -0.7677  |
| PCLAF | 0.899258    | 0.817579         | PKD2        | 0.929768         | -0.75857 |
| CRABP2| 0.87722     | 0.877783         | GSDME       | 0.91672          | -0.80166 |
| CCNB2 | 0.897672    | 0.848503         | IGFBP6      | 0.892364         | -0.79082 |
| LSM4  | 0.871655    | 0.664114         | THBD        | 0.869457         | -0.80323 |
| CDC20 | 0.893728    | 0.758763         | KDR         | 0.884189         | -0.74272 |
| UBE2C | 0.876499    | 0.874623         | TRY         | 0.899738         | -0.80019 |
| CDK1  | 0.876549    | 0.8119           | GNG11       | 0.930805         | -0.75466 |
| EZH2  | 0.862178    | 0.795034         | ABCA8       | 0.944703         | -0.89654 |
| MAD2L1| 0.889283    | 0.76929          | GPRAS1      | 0.899009         | -0.87775 |
| PTG1  | 0.870574    | 0.824798         | GPT2        | 0.866683         | -0.7817  |
| BUB1B | 0.892962    | 0.819369         | RNASE4      | 0.903103         | -0.80729 |
| DLGAP5| 0.877017    | 0.82571          | CALB2       | 0.897513         | -0.85688 |
| ZWINT | 0.922712    | 0.79987          | BCHE        | 0.922752         | -0.81917 |
| TRIP13| 0.884367    | 0.77344          | NPY1R       | 0.882911         | -0.85536 |
| RAD51AP1| 0.858395  | 0.802612         | GHR         | 0.888245         | -0.71151 |
| NDC80 | 0.86773     | 0.721595         | ECM2        | 0.850415         | -0.66158 |
| CKS2  | 0.94052     | 0.829763         | ARHGAP6     | 0.869659         | -0.73929 |
| KIF11 | 0.885283    | 0.822206         | WNT2B       | 0.877108         | -0.80986 |
| NEK2  | 0.89035     | 0.839637         | PTG1        | 0.883089         | -0.79366 |
| KIF23 | 0.853223    | 0.7629           | LGALS2      | 0.85777          | -0.75854 |
| FEN1  | 0.909533    | 0.760884         | MAF         | 0.898868         | -0.73781 |
| TTK   | 0.894859    | 0.855031         | SYNE1       | 0.854181         | -0.8207  |
| MELK  | 0.900554    | 0.813158         | PLPP1       | 0.920489         | -0.7545  |
| STIL  | 0.867957    | 0.835405         | TCEAL2      | 0.877183         | -0.81491 |
| SAC3D1| 0.874805    | 0.712282         | TBC1D2B     | 0.862424         | -0.70643 |
Supplementary Table 4 continued. Hub genes in blue and turquoise module ($|MM|>0.85$ and $|GS|>0.3$).

| Gene      | Blue module | Turquoise module |
|-----------|-------------|------------------|
|          | MM          | GS               |          | MM          | GS               |
| HMGAI     | 0.869641    | 0.806441         | PDE8B    | 0.878248    | -0.91687         |
| GINS1     | 0.893705    | 0.7861           | ATP1OD   | 0.88824     | -0.74773         |
| CENPF     | 0.880011    | 0.84842          | TFPI     | 0.86169     | -0.75802         |
| AURKA     | 0.914224    | 0.801243         | CHN2     | 0.855714    | -0.81706         |
| EIF4G1    | 0.869979    | 0.737708         | BICC1    | 0.864204    | -0.82881         |
| NR2F6     | 0.887103    | 0.752212         | DIXDC1   | 0.858203    | -0.8161          |
| BUB1      | 0.913764    | 0.809066         | DIRR2    | 0.875331    | -0.83324         |
| PUF60     | 0.866461    | 0.739985         | OLFML1   | 0.886786    | -0.70603         |
| TPX2      | 0.862716    | 0.819783         | CSGLNACT1| 0.90082     | -0.83234         |
| RPL39L    | 0.857416    | 0.742065         | PDGFD    | 0.866505    | -0.68393         |
| EIF6      | 0.86057     | 0.671087         | RADX     | 0.868631    | -0.84742         |
| XPOT      | 0.857586    | 0.684634         | KLF2     | 0.894365    | -0.78132         |
| SCRB      | 0.867393    | 0.770112         | SMPD3    | 0.867973    | -0.78767         |
| CCNA2     | 0.868501    | 0.7858           | PPP1R3B  | 0.85511     | -0.72434         |
| CCNB1     | 0.904717    | 0.755995         | OGN      | 0.88218     | -0.80725         |
| PRK1      | 0.886914    | 0.802267         | ABI3BP   | 0.872052    | -0.78816         |
| MRPL15    | 0.853847    | 0.663766         | ITLN1    | 0.884383    | -0.81818         |
| NUSAP1    | 0.867073    | 0.795277         | MGARP    | 0.903375    | -0.8306          |
| SLC5A2    | 0.868392    | 0.803613         | ARHGAP18 | 0.901991    | -0.80114         |
| TACC3     | 0.875177    | 0.752928         | DDR2     | 0.852443    | -0.67497         |
| KIF4A     | 0.871167    | 0.799792         | ANTXR2   | 0.858928    | -0.72284         |
| CEPS5     | 0.850744    | 0.827425         | LIX1L    | 0.878142    | -0.68937         |
| DTL       | 0.865255    | 0.780689         | MCC      | 0.870811    | -0.71797         |
| KIF20A    | 0.888407    | 0.825111         | TBRG1    | 0.85824     | -0.79791         |
| CENPU     | 0.863996    | 0.732426         | PTPN21   | 0.863405    | -0.7201          |
| KIF15     | 0.882733    | 0.786989         | CNRIP1   | 0.889837    | -0.80871         |
| ECT2      | 0.916732    | 0.831958         | PPM1K    | 0.895422    | -0.86713         |
| CDCA8     | 0.85916     | 0.797452         | MEDAG    | 0.866729    | -0.84797         |
| MCM4      | 0.914084    | 0.813176         | LINCO1279| 0.857322    | -0.66728         |
| RACGAP1   | 0.918941    | 0.779139         | PLEKHH2  | 0.865982    | -0.76522         |
| DSAT1     | 0.854015    | 0.826163         | SLC30A4  | 0.904592    | -0.70542         |
| UBE2T     | 0.850996    | 0.725215         | TCEAL3   | 0.856911    | -0.7193          |
| SLC25A33  | 0.868987    | 0.6974           | CDON     | 0.869831    | -0.69303         |
| CDCA3     | 0.857304    | 0.797015         | TCEAL7   | 0.870405    | -0.84944         |
**Supplementary Table 4 continued.** Hub genes in blue and turquoise module (|MM|>0.85 and |GS|>0.3).

| Gene       | MM   | GS         | Gene       | MM   | GS         |
|------------|------|------------|------------|------|------------|
| NUF2       | 0.91985 | 0.868254   | ERN1       | 0.880253 | -0.77167   |
| RCC2       | 0.909521 | 0.783731   | MUM1L1     | 0.899449 | -0.85402   |
| FAM83D     | 0.901393 | 0.849648   | RNASEL     | 0.872061 | -0.72795   |
| POClA      | 0.856044 | 0.816551   | DOCK5      | 0.893588 | -0.83622   |
| DEPDC1B    | 0.854264 | 0.764695   | RBMS3      | 0.850965 | -0.82463   |
| CENPL      | 0.857368 | 0.768059   | HAND2-AS1  | 0.858159 | -0.89488   |
| KIF14      | 0.902918 | 0.830047   | DTWD1      | 0.879321 | -0.75556   |
| FANCD2     | 0.890071 | 0.736399   | IFFO1      | 0.857298 | -0.73892   |

MM – module membership; GS – gene significance.

**Supplementary Table 5.** The Gene Set Enrichment Analysis (GSEA) of hub genes.

| Description      | setSize | enrichmentScore | NES    | p.adjust | core_enrichment                                                                 |
|------------------|---------|-----------------|--------|----------|---------------------------------------------------------------------------------|
| ALLOGRAFT REJECTION | 120    | 0.438738        | 1.761702 | 0.007716 | CDKN2A/NME1/G2M2/MMP9/CXCL9/CCL5/CXCL13/L15/CCL11/EIF5A/TAPI/CXCL13/GZMA/SRGN/I2RG/CCL2/UB2K/CCL7/HLA-DOB/CTSS/CCL4/B2M/CD3D/PRF1/CD2/TLR/NK1/S1T1/L2RA/CD7/HLA-G/CD8A/CD38/ST5A4/CDB6/FGR2B/FNG/IL12A/CXCR3/LY86/CDBB/RIP2/UB2E1/PSD2/HLA-DQA1/MRPL3/CDB8/WARS/CD79A/CC1/LCK/HDAC3/IGSF6/  |
| E2F TARGETS     | 105    | 0.805846        | 3.155588 | 0.007716 | MAD2L1/CDKN2A/BIRC5/C5S2/C5S1B/C5C1E/CCNE1/TK1/UBE2S/PTTG1/UBE2T/MYBL2/NME1/  |
| MAD2L1          |         |                 |        |          | CENB2/AURKB/PK1/DEPD1/KPNA2/CDC20/RMM2/CENPM/CDKN3/CDK1/PLK4/AURKA/PCNA/SNPB/  |
| G2M CHECKPOINT  | 104    | 0.759634        | 2.975861 | 0.007716 | M1L1/CCNA2/U2B2C/BIRC5/C5S2/C5S1B/U2B5/PTTG1/UBE2T/MYBL2/NME1/CCNB2/AURKB/  |
|                 |         |                 |        |          | PK1/KPNA2/CDC20/CDNPA/CDKN3/TTK/CDK1/PLK4/NEK2/AURKA/GINS2/KIF2C/JPT1/ORC6/  |
|                 |         |                 |        |          | H2AFX/CDC45/TOP2A/TOPAP/TACC3/CDC6/SNRPD1/TPX2/IF4A/GINS2/HPMRE/    |

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]
Supplementary Table 5 continued. The Gene Set Enrichment Analysis (GSEA) of hub genes.

| Description                  | setSize | enrichmentScore | NES     | p.adjust | core_enrichment                                                                 |
|------------------------------|---------|-----------------|---------|----------|--------------------------------------------------------------------------------|
| G2M CHECKPOINT               | 104     | 0.901015        | 3.369604| 0.006028 | MYBL2/KIF15/TPX2/TOP2A/KIF2C/UBE2C/BIRC5/ESPL1/MA2D1/HRMTR/FK/FK4A/PLK1/TROAP/TTK/BUBI/CD20/PLQLQ/NUSAP1/RACGAP1/CCNB2/AURKB/CENPA/MK167/CNA2/KNL1/CDK1/TACC3/TRAIP/PLKX/E2F2/CENPF/AURKA/KIF3/FK11/BRCA2/NEK2/CDCA5/NDZC80/EXO1/CD25A/E2F1/CDS1/CD6/UBE2S/PRC1/KNP2A/RADS4L/CKS2/CENPE/SMC2/STL/CCNF/LMIN1/CDKN3/PTTG1/STMN1/EZH2/ORC6/GINS2/CD7/FFXOS/MCM2/ODC1/NDZ2/H2AX/KIF22/MC6/INCENP/SMC4/CHEK1/DDX39A/KIF20B/BARD1/DYMK/CHA1/SLV9H1 |
| E2F TARGETS                  | 105     | 0.889041        | 3.34287 | 0.006028 | MYBL2/CDKN2A/DEPDC1/MEK/CCNE1/ASF1B/TP2A/KIF2C/TRIP13/BIRC5/ESPL1/BUUB/MA2D1/AHRMR/KIF4A/PLK1/CD20/CIT/CDCA8/SPA/S6/P2/C4/RACGAP1/CCNB2/AURKB/RMR2/TK1/MK167/KIF18B/PLGAPS/CDK1/TACC2/PLK3/CDK1/GINS4/AURKA/BRCA2/E2F8/RFC3/DIAH3/SPC25/CCN12/CKS2/TIMELESS/UBE25/RADS1AP1/KNP2A/CKS2/CENPE/GINS1/LMN1/CDKN3/PTTG1/STMN1/E2BT/CENPM/EZH2/ORC6/ATAD2/MCM2/MCM4/NCAPD2/H3LS/RNASEH2A/PCNA/H2AX/KIF22/MC6/SMC4/CHK1/DDX39A/BARD1/DSCC1/GINS3/TCF19/UV3H1/RFC2/CSE1L/UNG/MH2/SNRPB/PRC1/PIF1/CENPE/LMN1/FBXOS/KIF22/INCENP/SMC4/KIF20B/CENPF/SA556 |
| FANCD2                       |         |                 |         |          |                                                                                  |
| MITOTIC SPINDLE              | 80      | 0.798921        | 2.888574| 0.006028 | KIF15/TPX2/TOP2A/KIF2C/BIRC5/ESPL1/KIF4A/PLK1/TTK/BUUB1/ANL1/NUSAP1/RACGAP1/CCNB2/E2T2/PLGAPS/CDK1/CENPF/AURKA/KIF13/FK11/BRCA2/NEK2/NDZC830/PRC1/PIT1/CENPE/LMN1/FBXOS/KIF22/INCENP/SMC4/KIF20B/CENPF/SA556 |
| KRAS SIGNALING UP            | 118     | 0.656511        | 2.270659| 0.004293 | MMP11/PRRX1/PLAU/TMEM158/ETV1/CFH/GFP2T/LI/PLAT/SACCO1/ADGR2/TMEM17A/MM/ATM/LPTM5/ITGB2/PCSK1N/TMEM17B/RGS16/EBR1/ENG/NRP1/TFAP3/L2RG/ADP/MALL/EPHB2/K2F1/PLAUR/WNT7A/MABF/TFPI/AKAP12/TRA2B/KL/F4A/CXCL10/SP1/STP/CLC3/AIR1/PPON1/EVT5/ADAMDEC1/LCP1/PCR1/GTFL4/GTP5/GO52/TRA1/UV5/DUSP6/CTSS/ADAM8/SOX9/PPP1R5A/MM/IRF8 |
| PKD2                         |         |                 |         |          |                                                                                  |
### Supplementary Table 5 continued. The Gene Set Enrichment Analysis (GSEA) of hub genes.

| PKD2                  | Description                        | setSize | enrichmentScore | NES     | p.adjust | core_enrichment                                                                 |
|-----------------------|------------------------------------|---------|-----------------|---------|----------|--------------------------------------------------------------------------------|
|                       | TNFA SIGNALING VIA NFKB             | 109     | 0.696773        | 2.395892| 0.004293| SERPINE1/PLAU/FO5B/KLF2/ICAM1/GFPT2/LIF/EGR1/SLC2A3/FO5/ZFP36/DUSP1/TFNAIP6/NRA3/LIF/EGR2/TFNAIP3/LDLR/TFNAIP2/GADD45B/PLAU/Plek/NFAT5/CDKN1A/CCL2/KDM6B/KLF4/CXCL11/CXCL10/BMP2/IL6ST/DUSP4/FO5S2/CCL4/CXCL11/IER3/OSG1/TRAFA1/JUNB/F3/CD44/PP1R15A/SERPINB2/RHOB/NRA42/KLF9/SGK1/PTGER4/IFIT2/B4GALT5/MAFF/IER5/CCL6/ETSA/PER1/BCL6/TAPI1/STFR59/SMA3/ID2/PLPP3/IL1B/PTX3/SLC2A6/RNF19B/BIRC3/IFIH1 |
|                       | INTERFERON GAMMA RESPONSE           | 107     | 0.628308        | 2.160028| 0.004293| C1S/CFH/CXCL9/ICAM1/C1R/XAF1/TFNAIP6/OAS2/IL2RB/LATS2/CCL5/CSF2RB/LCP2/IF4F4L/OAS3/HLA-DQA1/RSAD2/TFNAIP3/HLA-B/TFNAIP2/MX1/HLELZ/SLAMF7/CDKN1A/CCL2/STAT1/CXCL10/FA5/EPST11/IFIT3/CD38/PIM1/TAPBP/CXCL11/SEP/C74/WARS/ST8SIA4/IRF8/ST3GALS5/IF44/LY6E/CD86/MLG53BP/IFIT2/FCGR1A/OAS1/EIF2AK2/MDMB7/F3/CTB/TAPI1/IFIT3/CMPK2/B2M/HLA-DRB1/PMI/IFIH1/STN/IFIT2/HLA-G/JAK2/TRIM14 |

| TBRG1                 | ALLOGRAFT REJECTION                | 120     | 0.489708        | 1.924057| 0.008729| IL18/THY1/LIF/CD74/HLA-DOA/HLA-DMA/HLA-DQA1/C2/HLA-DRA/LTB/IL2R/FGAS/ELF4/PRKCB/C47/PRKCG/B2M/CD3E/LY75/ICAM1/INHBB/TAI1/TAPBP/IL2RB/HDA9/C2/D2/IL16/CCL5/GZMA/FYB1/C906/CD4/JAK2/CXCL9/IL15/STAB1/CD77/CCL4/ITGa/HLA-DOB/IGSF6/IKKB/HLA-G/ITGB2/LYN/TFN/IL2A/SP1/PTPRC/CRTAM/CDBA/PKR1/C2CL7/WAS/LC2P/CTS3/CD3D/TFSLG/CXCR3 |

|                      | TNFA SIGNALING VIA NFKB             | 109     | 0.549517        | 2.128181| 0.008729| BIRC3/IL18/CCND1/FOS/FO5B/LIF/CCL20/GADD45B/EGFR/CBE/EDN1/JUNB/SGK1/CCNL1/NRA41/NFAT5/ZFP36/F3/IRF1/KLF2/TFNAIP2/IFIT2/CCL1F/SMAD3/ETSA/DUSP1/ICAM1/TAPI1/LAMB3/MAFF/SERPINE2/PLAU/TRBI1/EGR3/BTG2/CCL5/TFNAIP/TAPI3/IR3/TIPARP/EGR2/BMP2/TNF |

|                      | INTERFERON GAMMA RESPONSE           | 107     | 0.524759        | 2.03116| 0.008729| CFB/XAF1/C74/HLA-DOA/HLA-DQA1/IFIT3/M/HLA-DRB1/MX2/RTP4/PSMB8/IFIT2/PSMB9/IRF1/IOD1/IFIT3/IFIT1/LY6E/FGAS/TFNAIP2/IFIT2/EPST11/B2M/ZBP1/TXNIP/ICAM1/TAPI1/TAPBP/IL2RB/PML/TFNSF10/ITGB7/HLA-B/CCL5/CASP8/GZMA/SLC2A5A8/JAK2/CIR/CXCL9/IL15/NI/SECTM1/MX1/HLA-G/TFNAIP3/UB2EL6/C15/PARIP12 |
References:

1. Siegel RL, Miller KD, Jemal A: Cancer statistics, 2019. Cancer J Clin, 2019; 69: 7–34
2. Servant N, Roméjon J, Gestraud P et al: Bioinformatics for precision medicine in oncology: Principles and application to the SHIVA clinical trial. Front Genet, 2014; 5: 152
3. Langfelder P, Horvath S: WGCNA: An R package for weighted correlation network analysis. BMC Bioinformatics, 2008; 9: 559
4. Zhou XG, Huang XL, Liang SY et al: Identifying miRNA and gene modules associated with pathological stage by weighted gene co-expression network analysis. OncoTargets Ther, 2018; 11: 2815–30
5. Chen Y, Bi F, An Y et al: Identification of pathological grade and prognosis of pancreatic carcinoma identified by co-expression analysis as a hub candidate biological target in human epithelial ovarian cancer. J Surg Pathol, 2011; 35: 904–12
6. King ER, Tung CS, Tsang YT et al: The anterior gradient homolog 3 (AGR3) gene is associated with differentiation and survival in ovarian cancer. Am J Surg Pathol, 2011; 35: 904–12
7. Yeung TL, Leung CS, Wong KK et al: ELF3 is a negative regulator of epithelial-mesenchymal transition in ovarian cancer cells. Oncotarget, 2017; 8: 16951–63
8. Tone AA, Virtanen C, Shaw PA et al: Decreased progesterone receptor isoform expression in luteal phase fallopian tube epithelium and high-grade serous carcinoma. Endocr Relat Cancer, 2011; 18: 211–34
9. Tung CS, Mok SC, Tsang YT et al: The anterior gradient homolog 3 (AGR3) gene is associated with differentiation and survival in ovarian cancer. Am J Surg Pathol, 2011; 35: 904–12
10. King ER, Tung CS, Tsang YT et al: The anterior gradient homolog 3 (AGR3) gene is associated with differentiation and survival in ovarian cancer. Am J Surg Pathol, 2011; 35: 904–12
11. Mok SC, Bonome T, Vathipadiekal V et al: A gene signature predictive for outcome in advanced ovarian cancer identifies a survival factor: Microfibril-associated glycoprotein 2. Cancer Cell, 2009; 16: 521–32
12. Gladhaug IP, Westgaard A, Schjolberg AR et al: Spindle proteins in resected pancreatic head adenocarcinomas: BubR1 is an independent prognostic factor in pancreaticobiliary-type tumours. Histopathology, 2010; 56: 345–55
13. Yung TL, Leung CS, Wong KK et al: ELF3 is a negative regulator of epithelial-mesenchymal transition in ovarian cancer cells. Oncotarget, 2017; 8: 16951–63
14. Tone AA, Virtanen C, Shaw PA et al: Decreased progesterone receptor isoform expression in luteal phase fallopian tube epithelium and high-grade serous carcinoma. Endocr Relat Cancer, 2011; 18: 211–34
15. Tung CS, Mok SC, Tsang YT et al: PK2 expression in low malignant potential ovarian tumors and low-grade ovarian serous carcinomas. Mod Pathol, 2009; 22: 1243–50
16. Carithers L, Ardaile K, Barcus M et al: A novel approach to high-quality post-mortem tissue procurement: The GTeX project. Biopreservation Biobank, 2015; 13: 311–19
17. Mounir M, Lucchetta M, Silva TC et al: New functionalities in the TCGAbilinks package for the study and integration of cancer data from GDC and GTEX. PLoS Comput Biol, 2019; 15: e1006701
18. Gautier L, Cope L, Bolstad BM et al: Affy—analysis of Affymetrix GeneChips data at the probe level. Bioinformatics, 2004; 20: 307–15
19. Leek JT, Johnson WE, Parker HS et al: The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics, 2012; 28: 882–83
20. Ritchie ME, Phipson B, Wu D et al: Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res, 2015; 43: e47
21. Wu M. et al.: Identification of hub genes in high-grade serous ovarian cancer. © Med Sci Monit, 2020; 26: e922107
22. Gladhaug IP, Westgaard A, Schjolberg AR et al: Spindle proteins in resected pancreatic head adenocarcinomas: BubR1 is an independent prognostic factor in pancreaticobiliary-type tumours. Histopathology, 2010; 56: 345–55
23. Genga KR, Filho FD, Ferreira FV et al: Proteins of the mitotic checkpoint and spindle are related to chromosomal instability and unfavourable prognosis in patients with myelodysplastic syndrome. J Clin Pathol, 2015; 68: 381–87
24. Choi JW, Kim Y, Lee JH et al: High expression of spindle assembly checkpoint proteins CDC20 and MAD2 is associated with poor prognosis in urothelial bladder cancer. Virchows Arch, 2013; 463: 681–87
25. Nascimento AV, Singh A, Bousbaa H et al: Mad2 checkpoint gene silencing using epidermal growth factor receptor-targeted chitosan nanoparticles in non-small cell lung cancer model. Mol Pharm, 2014; 11: 3515–27
26. Li L, Xu DB, Zhao XL et al: Combination analysis of Bub1 and Mad2 expression in endometrial cancer: Act as a prognostic factor in endometrial cancer. Arch Gynecol Obstet, 2013; 288: 155–65
27. Furlong F, Fitzpatrick P, O’Toole S et al: Low MAD2 expression levels associate with reduced progression-free survival in patients with high-grade serous epithelial ovarian cancer. J Pathol, 2012; 226: 746–55
28. Nakano Y, Sumi T, Teramae M et al: Expression of the mitotic-arrest deficiency 2 is associated with chemotherapy resistance in ovarian serous adenocarcinoma. Oncol Rep, 2012; 28: 1200–4
29. Wang RP, Yu H, Deng CK: A requirement for breast-cancer-associated gene 1 (BRCA1) in the spindle checkpoint. Proc Natl Acad Sci U S A, 2004; 101: 17108–13
30. Torlakovic EE, Riddell R, Banerjee D et al: Canadian Association of Pathologists-Association canadienne des pathologistes National Standards Committee/Immunohistochemistry: Best practice recommendations for standardization of immunohistochemistry tests. Am J Clin Pathol, 2010; 133: 354–65
31. Balascu Q, Balascu L, Tudoran O et al: Gene expression profiling reveals activation of the FA/BRCA pathway in advanced squamous cervical cancer with intrinsic resistance and therapy failure. BMC Cancer, 2014; 14: 246
32. van der Groep P, Hoezel M, Buerger H et al: Loss of expression of FANCD2 protein in sporadic and hereditary breast cancer. Breast Cancer Res Treat, 2008; 107: 41–47
33. Ozawa H, Iwatsuki M, Mimori K et al: FANCD2 mRNA overexpression is a bona fide indicator of lymph node metastasis in human colorectal cancer. Ann Surg Oncol, 2010; 17: 2341–48
34. Han SS, Tompkins VS, Son DJ et al: CDKN1A and FANCD2 are potential oncotargets in Burkitt lymphoma and multiple myeloma. Exp Hematol Oncol, 2015; 4: 9
35. Komatsu H, Masuda T, Iguchi T et al: Clinical significance of FANCD2 gene expression and its association with tumor progression in hepatocellular carcinoma. Anticancer Res, 2017; 37: 1083–90
36. Moes-Sosnowska J, Rzepecka IK, Chodzynska J et al: Clinical importance of FANCD2, BRIP1, BRCA1, BRCA2 and FANC expression in ovarian carcinomas. Cancer Biol Ther, 2019; 20: 843–54
37. Kais Z, Rondinelli B, Holmes A et al: FANCD2 maintains fork stability in BRCA1/BRCA2-deficient tumors and promotes alternative end-joining DNA repair. Cell Rep, 2016; 15: 2488–99