Combinatorial Bioinformatics Analysis Reveals Novel Biomarkers for Improved Ovarian Cancer Prognosis

Xiaoyi Zhang  
Shandong Provincial Hospital

Shihong Zhu  
Shandong Provincial Hospital

Min Peng  
Shandong Provincial Hospital

Hongbo Ma  (Lotuswind123@163.com)  
Shandong Provincial Hospital  https://orcid.org/0000-0003-0611-7026

Research

Keywords: Ovarian cancer, Microarray, Differentially expressed genes, Bioinformatics, Cell cycle, p53 signaling pathway

DOI: https://doi.org/10.21203/rs.3.rs-553852/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

**Background**: Given the known lethality of highly frequent ovarian cancer (OC) among females, it is imperative to investigate potential biomarkers of prognostic and therapeutic significance. The objective of this study was to identify significant differentially expressed genes (DEGs) with poor prognosis and to explore their underlying mechanisms.

**Methods**: We acquired three microarray datasets (GSE14407, GSE36668 and GSE18520), available from the public database GEO. We compared a total of 72 cancerous and 26 normal samples originating from ovarian tissues. GEO2R and Venn diagram tools were used to obtain DEGs, followed by the gene ontology (GO) and Kyoto Encyclopedia of Gene and Genome (KEGG) analysis via Database for Annotation, Visualization and Integrated Discovery (DAVID). Subsequently, protein-protein interaction (PPI) network was constructed and visualized in Cytoscape.

**Results**: Among three analyzed datasets, a total of 232 DEGs were common. The upregulated 108 genes were significantly enriched in the cell adhesion, cellular response to interleukin-1, positive regulation of transcription from DNA/RNA, and transcription, extracellular matrix/region, anchored membrane component, cell junction and golgi membrane, sequence-specific DNA binding, transcription factor activity, RNA polymerase II regulatory region, and DNA binding. The PPI network analysis via MCODE plug-in revealed a total of 14 upregulated genes. Kaplan-Meier plotter analysis revealed that 9 genes were associated with significantly worse survival among OC patients while 4 genes exhibited no significant effect. Gene Expression Profiling Interactive Analysis (GEPIA) showed that 13 DEGs had significantly higher expression in the ovarian cancerous tissues compared to the normal ones. Repeated KEGG analysis showed that 11 genes (CDC6, CCNE1, BUB1B, CCNB2, BUB1, SFN, TTK, CDC20, PTTG1, CDK1 and CDKN2A)) were mainly associated with cell cycle while 2 genes (SFN and RRM2) were related to p53 signaling pathway.

**Conclusion**: Our findings identify potential upregulated DEGs of prognostic value among OC patients. This will facilitate to understand the underlying OC mechanisms and to implement targeted therapeutic measures.

Background

Among women, ovarian cancer (OC) is the fifth leading cause of cancerous deaths all over the world [1]. Though few prognostic biomarkers have been identified to date, given the rapid dissemination, distant metastasis and challenges in early diagnosis, the overall OC survival rate remains low [2, 3]. Thus, exploration of novel prognostic biomarkers is imperative for a better understanding of the OC mechanisms and subsequently, to facilitate effective treatment. Over a decade, gene chip is being used as a reliable technique [4] for rapid detection of the differentially expressed genes (DEGs), generating a lot of slice data that is stored in the public databases. Therefore, these data can be used to further explore several valuable research questions. In recent years, several bioinformatical studies regarding OC have been published [5–7], which implies that the combinatorial bioinformatical techniques could potentially be helpful to study and to better explore the underlying disease mechanisms.

In the current study, we acquired three datasets (GSE14407, GSE36668 and GSE18520) from Gene Expression Omnibus (GEO) database and analyzed the common DEGs among these datasets using online Venn diagram
Afterwards, the Database for Annotation, Visualization and Integrated Discovery (DAVID) was employed to analyze the common DEGs including their role in biological process (BP), cellular component (CC), molecular function (MF) and Kyoto Encyclopedia of Gene and Genome (KEGG) pathway. Furthermore, to rule out the core genes, we performed integrated analysis using protein-protein interaction (PPI) network combined with Cytotype Molecular Complex Detection (MCODE). The significant prognostic \((p < 0.05)\) details of the genes were obtained by processing candidate DEGs in online database, Kaplan Meier Plotter. Finally, the validation of DEGs expression between normal and cancerous ovarian tissues was performed \((p < 0.05)\) using Gene Expression Profiling Interactive Analysis (GEPIA). Taken together, the initial analysis revealed only 26 candidate genes. Repeated KEGG analysis of these genes showed that 11 DEGs (CDC6, CCNE1, BUB1B, CCNB2, BUB1, SFN, TTK, CDC20, PTTG1, CDK1, CDKN2A) in the cell cycle while 02 DEGs (SFN and RRM2) were generated and significantly enriched in p53 signaling pathway. In conclusion, our combinatorial bioinformatics analysis reveals potential OC biomarker genes which could be useful for targeted and effective prognosis among OC patients.

## Results

### Identification of DEGs in OC

The data from three expression profiles (GSE14407, GSE36668 and GSE18520) consisted of 12 OC vs 12 normal tissues, 8 OC vs 4 normal tissues and 53 OC vs 10 normal tissue samples. All microarray datasets were based on platform GPL570 (Affymetrix Human Genome U133 Plus 2.0 Array) (Table 1). The Venn diagram-based results revealed that a total of 232 DFEGs were common among the three datasets. Among these genes, 108 were upregulated \((FC > 0)\) while 124 genes were downregulated \((FC < 0)\) in the OC tissues compared to the normal ones (Fig. 1 & Table 2).

| GEO Accession | Year | Platform | Annotation Platform | No. of Samples | Sample Type          |
|---------------|------|----------|---------------------|----------------|----------------------|
| GSE14407      | 2009 | GPL570   | Affymetrix Human Genome U133 Plus 2.0 Array | 12 12          | Ovarian Surface Epithelia |
| GSE36668      | 2012 | GPL570   | Affymetrix Human Genome U133 Plus 2.0 Array | 08 04          | Ovarian Surface Epithelia |
| GSE18520      | 2009 | GPL570   | Affymetrix Human Genome U133 Plus 2.0 Array | 53 10          | Ovarian Surface Epithelia |
Table 2

A total of 232 differentially expressed genes (DEGs) were detected using 3 datasets, including 108 upregulated and 124 downregulated genes in ovarian cancer (OV) tissues compared to the normal (OV) tissues.

| DEGs | Gene Symbol |
|------|-------------|
| Upregulated | MUM1L1 LAMA2 LOC100507387///FAM153A///FAM153B NAP1L2 ITLN1 GADL1 NEFH TEMEM255A PPM1K BAMBI LOC101930363///LOC101928349///LOC100507387///FAM153A///FAM153B DPYD HLF PRSS35 GABRG1 THBD PRRX1 ABCA8 ADAMTS9-AS1 SPOCK3 TCF21 PDGFD KLF2 SNAIP NEGR1 HBD NT5E CLEC4M SNCA PLEKH2 GAS1RR MTUS1 GPM6A CPED1 MGARP SEMA6A LSAMP EFEMP1 B3GALT2 HBA2///HBA1 CHGB DIRAS3 PRKAR2B KCNT2 TEMEM150C ECM2 RORA OGN GIPC2 SNX292P2 ARX ARHGAP18 TCEAL2 NAP1L3 TCEAL7 CXorf57 CSGALNACT1 MCOLN3 CNTN1 AKAP12 HBG2///HBG1 COL14A1 CALCRL STAR ALDH1A1 SMPD3 TBX3 FGFR1OP2 SYNE1 BEX1 NR2F1-AS1 MCC CBLN4 HBB ITM2A GNG11 TFP1 GPRASP1 PEG3 PCDH9 HAND2-AS1 NKAIN2 RBMS3 PRDM5 HHIP CCND2 MAF PDGFRA PDE8B SIGLEC11 TLE4 DCN PEX5L ABCA5 HAS1 BNC2 ARMCX4 NRK ME1 GATM RNF128 LHX9 AOX1 AKT3 SFRP1 SYT4 RNASE4 GATA4 |
| Downregulated | MMP7 KLHL14 IGF2BP3 XK SUSD2 KLK8 CCNE1 ADGRF1 CDK1 LAMP3 SORT1 PCDH7 ASS1 AURKB SOX9 METTL7B KIF4A RASSF10 MCM10 CBS EPHX4 MELK AIF1L E2F7 GPR160 NUF2 SCGB1D2 PTTG1 NRTN Cdcas5 SLC2A1 CHMP4C EPCAM RGS1 UCP2 ECT2 ADGRG1 COL9A1 CRABP2 TSPAN12 KLK7 CLDN3 GALNT6 CCNB2 LRPP8 CEP55 S100A2 WFDC2 TFAP2A TIMELESS NR2F6 MECOM ESRP1 MKI67 LOC101928554 KIAA0101 XK4R ST6GALNA1 C1orf106 MPZL2 EHF KLK6 FOXQ1 PROM2 DEFB1 SMIM22 MUC1 KIF11 PART1 ELF3 PIGR F0LR1 MAL SYNE4 STON2 KCCAT333 LOC101929219///LOC100505650///C1orf186 KIAA1217 NLRP7 SOX17 KLK12 CDH6 SFN CDC20 CXXC5 DCDC2 INHBB SCGB2A1 BUB1 LYPD1 CD24 UBE2C SDC1 LEMD1 RRM2 CX3CL1 TOP2A WDR72 PAX8 BUB1B LYNX1 CDC8c RNF157-AS1 DTL HMMR FAM107A KIF20A STC2 BCL11A LCN2 LOC101929219///C1orf186 BCAT1 CDKN2A HMGA1 TTK SLCA4A11 GPR39 NCAPG ATP6V1B1 SLC52A2 CP CENPF KRT23 TNN1 |

**DEGs functional enrichment and KEGG analysis**

Next, all 232 DEGs were subjected to GO analysis by DAVID. The results indicated that for BP, the upregulated genes were mainly enriched in cell adhesion, cellular response to interleukin-1, positive regulation of transcription from DNA/RNA, and DNA binding. Whereas, for CC, these genes were enriched in extracellular matrix/region, anchored membrane component, cell junction and golgi membrane. As for MF, the upregulated genes were mainly enriched in sequence-specific DNA binding, transcription factor activity, RNA polymerase II regulatory region, and DNA binding.

On the other hand, GO analysis for downregulated genes showed that for BP, these DEGs were generated and enriched in the mitotic nuclear division, cell division, mitotic cell cycle/cell cycle transition, protein catabolic process and movement. Further, for CC, the downregulated genes exhibited enrichment in anaphase-promoting complex, kinesin complex, microtubule, bicular cellular tight junction and integral membrane component. Finally, for MF, these DEGs were mainly enriched in the ATP binding, microtubule motor activity/binding, protein serine/threonine kinase activity, serine-type endopeptidase activity and G-protein coupled receptor activity (Table 3).
| Expression     | Classification         | Term                                         | Count | \( p \)-Value | FDR  |
|----------------|------------------------|----------------------------------------------|-------|---------------|------|
| Upregulated    | GOTERM_BP_DIRECT       | GO:0007155 ~ cell adhesion                   | 5     | 2.20E-01      | 1.00E+00 |
|                |                        | GO:0071347 ~ cellular response to interleukin-1 | 4     | 6.10E-03      | 1.00E+00 |
|                |                        | GO:0045893 ~ positive regulation of transcription from DNA | 6     | 1.30E-01      | 1.00E+00 |
|                |                        | GO:0045944 ~ positive regulation of transcription from RNA | 10    | 7.10E-02      | 1.00E+00 |
|                |                        | GO:0001112 ~ transcription, DNA-templated    | 11    | 5.80E-01      | 1.00E+00 |
|                | GOTERM_CC_DIRECT       | GO:0031012 ~ extracellular matrix            | 7     | 5.00E-03      | 3.10E-01 |
|                |                        | GO:0005576 ~ extracellular region            | 15    | 4.60E-02      | 1.00E+00 |
|                |                        | GO:0031225 ~ anchored component of membrane  | 6     | 3.40E-04      | 4.20E-02 |
|                |                        | GO:0030054 ~ cell junction                   | 4     | 4.40E-01      | 1.00E+00 |
|                |                        | GO:0000139 ~ golgi membrane                  | 5     | 3.90E-01      | 1.00E+00 |
|                | GOTERM_MF_DIRECT       | GO:0043565 ~ sequence-specific DNA binding   | 9     | 6.00E-03      | 1.00E+00 |
|                |                        | GO:0001010 ~ transcription factor activity   | 7     | 4.00E-01      | 1.00E+00 |
|                |                        | GO:1903026 ~ RNA polymerase II regulatory region | 3     | 3.00E-01      | 1.00E+00 |
|                |                        | GO:0001010 ~ transcription factor activity   | 7     | 4.00E-01      | 1.00E+00 |
|                |                        | GO:0003677 ~ DNA binding                     | 6     | 9.50E-01      | 1.00E+00 |
| Downregulated  | GOTERM_BP_DIRECT       | GO:0007067 ~ mitotic nuclear division        | 13    | 6.60E-08      | 5.80E-05 |
|                |                        | GO:0051301 ~ cell division                   | 14    | 3.90E-07      | 1.70E-04 |
|                |                        | GO:0000278 ~ mitotic cell cycle              | 3     | 9.50E-03      | 4.20E-01 |
| Expression | Classification | Term                                      | Count | \(p\)-Value | FDR  |
|------------|----------------|-------------------------------------------|-------|--------------|------|
|            |                | GO:0000278 ~ mitotic cell cycle            | 4     | 1.10E-02     | 4.30E-01 |
|            |                | GO:0044772 ~ mitotic cell cycle transition | 4     | 1.30E-02     | 5.00E-01 |
|            |                | GO:0043161 ~ protein catabolic process     | 4     | 1.40E-01     | 9.90E-01 |
|            |                | GO:0007018 ~ microtubule-based movement    | 3     | 9.60E-02     | 9.90E-01 |
|            | GOTERM_CC_DIRECT | GO:0005680 ~ anaphase-promoting complex   | 3     | 9.00E-03     | 1.60E-01 |
|            |                | GO:0005871 ~ kinesin complex               | 3     | 4.30E-02     | 3.40E-01 |
|            |                | GO:0015630 ~ microtubule                   | 3     | 5.80E-01     | 9.60E-01 |
|            |                | GO:0005923 ~ bicellular tight junction     | 3     | 1.60E-01     | 7.00E-01 |
|            |                | GO:0016021 ~ integral component of membrane | 30   | 7.80E-01     | 9.60E-01 |
|            | GOTERM_MF_DIRECT | GO:0005524 ~ ATP binding                   | 15    | 9.00E-02     | 1.00E+00 |
|            |                | GO:0003777 ~ microtubule motor activity    | 3     | 9.20E-02     | 1.00E+00 |
|            |                | GO:0008017 ~ microtubule binding           | 3     | 3.80E-01     | 1.00E+00 |
|            |                | GO:0004674 ~ protein serine/threonine kinase activity | 6     | 9.10E-02     | 1.00E+00 |
|            |                | GO:0004252 ~ serine-type endopeptidase activity | 4     | 2.20E-01     | 1.00E+00 |
|            |                | GO:0004930 ~ G-protein coupled receptor activity | 4     | 8.30E-01     | 1.00E+00 |

KEGG analysis revealed that 12 of these genes (CDC6, CCND2, CCNE1, BUB1B, CCNB2, BUB1, SFN, TTK, CDC20, PTTG1, CDK1 and CDKN2A) were particularly enriched in the cell cycle pathway, 7 genes (CCND2, CCNE1, CCNB2, SFN, RRM2, CDK1 and CDKN2A) in p53 signaling pathway and likewise, 7 genes (AURKA, CCNE1, CCNB2, BUB1, CDC20, PTTG1, CDK1) were significantly enriched in the oocyte meiosis pathway (\(p<0.05\)) (Table 4).
Table 4
KEGG-based pathway analysis of DEGs related to ovarian cancer

| Pathway ID | Description                  | Count | %       | p-Value       | Genes                                                                 |
|------------|------------------------------|-------|---------|---------------|----------------------------------------------------------------------|
| hsa04110   | Cell cycle                   | 12    | 9.75    | 9.57E-06      | CDC6, CCND2, CCNE1, BUB1B, CCNB2, BUB1, SFN, TTK, CDC20, PTTG1, CDK1, CDKN2A |
| hsa04115   | p53 signaling pathway        | 7     | 10.29   | 0.0023        | CCND2, CCNE1, CCNB2, SFN, RRM2, CDK1, CDKN2A                           |
| hsa04114   | Oocyte meiosis               | 7     | 6.03    | 0.0345        | AURKA, CCNE1, CCNB2, BUB1, CDC20, PTTG1, CDK1                         |

PPI network construction and module analysis

A total of 221 (up/down-regulated) genes with significant enrichment in GO analysis were then subjected to the PPI network analysis. The PPI network of these DEGs consisted of 374 nodes and 626 edges (Fig. 2A). Next, to find highly interconnected regions (clusters), we used MCODE app in Cytoscape which revealed a cluster that consisted of a total of 14 genes (Fig. 2B).

Analysis of candidate genes by Kaplan Meier plotter and GEPIA

The survival data of 14 candidate genes was acquired using KM-plotter (https://kmplot.com/analysis). The results showed that 9 genes exhibited a significantly (p < 0.05) worse survival effect while the remaining 5 didn't show a significant effect (p > 0.05) related to survival (Fig. 3). In the next step, we compared the mRNA expression levels between normal and OC-affected persons via GEPIA. We found that 13 of 14 analyzed genes had a significantly higher expression (p > 0.05) in OC tissues, compared to the normal tissues (Fig. 4).

Validation of KEGG pathway enrichment results

To gain further insights into possible pathways of 13 DEGs with higher expression, we re-analyzed them using DAVID software. Results showed that 11 DEGs (CDC6, CCNE1, BUB1B, CCNB2, BUB1, SFN, TTK, CDC20, PTTG1, CDK1, CDKN2A) were significantly enriched in the cell cycle pathway (p = 1.6E-16) (Fig. 5 & Table 5) while two (SFN and RRM2) were significantly enriched ((p = 1.3E-8) in p53 signaling pathway (Fig. 6 & Table 5).

Table 5
Repeated KEGG pathway analysis of 14 selected genes

| Pathway ID | Description                  | Count | %       | p-Value       | Genes                                                                 |
|------------|------------------------------|-------|---------|---------------|----------------------------------------------------------------------|
| hsa04110   | Cell cycle                   | 11    | 78.5    | 1.6E-16       | CDC6, CCNE1, BUB1B, CCNB2, BUB1, SFN, TTK, CDC20, PTTG1, CDK1, CDKN2A |
| hsa04115   | p53 signaling pathway        | 2     | 23.3    | 1.3E-8        | SFN, RRM2                                                             |

Discussion
The present study represents a focused effort to identify potential OC-related biomarkers that could be useful in effective OC prognosis. We used integrated bioinformatics tools to analyze three microarray datasets (GSE14407, GSE36668 and GSE18520) and processed the data from a total of 73 cancerous (OC) and 26 normal ovarian tissue samples. The analysis with GEO2R revealed a total of 232 common DEGs ($p$-value < 0.05 and $|\log FC| > 2$), including 108 upregulated ($\log FC > 0$) and 124 downregulated ($\log FC < 0$) genes. Further, GO and pathway enrichment analysis via DAVID showed that for BP, the upregulated genes were involved in cell adhesion, cellular response to interleukin-1, positive regulation of transcription from DNA/RNA, and transcription. Whereas, for CC, these genes were enriched in extracellular matrix/region, anchored membrane component, cell junction and golgi membrane. As for MF, the upregulated genes were mainly enriched in sequence-specific DNA binding, transcription factor activity, RNA polymerase II regulatory region, and DNA binding. The KEGG pathway analysis showed that 26 genes were mainly involved in cell cycle, p53 signaling and oocyte meiosis pathways. Re-analysis of these candidate genes that 13 DEGs were related to significantly higher expression in OC tissues compared to the normal tissues. Of these, 11 genes (CDC6, CCNE1, BUB1B, CCNB2, BUB1, SFN, TTK, CDC20, PTTG1, CDK1 and CDKN2A) were found to be significantly generated and enriched ($p<0.05$) in the cell cycle pathway, while 2 genes (SFN and RRM2) were significantly enriched ($p<0.05$) in p53 signaling pathway. Our findings reveal new potential candidate genes that could further be targeted for improved ovarian cancer prognosis.

The cell division cycle 6 or CDC6 is a novel initiator of DNA replication and it plays a vital role in the initiation and regulation of the cell cycle [8]. Several studies have reported that CDC6 plays a key role in human cancers including squamous cell carcinoma (SCC) of head and neck, nasopharyngeal carcinoma and lung cancer [9–11]. Remarkably, a significantly higher expression of CDC6 has been associated with epithelial ovarian cancer (EOC) tissues compared to the normal ones [12]. The prognostic significance of this gene in OC and other forms of cancers [13] makes it important target for therapeutic strategies.

Furthermore, our results revealed additional genes like cyclin-dependent kinase 1 (CDK1), cyclin E1 (CCNE1) and cyclin-dependent kinase inhibitor 2A (CDKN2A) that exhibited significantly higher expression in the OC tissues. A recent study by Yunoki and coworkers demonstrated that these three genes were significantly upregulated in the sebaceous gland carcinoma (SGC) of eyelid [14]. Interestingly, CDKN2A is the most widely studied gene for its tumor suppressive activity. Also, any mutation in this gene or disruption in its functional regulation are frequently associated with different types of cancers in human [15, 16].

BUB1 is a serine/threonine kinase that binds centromeres during the process of mitosis. Studies have demonstrated that upregulation of BUB1 is associated with various types of human cancers and subsequently their clinical prognosis [17]. Another study reported that a positively higher percentage of BUB1 protein denotes an advanced stage and higher degree of differentiation in the endometrial carcinoma patients [18]. On the other hand, BUB1B, a mammalian homolog of yeast Mad3, has been shown to elevate the proliferation of tumor and is related to worse survival rates in different forms of cancers including breast, colorectal, prostate and gastric cancer [19–21].

Remarkably, the higher expression of stratafin (SFN) has been designated as a universal abnormality and it is associated with progression of lung adenocarcinoma [22]. In case of human ovarian cancer, the prognostic importance of SFN has been demonstrated earlier. For example, regarding the clinical significance of SFN
mRNA, a study has shown that SFN is a cell cycle-related checkpoint gene that is associated with oncogenesis. The higher expression of SFN was observed in different cells of the OC patients and it was reported that higher expression of SFN has association with age and cancer levels [23]. In short, SFN plays a vital role in the regulation of cell cycle and OC pathogenesis [24].

The RRM2 or ribonucleotide reductase subunit M2 is associated with 2p25 chromosome which lacks the structural variations in cervical cancer samples [25]. In addition to being a potential indicator of poor prognosis [26–30], the overexpression of RRM2 has frequently been observed in various forms of cancers including gastric, lung, adrenocortical, nasopharyngeal cancer and neuroblastoma [31–34]. While in our analysis we found that RRM2 is associated with p53 signaling pathway. Given the well-known role of p53 signaling pathways in cancers [35], it could be a potential prognostic biomarker in case of OC.

As several studies have demonstrated that the variable expression of the above-described genes is not only associated with different forms of cancers (including OC), but also, these genes could be a potential candidates for improved OC prognosis. While our study provides the useful information on potential OC biomarkers, in future, these findings should be validated by proper experimentation.

**Conclusion**

Taken together, our comparative bioinformatics analysis of normal and OC tissues revealed a total of 12 DEGs (CDC6, CCNE1, BUB1B, CCNB2, BUB1, SFN, TTK, CDC20, PTTG1, CDK1, RRM2 and CDKN2A) that were significantly generated and enriched in the cell cycle and p53 signaling pathways. As supported by several studies, these genes could be further targeted for their OC prognostic value. The findings of our study will be useful to investigate the pathogenesis of the OC and to develop better prognosis of improved therapeutic significance.

**Methods**

**Acquisition of microRNA expression datasets**

In this study, we took advantage of the free, public database and obtained three datasets (GSE14407, GSE36668 and GSE18520) from online GEO (NCBI) database (https://www.ncbi.nlm.nih.gov/geo/). A total of 73 OC samples were compared with 26 normal ovarian samples originating from all datasets. The detailed information on the annotation platform, number and type of analyzed samples is given in the Table 1.

**Data processing and identification of DEGs**

GEO2R is a web-based interactive tool that can efficiently perform the comparison of two different datasets originating from the same experimental conditions [36]. The comparative analysis of all datasets was performed in the above-mentioned online tool and subsequently, the results were exported in TXT format and processed via Microsoft Excel. We used Venn, an online tool for the identification of the common DEGs. The DEGs with a value of FC > 0 were considered as upregulated genes, while the DEGs with FC < 0 were designated as downregulated genes.

**Gene ontology and functional enrichment analysis of DEGs**
GO functional analysis [37] and KEGG pathway enrichment analysis [38] were performed to predict the potential functions of the candidate DEGs via DAVID (https://david.ncifcrf.gov/tools.jsp). A cut-off value of $p < 0.05$ was implemented to rule out noteworthy BP, CC and MF of DEGs.

**Construction of PPI network and module analysis**

To evaluate the PPI information of the target DEGs, online Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) was used (https://string-db.org/) with a confidence score > 0.4. Subsequently, we used the MCODE app in Cytoscape [39] to find out the PPI network modules (degree cutoff = 2, depth = 100, k-core = 2 and node cutoff value = 0.2).

**Survival analysis and validation of expression of the candidate genes**

To assess the effect of multiple genes on survival based on GEO, we used a well-known web-based tool, Kaplan Meier-plotter [40]. The hazard ratio (HR) and logrank P value with confidence interval (CI) of 95% were calculated and displayed on the plot. Furthermore, the validation of candidate DEGs was performed via GEPIA web tool based on the RNA-seq expression data from thousands of TCGA and GTEx projects [41].

**Declarations**

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Funding**

This study did not receive any funding.

**Author contribution**

Zhang XY, the first author, conceptualized and designed the research, performed the experiments, analyzed the data, and was a major contributor in writing the manuscript. Zhu SH contributed in technical support. Peng M performed the experiments and critically revised the manuscript, Ma HB, the corresponding author, conceptualized and designed the research, supervised the research, and critically revised the manuscript. All authors read and approved the final manuscript.

**Acknowledgments**

Not applicable.
Availability of data and material

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics. 2020. CA: A Cancer Journal for Clinicians. 2020;70:7–30.
2. Cormio G, Rossi C, Cazzolla A, Resta L, Loverro G, Greco P, et al. Distant metastases in ovarian carcinoma. Int J Gynecol Cancer. 2003;13:125–9.
3. Weidle UH, Birzèle F, Kollmorgen G, Rueger R. Mechanisms and Targets Involved in Dissemination of Ovarian Cancer. Cancer Genomics Proteomics. 2016;13:407–23.
4. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Díaz LA Jr, Kinzler KW. Cancer genome landscapes. Science. 2013;339:1546–58.
5. Kumar SU, Kumar DT, Siva R, Doss CGP, Zayed H. Integrative Bioinformatics Approaches to Map Potential Novel Genes and Pathways Involved in Ovarian Cancer. Frontiers in Bioengineering Biotechnology. 2019;7:391.
6. Yang D, He Y, Wu B, Deng Y, Wang N, Li M, et al. Integrated bioinformatics analysis for the screening of hub genes and therapeutic drugs in ovarian cancer. Journal of Ovarian Research. 2020;13:10.
7. Zheng MJ, Li X, Hu YX, Dong H, Gou R, Nie X, et al. Identification of molecular marker associated with ovarian cancer prognosis using bioinformatics analysis and experiments. Journal of cellular physiology. 2019;234:11023–36.
8. Kim YH, Byun YJ, Kim WT, Jeong P, Yan C, Kang HW, et al. CDC6 mRNA Expression Is Associated with the Aggressiveness of Prostate Cancer. J Korean Med Sci. 2018;33(47):e303.
9. Chen X, Chen S, Pei N, Mao Y, Wang S, Yan R, et al. AAV-Mediated angiotensin 1–7 overexpression inhibits tumor growth of lung cancer in vitro and in vivo. Oncotarget. 2016;8:354–63.
10. Liu M, Zhu K, Qian X, Li W. Identification of miRNA/mRNA-negative regulation pairs in nasopharyngeal carcinoma. Medical science monitor: international medical journal of experimental clinical research. 2016;22:2215.
11. Saloura V, Vougiouklakis T, Zewde M, Kiyotani K, Park J-H, Gao G, et al. WHSC1L1 drives cell cycle progression through transcriptional regulation of CDC6 and CDK2 in squamous cell carcinoma of the head and neck. Oncotarget. 2016;7:42527.
12. Deng Y, Jiang L, Wang Y, Xi Q, Zhong J, Liu J, et al. High expression of CDC6 is associated with accelerated cell proliferation and poor prognosis of epithelial ovarian cancer. Pathol Res Pract. 2016;212:239–46.
13. Mahadevappa R, Neves H, Yuen SM, Bai Y, McCrudden CM, Yuen HF, et al. The prognostic significance of Cdc6 and Cdt1 in breast cancer. Sci Rep. 2017;7:985.

14. Yunoki T, Hirano T, Tabuchi Y, Furusawa Y, Torigoe M, Nakajima T, et al. CDKN2A, CDK1, and CCNE1 overexpression in sebaceous gland carcinoma of eyelid. Int Ophthalmol. 2020;40:343–50.

15. Foulkes WD, Flanders TY, Pollock PM, Hayward NK. The CDKN2A (p16) Gene and Human Cancer. Mol Med. 1997;3:5–20.

16. Jiao Y, Feng Y, Wang X. Regulation of Tumor Suppressor Gene CDKN2A and Encoded p16-INK4a Protein by Covalent Modifications. Biochemistry. 2018;83:1289–98.

17. Wang Z, Katsaros D, Shen Y, Fu Y, Canuto EM, Benedetto C, et al. Biological and Clinical Significance of MAD2L1 and BUB1, Genes Frequently Appearing in Expression Signatures for Breast Cancer Prognosis. PLoS One. 2015;10:e0136246.

18. Zhao Q, Bian AP, Zhang Y, Qin L, Shi HR, Su K. Expression of budding uninhibited by benzimidazoles-1 and mitotic arrest deficient-2 in endometrial carcinoma and its significance. Eur J Gynaecol Oncol. 2014;35:44–7.

19. Fu X, Chen G, Cai Z-D, Wang C, Liu Z-Z, Lin Z-Y, et al. Overexpression of BUB1B contributes to progression of prostate cancer and predicts poor outcome in patients with prostate cancer. Onco Targets Ther [Internet]. 2016 2016; 9:[2211-20 pp].

20. Hahn MM, Vreede L, Bemelmans SA, van der Looij E, van Kessel AG, Schackert HK, et al. Prevalence of germline mutations in the spindle assembly checkpoint gene BUB1B in individuals with early-onset colorectal cancer. Genes Chromosomes Cancer. 2016;55:855–63.

21. Hudler P, Britovsek NK, Grazio SF, Komel R. Association between polymorphisms in segregation genes BUB1B and TTK and gastric cancer risk. Radiol Oncol. 2016;50:297–307.

22. Shiba-Ishii A, Kano J, Morishita Y, Sato Y, Minami Y, Noguchi M. High expression of stratifin is a universal abnormality during the course of malignant progression of early-stage lung adenocarcinoma. Int J Cancer. 2011;129:2445–53.

23. Hu Y, Zeng Q, Li C, Xie Y. Expression profile and prognostic value of SFN in human ovarian cancer. Biosci Rep. 2019;39:BSR20190100.

24. Sarno J, Welch W, Mok S, Garner E. Stratifin expression in epithelial ovarian cancer: the role of methylation in vitro and in vivo. AACR; 2007.

25. Wang N, Zhan T, Ke T, Huang X, Ke D, Wang Q, et al. Increased expression of RRM2 by human papillomavirus E7 oncoprotein promotes angiogenesis in cervical cancer. Br J Cancer. 2014;110:1034–44.

26. Dai L, Lin Z, Qiao J, Chen Y, Flemington EK, Qin Z. Ribonucleotide reductase represents a novel therapeutic target in primary effusion lymphoma. Oncogene. 2017;36:5068–74.

27. D'Angiolella V, Donato V, Forrester FM, Jeong YT, Pellacani C, Kudo Y, et al. Cyclin F-mediated degradation of ribonucleotide reductase M2 controls genome integrity and DNA repair. Cell. 2012;149:1023–34.

28. Kang W, Tong JH, Chan AW, Zhao J, Wang S, Dong Y, et al. Targeting ribonucleotide reductase M2 subunit by small interfering RNA exerts anti-oncogenic effects in gastric adenocarcinoma. Oncol Rep. 2014;31:2579–86.
29. Morikawa T, Maeda D, Kume H, Homma Y, Fukayama M. Ribonucleotide reductase M2 subunit is a novel diagnostic marker and a potential therapeutic target in bladder cancer. Histopathology. 2010;57:885–92.

30. Shah KN, Wilson EA, Malla R, Elford HL, Faridi JS. Targeting Ribonucleotide Reductase M2 and NF-κB Activation with Didox to Circumvent Tamoxifen Resistance in Breast Cancer. Mol Cancer Ther. 2015;14:2411–21.

31. Grolmusz VK, Karászi K, Micsik T, Tóth EA, Mészáros K, Karvaly G, et al. Cell cycle dependent RRM2 may serve as proliferation marker and pharmaceutical target in adrenocortical cancer. Am J Cancer Res. 2016;6:2041–53.

32. Grossi F, Dal Bello MG, Salvi S, Puzone R, Pfeffer U, Fontana V, et al. Expression of Ribonucleotide Reductase Subunit-2 and Thymidylate Synthase Correlates with Poor Prognosis in Patients with Resected Stages I-II Non-Small Cell Lung Cancer. Dis Markers. 2015;2015:302649.

33. Han P, Chen RH, Wang F, Zeng JY, Yu ST, Xu LH, et al. Novel chimeric transcript RRM2-c2orf48 promotes metastasis in nasopharyngeal carcinoma. Cell Death Dis. 2017;8(9):e3047.

34. Zhong Z, Cao Y, Yang S, Zhang S. Overexpression of RRM2 in gastric cancer cell promotes their invasiveness via AKT/NF-κB signaling pathway. Pharmazie. 2016;71:280–4.

35. Harris SL, Levine AJ. The p53 pathway: positive and negative feedback loops. Oncogene. 2005;24:2899–908.

36. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets–update. Nucleic Acids Res. 2013;41(Database issue):D991-5.

37. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25:25–9.

38. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28:27–30.

39. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498–504.

40. Szász AM, Lánczky A, Nagy Á, Förster S, Hark K, Green JE, et al. Cross-validation of survival associated biomarkers in gastric cancer using transcriptomic data of 1,065 patients. Oncotarget. 2016;7:49322–33.

41. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res. 2017;45:W98-w102.

Figures
Venn diagrams represent common DEGs among 3 datasets (GSE14407, GSE36668 and GSE18520); (a) 108 upregulated (logFC>0) and (b) 124 downregulated (logFC<0) genes. Different colors represent different datasets. The diagrams were generated through online software (http://bioinformatics.psb.ugent.be/webtools/Venn/).
Figure 2

PPI network construction by STRING tool and modular analysis; (a) a total of 221 DEGs composed the PPI network. The nodes denote proteins while edges represent the protein interactions, (b) Result of the MCODE-based module analysis using Cytoscape (k-core = 2, max. depth = 100, node score cutoff = 0.2 and degree cutoff = 2)
The online tool KM-plotter was used to acquire the prognostic information of 13 core genes. A total of 9 genes exhibited a significantly (p < 0.05) worse survival rate while 4 genes did not have a significant survival rate (p > 0.05).

Figure 3
Figure 4

The comparative analysis of significant gene expression between normal and OC-affected patients was performed using an online web resource, GEPIA. A total of 13 genes were significantly upregulated in the OC tissues compared to the normal ovarian tissues (*p > 0.05). The grey color denotes normal tissues and the red color represents the tumor tissues.
Figure 5

Repeated KEGG pathway analysis of 14 significantly upregulated DEGs from OC tissues. A total of 11 genes were generated and significantly enriched in the cell cycle pathway in G1, S G2 and M phases. Ink4a and ARF represent CDKN2A, CycE means CCNE1, Mps1 denotes TTK, and CybB means CCNB2.
Repeated KEGG pathway analysis of 14 significantly upregulated DEGs from OC tissues. A total of 06 genes were generated and significantly enriched in the p53 signaling pathway in different phases. In addition to the gene associated with cell cycle pathway, p53R2/RRM2 and 14-3-3-σ/SFN were involved in p53 signaling pathway.