

# Differences in Gene Expression Between High- and Low-Grade Serous Ovarian Cancers: Implications for Diagnosis and Prognosis

**Jianhao Xu**

Kunshan First People's Hospital <https://orcid.org/0000-0002-2064-4597>

**Qian Wang**

Suzhou University: Soochow University

**Fang Cao**

Kunshan First People's Hospital

**Zhiyong Deng**

Kunshan First People's Hospital

**Xiaojiao Gao**

Kunshan First People's Hospital

**Tingting Gu**

Kunshan First People's Hospital

**Tingting Liu**

Kunshan First People's Hospital

**Song Xu** (✉ [xs19780116@163.com](mailto:xs19780116@163.com))

Kunshan First People's Hospital

**Wenjuan Gan**

Dushu Lake Hospital Affiliated to Soochow University

---

## Research

**Keywords:** high-grade serous ovarian cancer, low-grade serous ovarian cancer, diagnosis, prognosis, differentially expressed genes

**Posted Date:** October 19th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-963632/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

The clinical presentations of high-grade serous ovarian cancer (HGSOC) and low-grade serous ovarian cancer (LGSOC) differ. In this study, we aimed to identify the essential molecules for the diagnosis and prognosis of these OC subtypes.

## Methods

Differentially expressed genes (DEGs) between HGSOC and LGSOC were identified using three GEO series. The functional enrichment analysis was performed to investigate different biological processes and pathways. The protein–protein interaction network was constructed, and hub genes were screened to narrow the focus of the study. The discovered hub genes were frequently validated using prognostic correlation, co-expression, and immunohistochemistry (IHC) in GTEx, Oncomine, GEPIA, cBioportal, HPA, and other databases.

## Results

In comparison with LGSOC, 79 upregulated genes and 85 downregulated genes were identified in HGSOC, and the biological roles of these genes were mostly centered on the cell cycle process and chromosomal segregation. Among the 10 hub genes chosen, BIRC5 was favorably linked with overall survival of patients with ovarian cancer ( $p = 0.014$ ), whereas RRM2 was negatively correlated with the ovarian cancer stage ( $p = 0.0251$ ). In IHC studies, the intensity of BIRC5 expression in ovarian cancer was greater than that in normal ovarian tissues; however, RRM2 was not substantially expressed in either ovarian cancer tissues or normal ovarian tissues.

## Conclusions

BIRC5 is a potential marker that can distinguish HGSOC from LGSOC, guide prognosis, and be utilized in clinical IHC.

### 1. Background

Serous carcinoma, mucinous carcinoma, endometrioid carcinoma, and clear cell carcinoma are the histological subtypes of epithelial ovarian cancer (OC)(1). Serous OC, the most frequent subtype, is further subdivided into high-grade and low-grade serous OCs (abbreviated as HGSOC and LGSOC, respectively)(2). The clinical characteristics of the two serous OCs differ(2). For example, we used the Genomic Data Commons (GDC) database to evaluate survival data for HGSOC and unclassified SOC and found that patients with HGSOC exhibited worse disease-free survival (DFS,  $p=0.0151$ , Figure 1A). Despite

a trend in overall survival (OS) analysis, the difference was not significant ( $p=0.1370$ , Figure 1B). The data set GSE151335 was used to verify the differences in progress-free survival (PFS) and OS between HGSOC and LGSOC. Platform for GSE151335 is GPL28589 [Oxford Classifier of Carcinoma of the Ovary]. Patients with HGSOC exhibited worse PFS ( $p=0.011$ , Figure 1C) and OS ( $p=0.0027$ , Figure 1D).

Furthermore, the chemotherapy regimens recommended for HGSOC and LGSOC are not differentiated in the NCCN recommendations (2021 version) (<https://www.nccn.org/guidelines/guidelines-detail?category=1&id=1453>). Therefore, treatment resistance in patients with HGSOC and LGSOC is comparable. According to research, LGSOC is often resistant to chemotherapy; however, the response of HGSOC to chemotherapy varies(3). It implies that even in the HGSOC patient population, the prognosis may be examined in depth across different subgroups. In conclusion, pathological differentiation of HGSOC and LGSOC is particularly important due to differences in clinical manifestations and prognosis between these OC subtypes. Additionally, the screening of the molecules that can guide prognosis and clinical application of these molecules in the pathology department may provide references for the treatment and follow-up of patients after surgery.

HGSOC is now known to be linked with TP53 mutations, whereas LGSOC is associated with BRAF and KRAS mutations as the benign/borderline adenoma progression(4, 5). Given that BRAF and KRAS mutations are often detected through next-generation sequencing(6), time and economic expenses associated with LGSOC diagnosis are high. Immunohistochemistry (IHC) techniques are more commonly utilized for differential diagnosis in real-world clinical settings. TP53 is a precise marker for the differential diagnosis of HGSOC(7), although it occasionally produces errors in clinical practice. Furthermore, the Gene Expression Profiling Interactive Analysis (GEPIA) website indicates that TP53 has little predictive value for the prognosis of OC patients. Thus, the major clinical objective of this research is to identify critical molecules that can distinguish diagnosis and inform prognosis based on HGSOC and LGSOC expression differences ( $pHR=0.93$ , Supplementary Figure 1).

Therefore, the primary clinical goal of this research is to screen key molecules that can differentiate diagnosis and guide prognosis based on differential genes expression of HGSOC and LGSOC.

## 2. Methods

### 2.1 Microarray data

Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo>) functions as a public functional genomics database of high throughput gene expression data, chips, and microarrays(8). Microarray data of the human OC dataset were screened through the GEO database. The search strategy was as follows: ("ovarian neoplasms"[MeSH Terms] OR ovarian cancer[All Fields]) AND "Homo sapiens"[porgn] AND ("gse"[Filter] AND "Expression profiling by array"[Filter] AND "attribute name tissue"[Filter]). Three gene expression profiles in HGSOC and LGSOC tissues were screened and downloaded from a total of 136 GEO Series (GSEs), namely GSE73638, GSE27651 and GSE14001. Platforms for GSE73638 is GPL20967

[ADXOCv1a520630] Almac OC Disease Specific Array, while for GSE27651 and GSE14001 is GPL570[HG-U133\_Plus\_2] Affymetrix Human Genome U133 Plus 2.0 Array. GSE73638 contains 13 HGSOC samples and 7 LGSOC samples. GSE27651 contains 21 HGSOC samples and 13 LGSOC samples. GSE14001 contains 10 HGSOC samples and 10 LGSOC samples.

## 2.2 Identification of DEGs

GEO2R is regarded as an interactive online tool that has been designed to compare two or more datasets in the GEO series to identify differentially expressed genes (DEGs)(9). DEGs between HGSOC and LGSOC tissues were identified using GEO2R, with the cut-off points of  $\text{adj.p} < 0.05$  and  $|\log\text{FC}| > 1$ . DEGs in GSE14001 were identified only with the threshold of  $\text{adj.p} < 0.05$  due to the limited DEGs by the aforementioned threshold to ensure sufficient DEGs for subsequent analyses. Therefore, we first detected the intersection DEGs between GSE73638 and GSE27651 and then used GSE14001 for a second confirmation. We used online Wien software to detect the intersection DEGs between the datasets. Volcano and the heat map were painted to picture these DEGs by using the R ggplot2 and heat map package.

## 2.3 Function enrichment analysis

The Gene Ontology (GO) datasets and Kyoto Encyclopedia of Gene and Genome (KEGG) pathway enrichment were used to analyze DEGs at the functional level with WEB-based Gene Set Analysis Toolkit (WebGestalt, <http://www.webgestalt.org/option.php>, version 2019). WebGestalt is a functional enrichment analysis web tool, which has on average 26,000 unique users from 144 countries and territories per year according to Google Analytics(10). The results were pictured in the R ggplot2 package.  $P < 0.05$  was considered statistically significant.

## 2.4 GSEA analysis

Gene Set Enrichment Analysis (GSEA) is a computational method that determines whether an a priori defined set of genes shows statistically significant, concordant differences between two biological states (e.g. HGSOC and LGSOC)(11). Since general GO and pathway analysis focus on a few genes that are significantly up-regulated or down-regulated, it is easy to miss some genes that are not significantly differentially expressed but have important biological significance. To make up for this defect, the GSEA approach was applied to reconfirm function enrichment analysis in GSEA software version 4.1.0, which uses predefined gene sets from the Molecular Signatures Database (MSigDB v7.4).

## 2.5 PPI network construction and module analysis

Search Tool for the Retrieval of Interacting Genes (STRING, <http://string-db.org>, version 11.5) is a database of known and predicted protein-protein interactions (PPI), which was applied to visualize the potential gene interaction network(12). The minimum required interaction score is set to 0.700, which is high confidence. Cytoscape (a public bioinformatics software, <https://cytoscape.org/>, version 3.8.2) is an open-source software platform for visualizing complex networks and integrating these with any type of attribute data(13). The plug-in Molecular Complex Detection (MCODE, version 2.0.0) app of Cytoscape

clusters a given network based on the topology to find densely connected regions with the following criteria: degree cutoff = 2, node score cutoff = 0.2, K-Core = 2, Max. Depth = 100(14).

## 2.6 Hub genes selection and analysis

Another plug-in app cytoHubba (version 0.1) of Cytoscape predicts and explores important nodes and sub-networks in a given network by different topological algorithms(15). In the present study, the top nodes were ranked by Radiality algorithms. The mountain map and the grouped violin map are used to observe the expression of the selected hub genes in GSE73638, GSE27651, and GSE14001, which were pictured by the Sangerbox tools (<http://www.sangerbox.com/tool>)(16).

First, differential expression of 10 hub genes in normal ovarian tissue and OC tissue was analyzed by Genotype-Tissue Expression (GTEx, <https://www.gtexportal.org/>) website(17) and Oncomine (<https://www.oncomine.org/>) database(18). Second, the OS and stage analysis of 10 hub genes in OC was conducted by GEPIA (<http://gepia.cancer-pku.cn/index.html>) database(19). Third, Mutation profiles of 10 hub genes in OC were analyzed by cBioportal (<https://www.cbioportal.org/>) database(20).

## 2.7 Further verification of selected genes with potential clinical value

According to the correlation between hub genes and the OS and stage of OC in GEPIA, Genes with potential clinical value were selected for further verification.

First, co-expression genes were analyzed using the Oncomine database(18), which were compared with overlapping DEGs.

Second, IHC verification was conducted using OC pathological tissue slices in Human Protein Atlas (HPA, <https://www.proteinatlas.org/>, Version: 20.1) database(21).

Third, the upstream target miRNAs were screened using miRWalk (<http://mirwalk.umm.uni-heidelberg.de/>, version new\_update\_2021) database(22), and Targetscan ([http://www.targetscan.org/vert\\_72/docs/help.html](http://www.targetscan.org/vert_72/docs/help.html), version 7.2) database(23). The intersection target miRNAs were verified by GSE61741 from GEO, which is miRNA Profiles in OC and controls. Platform for GSE61741 is GPL9040 [febit Homo Sapiens miRBase 13.0]. GSE61741 contains 24 OC blood samples and 94 normal blood samples.

## 3. Results

Supplementary Figure 2 summarizes the technical roadmap of this research, and the results are split into five sections as follows.

### 3.1 Identification of DEGs between HGSOC and LGSOC

Among the 136 GSEs screened, we evaluated a total of four series that comprised HGSOc and LGSOc broad gene expression microarray data, namely GSE73638, GSE73551, GSE27651, and GSE14001. Because the series GSE73638 and GSE73551 belong to the same study, the current study included only GSE73638 because it had a larger sample size. Table 1 summarizes the essential information for the three GSEs.

Table 1  
Sample details of the selected GEO Series.

GEO Series	Publication date	Platform	Samples	Source	Cell type	Case number
GSE73638	8-Nov-16	GPL20967	102	Primary tumor	Clear cell	12
					Endometrioid	9
					Mucinous	9
					Serous low-grade	7
					Serous high-grade	13
				Ovarian cells	52	
GSE27651	4-Mar-11	GPL570	49	Primary tumor	Serous low-grade	13
					Serous high-grade	21
					Low-malignant tumor of the ovary	9
					Normal ovarian surface epithelial cells	6
GSE14001	31-May-09	GPL570	23	Primary tumor	Serous low-grade	10
					Serous high-grade	10
					Normal ovarian surface epithelial cells	3

We utilized the GEO2R online program to identify 1465, 9914, and 230 distinct genes from GSE73638, GSE27651, and GSE14001, respectively (Figure 2A, adj.P 0.01,  $|\log_{2}FC| > 1$ ). The DEGs in the GSE14001 are restricted by the aforementioned threshold. To guarantee the availability of sufficient DEGs for further research, 9500 DEGs in GSE14001 were identified with a threshold adj.P 0.05 and utilized to confirm the overlapping DEGs between GSE73638 and GSE27651. According to the Venn diagram, overlapping DEGs between GSE73638 and GSE27651 comprised 157 upregulated and 204 downregulated genes (Figure 2B). Following validation by GSE14001, 79 upregulated and 85 downregulated genes were chosen for the

current study (Figure 2C). We used heat maps to depict the distribution of screened gene expression in each GSE between HGSOC and LGSOC (Figure 2D).

## 3.2 Enrichment analysis for DEGs

We used the WebGestalt web tool to perform GO and KEGG enrichment analyses to identify the most important biological processes (BPs) and pathways. In total, 164 DEGs were mostly enriched in the BPs associated with mitotic cell cycle, organelle fission, and nuclear division (Figure 3A) and pathways such as hepatitis C, micro RNAs in cancer, and chronic myeloid leukemia (Figure 3B). In addition, we used GSEA to validate the GOBP and KEGG pathways. The GOBP or KEGG pathways that were substantially represented with a normalized p value of  $<0.05$  were presented. In the HGSOC group, 363, 8, and 31 BPs were enriched in GSE73638, GSE27651, and GSE14001, respectively, and three BPs were enriched in all three GSEs (Figure 3C). In the LGSOC group, 88, 78, and 83 BPs were enriched in GSE73638, GSE27651, and GSE14001, respectively, whereas 10 BPs were enriched in two GSEs (Figure 3C). Of the three GSEs, only GSE73638 was enriched in 13 KEGG pathways in the HGSOC group (Figure 3D). In the LGSOC group, the number of KEGG pathways enriched in GSE73638, GSE27651, and GSE14001 was 1, 1, and 5, respectively, and no overlap was found among the GSEs (Figure 3D).

Meiotic cell cycle process, homologous chromosome segregation, and meiosis I cell cycle process were the BPs confirmed by the three GSEs (Figure 4). Cell cycle process and chromosomal segregation were the GOBPs confirmed through WebGestalt and GSEA. Chemokine signaling route and oocyte meiosis were the KEGG pathways confirmed through WebGestalt and GSEA.

## 3.3 PPI network construction and significant module identification

We utilized the String database to estimate the protein-level connection of the overlapped DEGs (Figure 5A). We improved the visualization with Cytoscape software and constructed a PPI network with 115 nodes and 894 edges (Figure 5B). MCODE was used to divide the PPI network into four modules (Figure 5C); the first module had 38 nodes and 661 edges (MCODE score 35.730), the second module had 7 nodes and 21 edges (MCODE score 7.000), the third module had 5 nodes and 10 edges (MCODE score 5.000), and the fourth module had 3 nodes and 3 edges (MCODE score 3.000). We utilized Cytohubba to filter the top 10 Hubba nodes, namely BIRC5, CDC20, CDK1, CDKN3, MKI67, NUSAP1, RRM2, TOP2A, TPX2, and UBE2C, for additional investigation using Radiality topological techniques (Figure 5D).

The expression differences between HGSOC and LGSOC could be displayed more naturally in the three GSEs by using the mountain map (Supplementary Figure 3A) and the grouped violin map (Supplementary Figure 3B).

## 3.4 Analysis of the hub genes

First, we used the GTEx website and the Oncomine database to evaluate the differential expression of 10 hub genes in normal ovarian tissues and OC tissues. Except for NUSAP1, which was moderately

expressed in normal ovarian tissues, the expression levels of the other nine hub genes were low (Figure 6A). Other normal female reproductive organs, such as the breast, cervix, and endometrium, exhibited a minimal expression of these 10 hub genes. All the 10 hub genes were highly expressed in OC tissues, similar to those in other female common malignancies such as breast cancer and cervical cancer (Figure 6B).

Second, we used the GEPIA database to investigate the connection between the 10 hub genes and OC OS/staging. Among the 10 hub genes, only BIRC5 was found to be favorably linked with OS in OC ( $p_{HR} = 0.014$ , Figure 7A), and only RRM2 was found to be negatively correlated with OC staging ( $p = 0.0251$ , Figure 7B).

Third, we used the cBioportal website to examine the alterations of the 10 OC hub genes. The oncoPrint indicated that the amplification of BIRC5, RRM2, and CDC20 has no evident concomitance, whereas the amplification of CDK1, MKI67, UBE2C, and TPX2 has a possible concomitance (Figure 8A). In OC, BIRC5 displayed more visible amplification and gain mutations, whereas RRM2 displayed less visible amplification and gain mutations (Figure 8B).

### **3.5 Re-analysis of the genes with a potential clinical value**

Based on the aforementioned results, BIRC5 and RRM2 with potential therapeutic utility were chosen from a list of 10 hub genes for further investigations.

First, we examined the co-expressed genes of BIRC5 and RRM2 in the Oncomine database and compared these genes with 164 overlapping DEGs. The DEGs, NUF2, KIF23, DEPDC1, TOP2A, and GPSM2, were found to co-express with BIRC5 (Supplementary Figure 4A), whereas DLGAP5, CDKN3, ZWINT, MELK, and CDC20 were found to co-express with RRM2 (Supplementary Figure 4B).

Second, we performed IHC on BIRC5 and RRM2 by using the HPA database. By using antibodies HPA002830 and CAB004270, we found that the expression of BIRC5 in OC tissues was greater than that in normal ovarian tissues (Figure 9A-B). RRM2 expression in OC tissues could not be identified with the antibody HPA056994 in both OC tissues and normal ovarian tissues (Figure 9A&C).

Third, we utilized the miRWalk and Targetscan databases to identify BIRC5 and RRM2 upstream miRNAs. The intersection of the two datasets revealed that 88 miRNAs regulate BIRC5 expression, whereas 76 miRNAs regulate RRM2 expression. Simultaneously, four miRNAs were found to control both BIRC5 and RRM2 expressions (Figure 10A). The mRNA–miRNA map was constructed using Cytoscape, with the size of the node representing the strength of the connection (Figure 10B). Then, we validated overlapping miRNAs by using GSE61741, which contains miRNA profiles from 24 OC blood samples and 94 normal blood samples. The GEO2R online tool was used to identify 185 differential miRNAs from GSE61741 (Figure 10C,  $adj.P < 0.05$ ). The Venn diagram (Figure 10D) depicts the distribution of intersection miRNAs for BIRC5 and RRM2 across the GSE61741, miRWalk database, and Targetscan database (Figure 10E). Although 5 overlapping miRNAs were screened, we inferred that the upstream miRNAs of BIRC5 and RRM2 exhibit low expressions in OC tissues due to the negative regulatory connection between miRNA

and mRNA. In addition, by using the heat map, hsa-miR-520a-3p, hsa-miR-1323, and hsa-miR-324-3p were found to match the requirements. Cytoscape was used to visualize the miRNA–mRNA map (Figure 10F).

## 4. Discussion

HGSOC and LGSOC have diverse clinical presentations, and the prognosis differs even within the same group. However, our literature survey indicated that only few studies have compared HGSOC with LGSOC. Thus, the goal of the present study was to identify and validate the essential molecules that can help in distinguishing diagnoses and guiding prognoses of the two OC subtypes.

We identified 164 robust DEGs by using the three GSEs. Based on WebGesalt and GSEA analysis, we discovered that the GOBPs differing between HGSOC and LGSOC are mainly the cell cycle process and chromosomal segregation, which is compatible with HGSOC's strong proliferative activity(24, 25). Chemokine signaling and oocyte meiosis are the two crucial KEGG pathways. Chen et al. discovered that the chemokine CCL20 increases paclitaxel resistance in CD44+CD117+ cells in OC through the Notch1 signaling pathway(26). Chiaramonte et al. discovered that the Notch pathway promotes OC development and migration through the CXCR4/SDF1alpha chemokine system(27). These findings highlight the growing relevance of the chemokine signaling system in OC research.

We then selected 10 hub genes out of 164 DEGs and conducted a series of studies on these genes. First, we used the GTEx and Oncomine databases to examine the expression of the 10 hub genes in normal ovarian tissues and OC tissues and observed substantial differences in the expression of all these genes, except NUSAP1. Overexpression of the remaining 9 hub genes, including BIRC5 and RRM2, might have a role in carcinogenesis. Second, we used GEPIA to examine OS and OC staging based on the 10 hub genes. We found that only BIRC5 is favorably linked with OS, and only RRM2 is inversely correlated with staging, which indicated the therapeutic significance of these two genes. Third, we used cBioportal to examine the mutations of the 10 OC hub genes and discovered that BIRC5 and RRM2 exhibit varying degrees of amplification and gain mutations. Consequently, we narrowed down our focus toward the essential genes BIRC5 and RRM2.

By using the Oncomine database, we identified the genes that were co-expressed with BIRC5 and RRM2 in OC and compared these genes with the 164 DEGs examined. We identified 5 DEGs that were co-expressed with BIRC5/RRM2, which indicated the dependability of the hub genes BIRC5 and RRM2. Furthermore, as stated in the introduction that the chosen molecules should be used in clinical IHC, we used the HPA website to assess the feasibility of BIRC5 and RRM2 as IHC markers. The findings indicated that BIRC5 can differentiate between normal ovarian tissues and OC tissues; however, RRM2 could not be detected. Finally, we investigated the upstream regulation of BIRC5 and RRM2 by using the miRNAs database. Simultaneously, GSE61741 was utilized to confirm the upstream regulatory miRNAs of BIRC5 and RRM2. The results demonstrated that hsa-miR-520a-3p and hsa-miR-1323 can downregulate BIRC5 expression in OC, whereas hsa-miR-324-3p can downregulate RRM2 expression in OC, which confirmed the

dependability of BIRC5 and RRM2 at the upstream regulatory level. In conclusion, BIRC5 offers a high clinical application value.

BIRC5 is an inhibitor of the apoptosis gene family that encodes negative regulatory proteins, which suppress apoptotic cell death. BIRC5 expression is high during fetal development and in most cancers but low in adult tissues(28). Yin et al. used bioinformatics tools to investigate the OC OVDM1 cell line and confirmed the clinical significance of BIRC5, which is consistent with the findings of our study(29). A meta-analysis by He X et al. indicated that the protein survivin is closely linked to FIGO staging and tumor grade of OC(30). BIRC5 has been considered a therapeutic target for cancers for more than 20 years(31). Ozreti et al. revealed that the Hedgehog signaling pathway is linked to OC pathogenesis and that BIRC5 might be a novel target of this pathway(32). Wang et al. utilized an orthotopic OC mouse model to confirm that miR-203 suppresses ovarian tumor metastasis by targeting BIRC5 to prevent EMT(33). In the present study, we discovered an intriguing phenomenon: OS of patients with HGSOC was found to be poorer than that of patients with LGSOC (Figure 1); however, BIRC5, a stronger predictor of OS, was found to overexpress in the HGSOC group (as shown in Figure 7A). This seemingly paradoxical behavior implies that the categorization of the HGSOC group into subgroups based on the degree of BIRC5 expression may guide the prognosis.

RRM2 encodes one of the two non-identical ribonucleotide reductase subunits. This reductase catalyzes the conversion of ribonucleotides to deoxyribonucleotides. RRM2 has been reported to be a key gene involved in various malignancies. Yang et al. recently revealed that RRM2 may prevent iron death in liver cancer and can be utilized as a tumor biomarker(34). In patients with lung adenocarcinoma, two investigations by Jin et al. and Ma et al. highlighted that the elevated RRM2 expression can be an independent predictor for poor prognosis(35, 36). Through bioinformatics study, Shen et al. discovered that RRM2 is substantially negatively associated with clinical staging in OC and that it represents an anti-cancer therapeutic target for patients with OC. The findings are similar to those of the current study(36). Figures 6 and 7B show that RRM2 is significantly expressed in OC tissues compared with normal ovarian tissues, which is an intriguing observation in our investigation. According to the HPA website analysis results (illustrated in Figure 9), RRM2 could not be identified in OC tissues through IHC, which might be due to an issue with the antibody itself, and more investigation in this regard is warranted.

Although TP53 is a reliable IHC index for differentiating HGSOC from LGSOC, none of the 164 DEGs examined in this study comprised the TP53 molecule, which necessitates the addition of novel differential diagnostic markers in further investigations. Furthermore, we confirmed the expression connection between TP53 and BIRC5/RRM2 in OC by using the GDC website. Compared with wild-type TP53, mutant TP53 is associated with increased expressions of BIRC5 and RRM2, with p values of 0.03 and 0.0063, respectively (Supplementary Figure 5). This finding implies that the expressions of BIRC5 and RRM2 in OC are congruent with that of TP53 and that the three IHC results may be complementary.

The present study has some limitation. First, despite the screening of BIRC5 as a possible HGSOC differential diagnostic and prognostic indicator, the experimental validation is lacking in this study.

Second, the selection criteria for DEGs in GSE14001 are only specified as  $\text{adj.p} < 0.05$ , which is utilized as a secondary verification for the summary results of DEGs in the other two datasets. Third, limited pathological specimens (only three normal ovarian tissues and 12 OC tissues; Figure 9) were included in the HPA database. The negative RRM2 IHC result in OC necessitates the inclusion of more pathological specimens for verification.

## **5. Conclusion**

BIRC5 is the key molecule identified in this investigation that can distinguish diagnosis, guide prognosis, and be applied in clinical settings. The importance of this study is to provide a novel marker to guide clinical practice. For the aforementioned limitations, we are currently gathering clinical data and pathology materials for a more in-depth experimental study on BIRC5, including the control of its upstream miRNA.

## **Abbreviations**

Abbreviation	Full name
HGSOC	High-grade serous ovarian cancer
LGSOC	Low-grade serous ovarian cancer
DEGs	Differentially expressed genes
DFS	Disease-free survival
PFS	Progress-free survival
GDC	Genomic Data Commons
OS	Overall survival
IHC	Immunohistochemistry
GEPIA	Gene Expression Profiling Interactive Analysis
GEO	Gene Expression Omnibus
GO	Gene ontology
KEGG	Kyoto Encyclopedia of Gene and Genome
WebGestalt	WEB-based Gene SeT AnaLysis Toolkit
GSEA	Gene Set Enrichment Analysis
STRING	Search Tool for the Retrieval of Interacting Genes
PPI	Protein-protein interactions
MCODE	Molecular Complex Detection
GTEx	Genotype-Tissue Expression
HPA	Human Protein Atlas
BP	Biological processes

## Declarations

### 7.1 Ethics approval and consent to participate

Not applicable.

### 7.2 Consent for publication

Not applicable.

### 7.3 Availability of data and materials

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

### 7.4 Competing interests

The authors declare that they have no competing interests.

### 7.5 Funding

The current study was supported by grants from the 2019 Kunshan Key R&D Plan (Ecological Agriculture and Social Development) -Social Development Science and Technology Project (KS1941).

### 7.6 Authors' contributions

Conceptualization: XS, GWJ; Database search: XJH, WQ, CF; Data analysis: XJH, GXJ; Original draft writing: XJH, WQ, GTT; Review, editing, and final approval: DZY, LTT; Research supervision: XS, GWJ, DZY.

### 7.7 Acknowledgements

Not applicable.

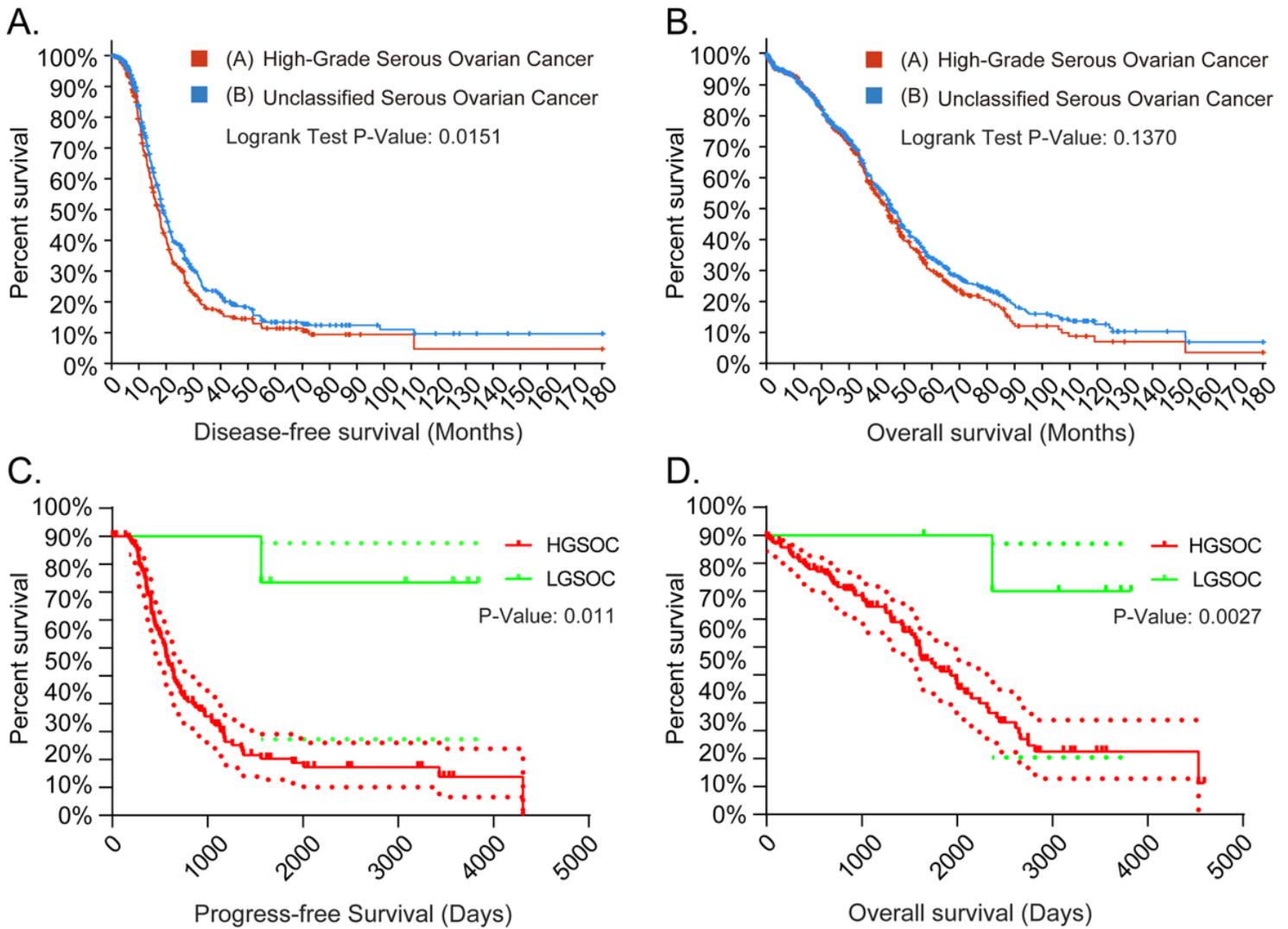
## References

1. Kaku T, Ogawa S, Kawano Y, Ohishi Y, Kobayashi H, Hirakawa T, et al. Histological classification of ovarian cancer. *Medical electron microscopy : official journal of the Clinical Electron Microscopy Society of Japan*. 2003;36(1):9-17.
2. Goulding EA, Simcock B, McLachlan J, van der Griend R, Sykes P. Low-grade serous ovarian carcinoma: A comprehensive literature review. *The Australian & New Zealand journal of obstetrics & gynaecology*. 2020;60(1):27-33.
3. Kaldawy A, Segev Y, Lavie O, Auslender R, Sopik V, Narod SA. Low-grade serous ovarian cancer: A review. *Gynecologic oncology*. 2016;143(2):433-8.
4. Ramalingam P. Morphologic, Immunophenotypic, and Molecular Features of Epithelial Ovarian Cancer. *Oncology (Williston Park, NY)*. 2016;30(2):166-76.
5. Kurman RJ. Origin and molecular pathogenesis of ovarian high-grade serous carcinoma. *Annals of oncology : official journal of the European Society for Medical Oncology*. 2013;24 Suppl 10:x16-21.
6. Surrey LF, Oakley FD, Merker JD, Long TA, Vasalos P, Moncur JT, et al. Next-Generation Sequencing (NGS) Methods Show Superior or Equivalent Performance to Non-NGS Methods on BRAF, EGFR, and KRAS Proficiency Testing Samples. *Archives of pathology & laboratory medicine*. 2019;143(8):980-4.
7. Köbel M, Piskorz AM, Lee S, Lui S, LePage C, Marass F, et al. Optimized p53 immunohistochemistry is an accurate predictor of TP53 mutation in ovarian carcinoma. *The journal of pathology Clinical*

- research. 2016;2(4):247-58.
8. 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 research*. 2013;41(Database issue):D991-5.
  9. . !!! INVALID CITATION !!! [8].
  10. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic acids research*. 2019;47(W1):W199-w205.
  11. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*. 2005;102(43):15545-50.
  12. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic acids research*. 2015;43(Database issue):D447-52.
  13. 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 research*. 2003;13(11):2498-504.
  14. Bandettini WP, Kellman P, Mancini C, Booker OJ, Vasu S, Leung SW, et al. MultiContrast Delayed Enhancement (MCODE) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: a clinical validation study. *Journal of cardiovascular magnetic resonance : official journal of the Society for Cardiovascular Magnetic Resonance*. 2012;14(1):83.
  15. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC systems biology*. 2014;8 Suppl 4(Suppl 4):S11.
  16. Hu J, Qiu D, Yu A, Hu J, Deng H, Li H, et al. YTHDF1 Is a Potential Pan-Cancer Biomarker for Prognosis and Immunotherapy. *Frontiers in oncology*. 2021;11:607224.
  17. The Genotype-Tissue Expression (GTEx) project. *Nature genetics*. 2013;45(6):580-5.
  18. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, et al. ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia (New York, NY)*. 2004;6(1):1-6.
  19. 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 research*. 2017;45(W1):W98-w102.
  20. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science signaling*. 2013;6(269):pl1.
  21. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. *Science (New York, NY)*. 2015;347(6220):1260419.
  22. Sticht C, De La Torre C, Parveen A, Gretz N. miRWalk: An online resource for prediction of microRNA binding sites. *PloS one*. 2018;13(10):e0206239.

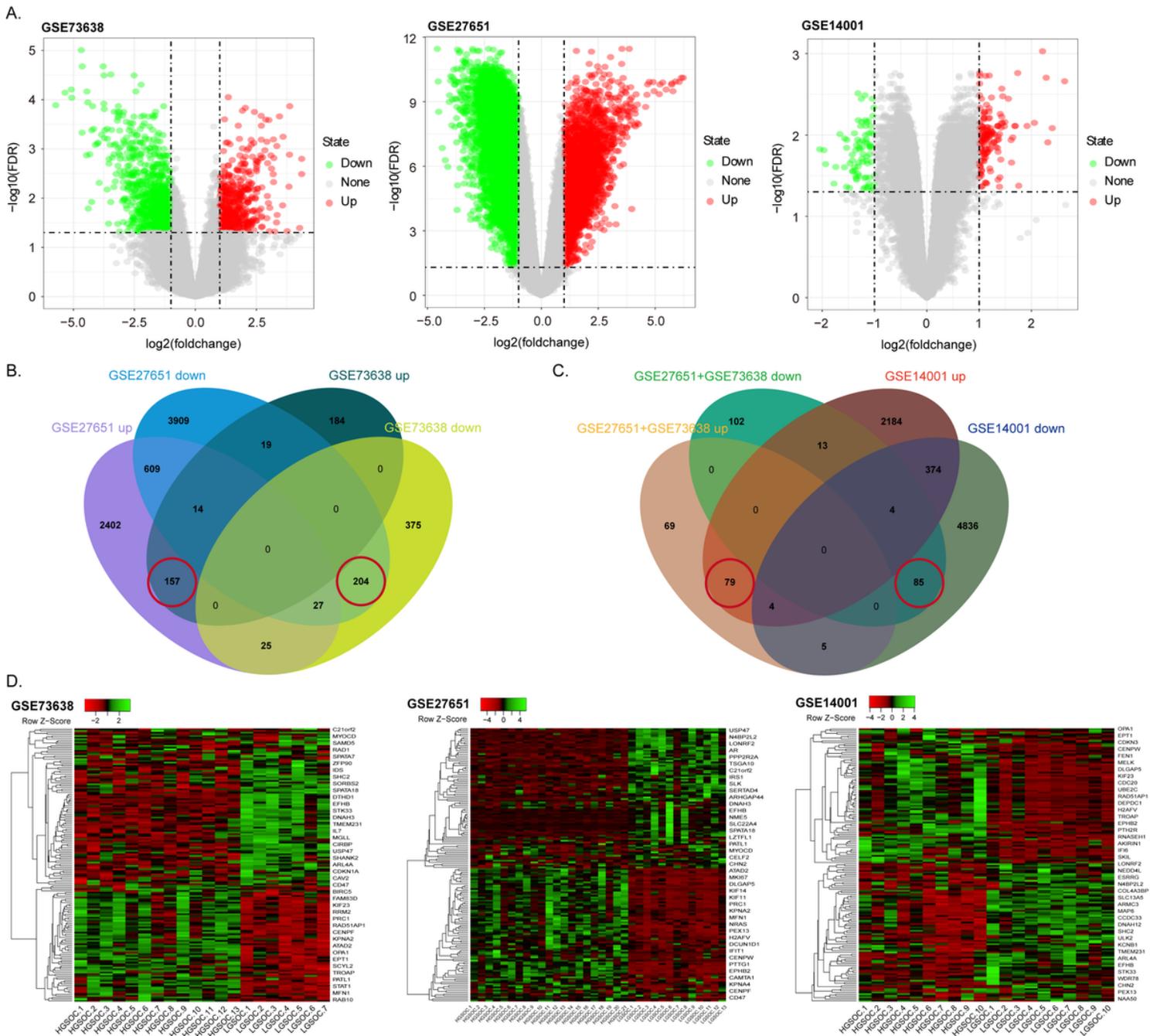
23. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. *eLife*. 2015;4.
24. Cunningham JM, Vierkant RA, Sellers TA, Phelan C, Rider DN, Liebow M, et al. Cell cycle genes and ovarian cancer susceptibility: a tagSNP analysis. *British journal of cancer*. 2009;101(8):1461-8.
25. Tanaka K, Hirota T. Chromosome segregation machinery and cancer. *Cancer science*. 2009;100(7):1158-65.
26. Chen M, Su J, Feng C, Liu Y, Zhao L, Tian Y. Chemokine CCL20 promotes the paclitaxel resistance of CD44(+)CD117(+) cells via the Notch1 signaling pathway in ovarian cancer. *Molecular medicine reports*. 2021;24(3).
27. Chiaramonte R, Colombo M, Bulfamante G, Falleni M, Tosi D, Garavelli S, et al. Notch pathway promotes ovarian cancer growth and migration via CXCR4/SDF1 $\alpha$  chemokine system. *The international journal of biochemistry & cell biology*. 2015;66:134-40.
28. Wheatley SP, Altieri DC. Survivin at a glance. *Journal of cell science*. 2019;132(7).
29. Yin S, Du J, Zhang J, Zhang X, Ma K. Identification of Key Genes and Pathway for Ovarian Neoplasms Using the OVDM1 Cell Line Based on Bioinformatics Analysis. *Medical science monitor : international medical journal of experimental and clinical research*. 2019;25:4305-13.
30. He X, Yang K, Wang H, Chen X, Wu H, Yao L, et al. Expression and clinical significance of survivin in ovarian cancer: A meta-analysis. *PloS one*. 2018;13(5):e0194463.
31. Li F, Aljahdali I, Ling X. Cancer therapeutics using survivin BIRC5 as a target: what can we do after over two decades of study? *Journal of experimental & clinical cancer research : CR*. 2019;38(1):368.
32. Ozretić P, Trnski D, Musani V, Maurac I, Kalafatić D, Orešković S, et al. Non-canonical Hedgehog signaling activation in ovarian borderline tumors and ovarian carcinomas. *International journal of oncology*. 2017;51(6):1869-77.
33. Wang B, Li X, Zhao G, Yan H, Dong P, Watari H, et al. miR-203 inhibits ovarian tumor metastasis by targeting BIRC5 and attenuating the TGF $\beta$  pathway. *Journal of experimental & clinical cancer research : CR*. 2018;37(1):235.
34. Yang Y, Lin J, Guo S, Xue X, Wang Y, Qiu S, et al. RRM2 protects against ferroptosis and is a tumor biomarker for liver cancer. *Cancer cell international*. 2020;20(1):587.
35. Jin CY, Du L, Nuerlan AH, Wang XL, Yang YW, Guo R. High expression of RRM2 as an independent predictive factor of poor prognosis in patients with lung adenocarcinoma. *Aging*. 2020;13(3):3518-35.
36. Ma C, Luo H, Cao J, Gao C, Fa X, Wang G. Independent prognostic implications of RRM2 in lung adenocarcinoma. *Journal of Cancer*. 2020;11(23):7009-22.

## Figures



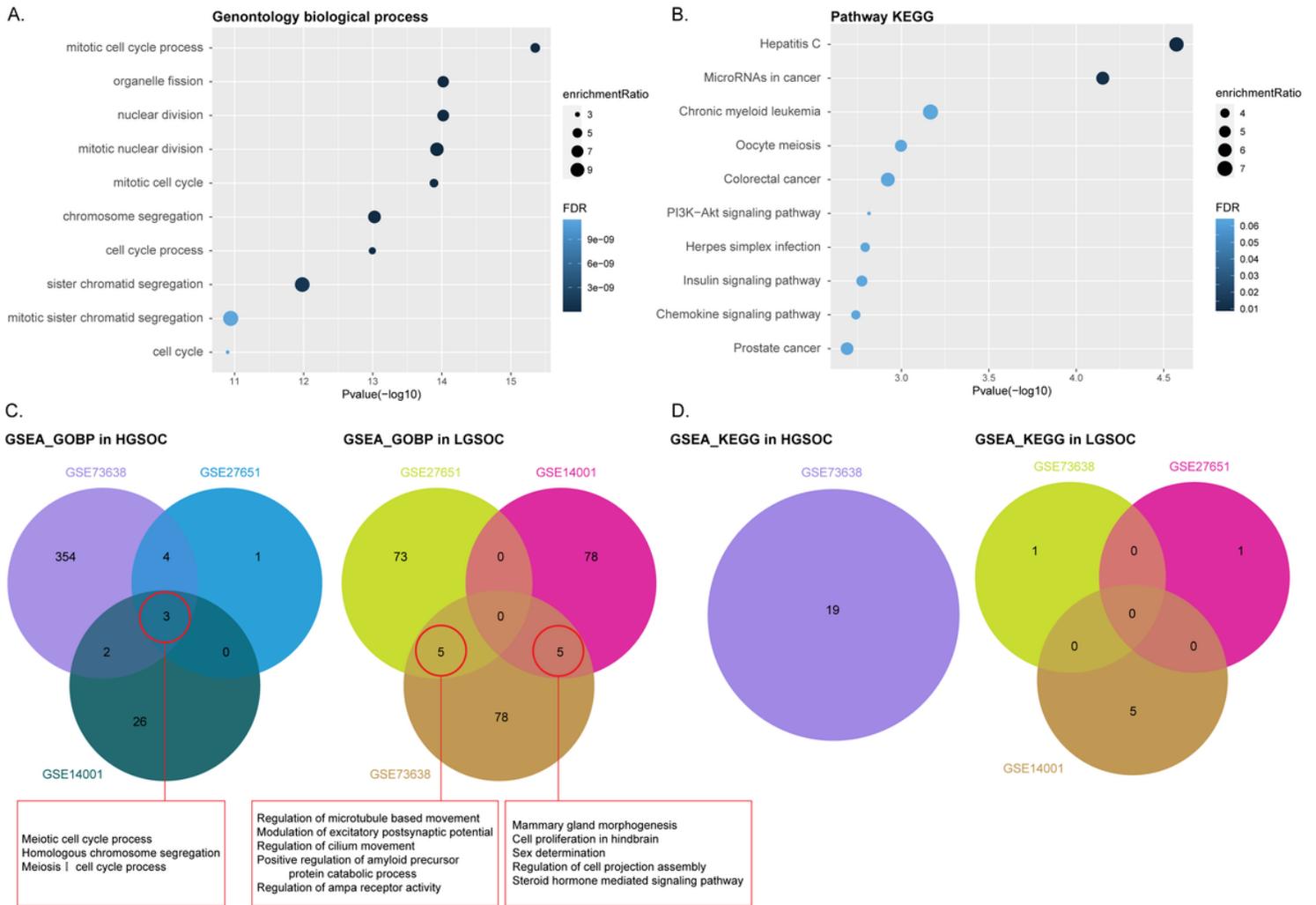
**Figure 1**

The prognostic information of the high-grade serous ovarian cancer. (A&B) Disease-free survival and overall survival between high-grade and unclassified serous ovarian cancer in GDC database. (C&D) Progress-free survival and overall survival between high-grade and low-grade serous ovarian cancer through GSE151335. Disease-free survival, progress-free survival, and overall survival are predicted by Kaplan-Meier analysis. A log-rank p-value < 0.05 was considered statistically significant. The median seq expression level was set as the cutoff for the KM plot. High-grade serous ovarian cancer sample number = 489, unclassified serous ovarian cancer = 1185 in GDC database. High-grade serous ovarian cancer sample number = 238, low-grade serous ovarian cancer = 6 in GSE151335. The red line indicates high-grade serous ovarian cancer, the blue line indicates unclassified serous ovarian cancer, and the green line indicates low-grade serous ovarian cancer.



**Figure 2**

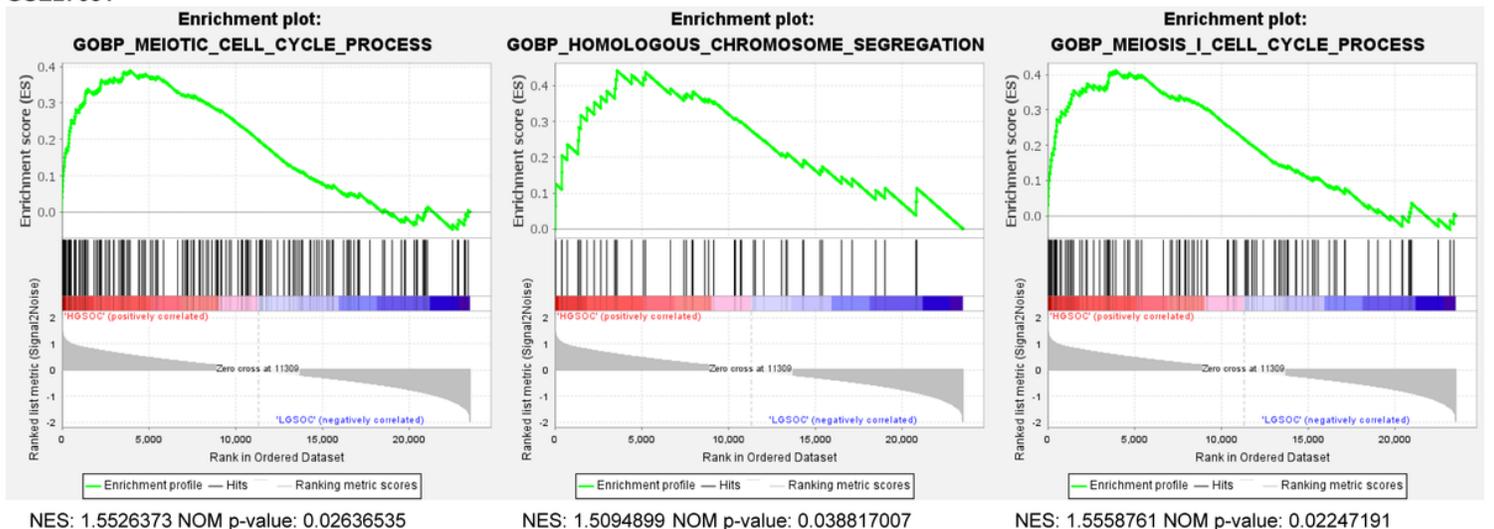
Screening of differential expressed genes between HGSOX and LGSOC. (A) Volcano plot of DEGs between HGSOX and LGSOC in each dataset. (B, C) Venn diagram of overlapping DEGs from GSE73638, GSE27651 and GSE14001. (D) Heatmap shows the distribution of DEGs data of each group extracted from GSE73638, GSE27651 and GSE14001.



**Figure 3**

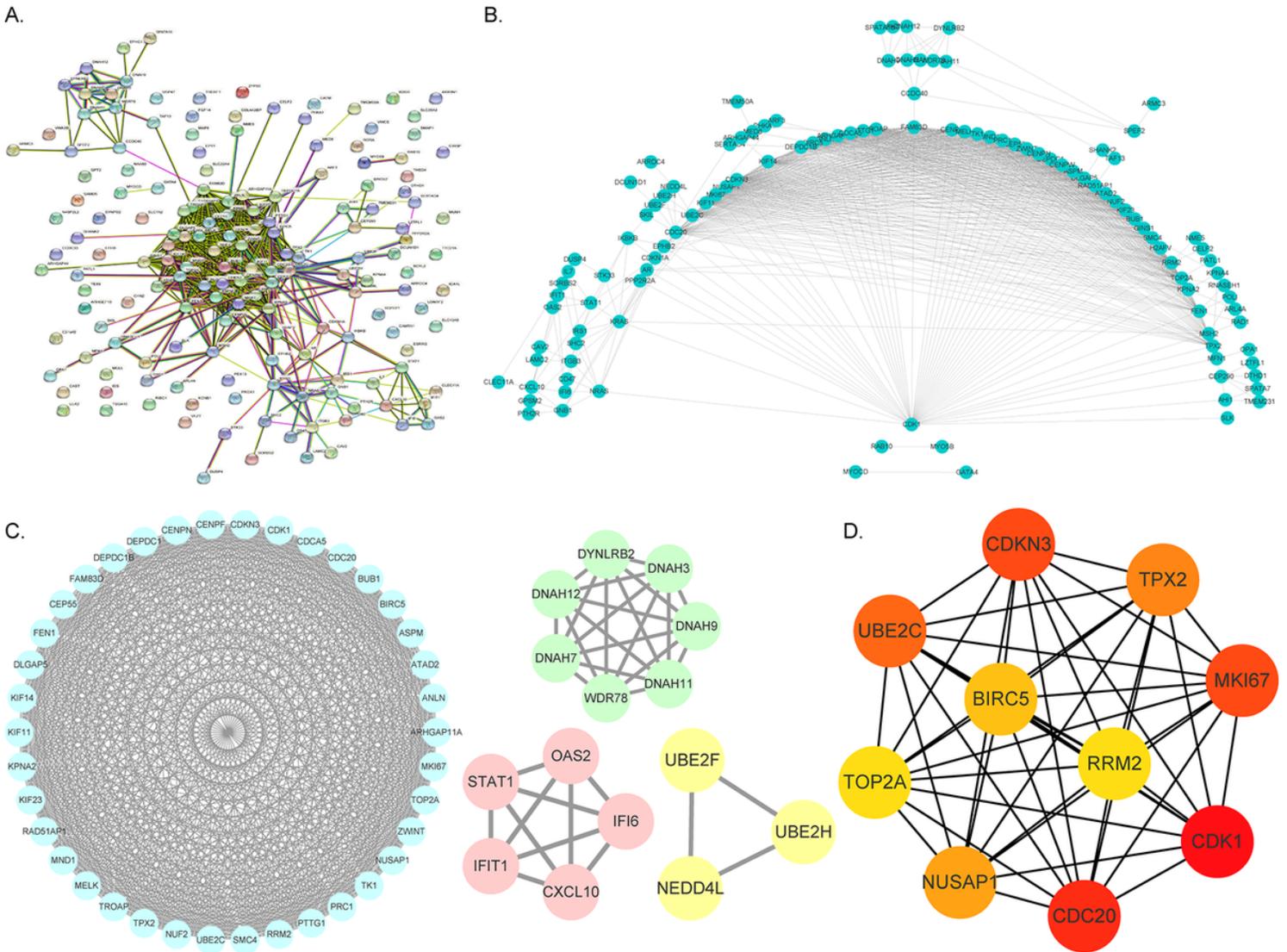
GO and KEGG analysis of the overlapping DEGs between HGSOC and LGSOC. (A) Top10 of GOBP enrichment via WebGestalt. (B) Top10 of KEGG pathway enrichment via WebGestalt. (C) Veen diagram of overlapping GOBP in HGSOC and LGSOC groups via GSEA. (D) Veen diagram of overlapping KEGG pathway in HGSOC and LGSOC groups via GSEA.

**GSE27651**



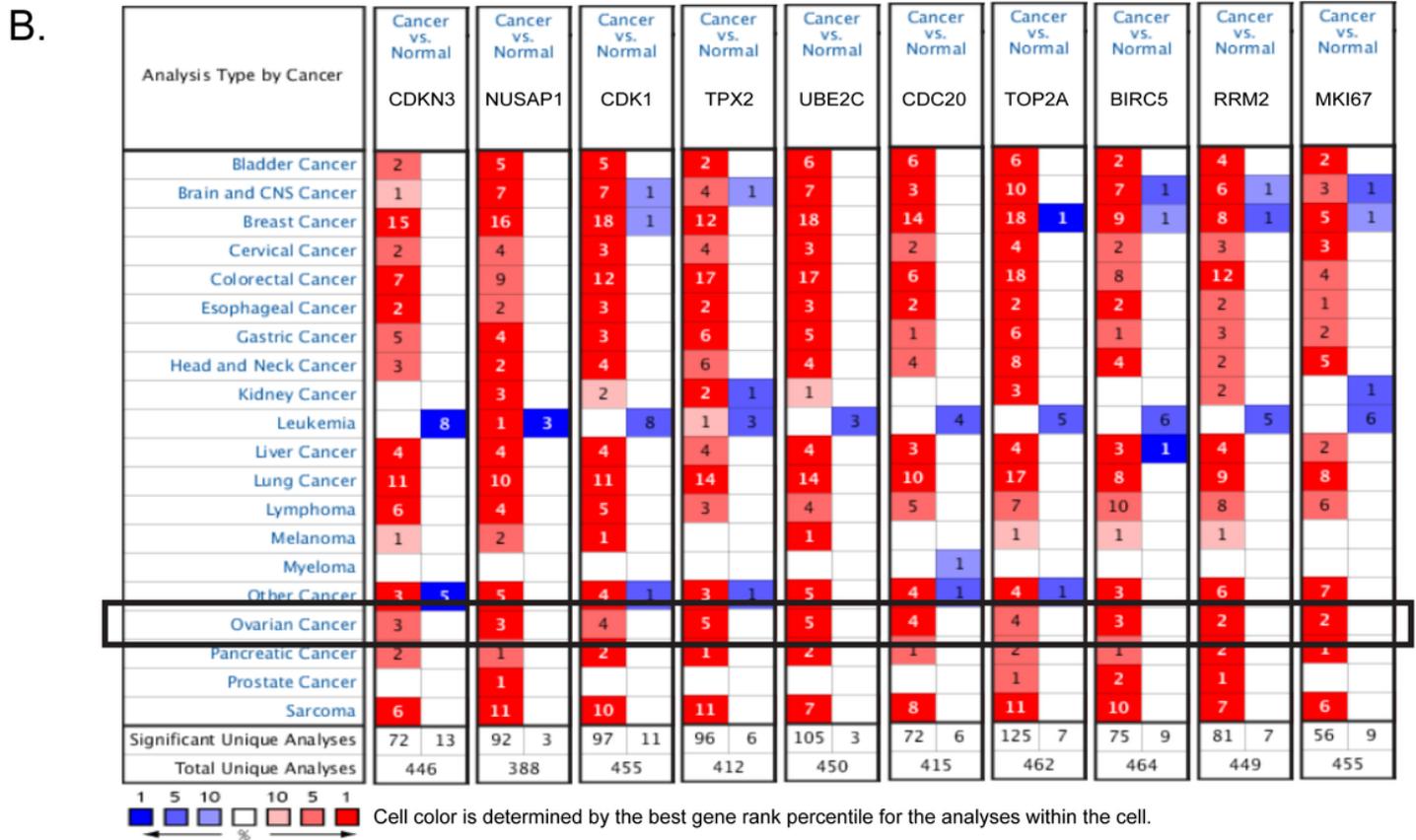
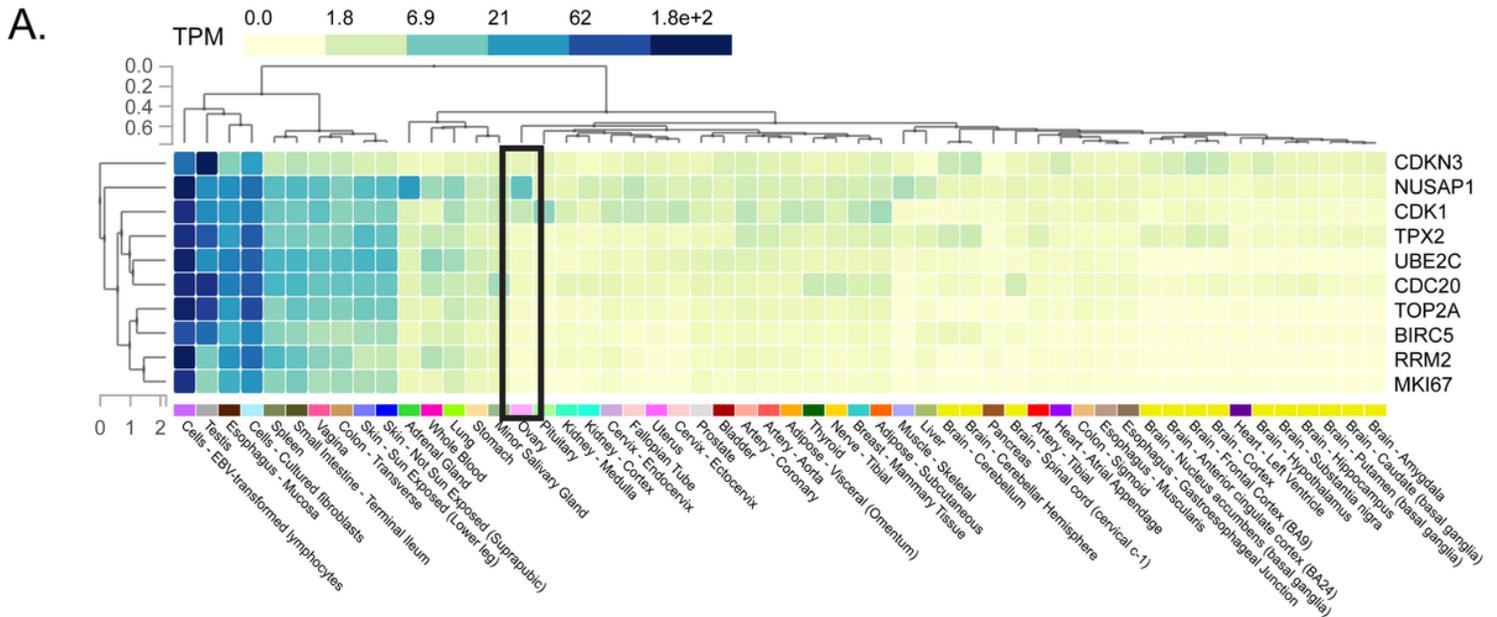
## Figure 4

The biological processes verified by the three GSEs. NES, normalized enrichment score; NOM p-value, Nominal p-value.



## Figure 5

DEGs PPI network construction and module analysis. (A) PPI network via STRING. (B) PPI network via Cytoscape. (C) Module analysis by MCODE: degree cutoff = 2, node score cutoff = 0.2, K-Core = 2, Max. Depth = 100. (D) Hub genes analysis by cytoHubba (top 10 nodes ranked by Radiality algorithms).



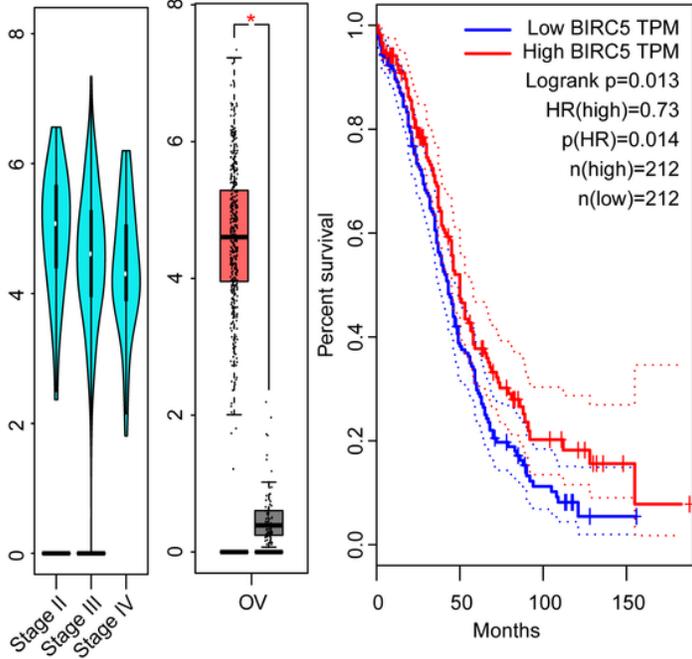
**Figure 6**

The transcription level of 10 hub genes in different normal tissues (A, GTEx) and cancers (B, OncoPrint).

### A. BIRC5

F value = 2.31  
Pr(>F) = 0.1

num(T)=426  
num(N)=88



### B. RRM2

F value = 3.72  
Pr(>F) = 0.0251

num(T)=426  
num(N)=88

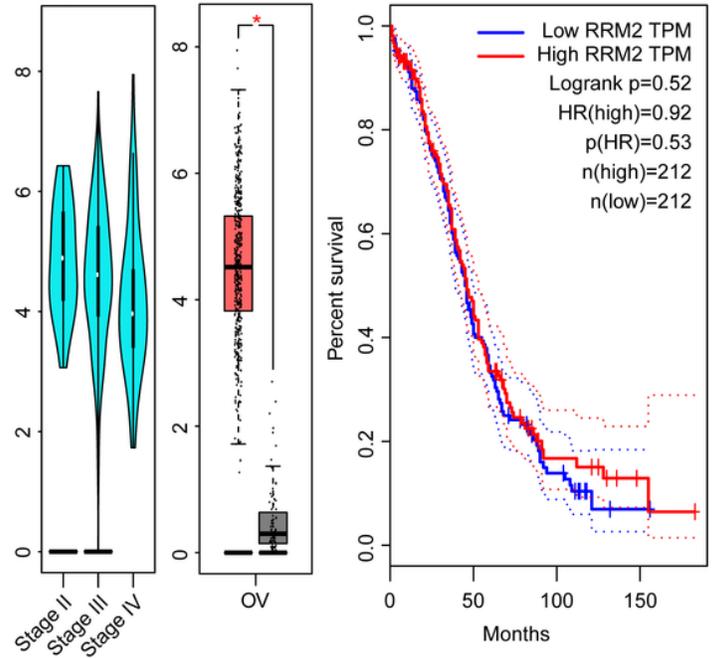
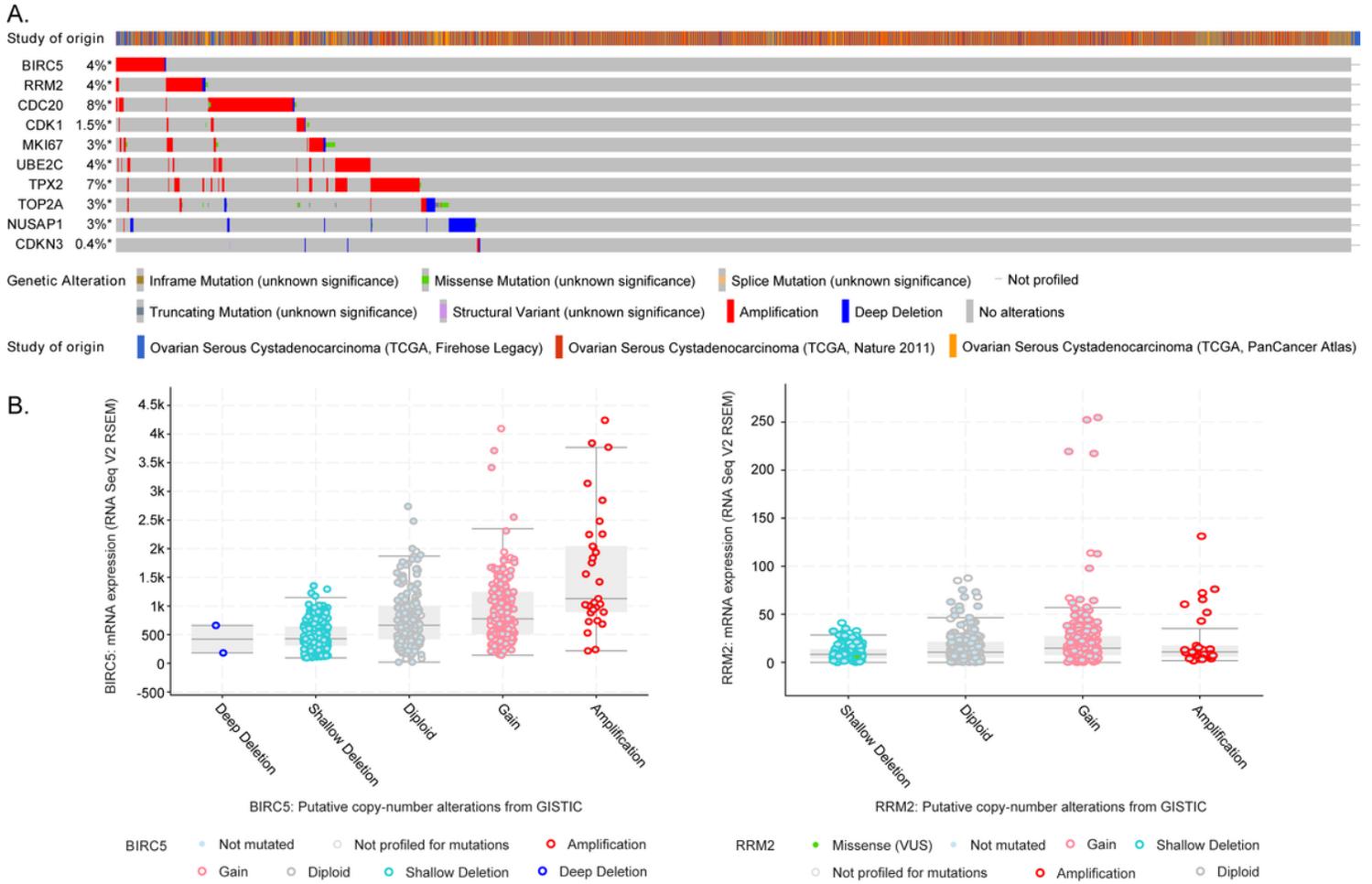


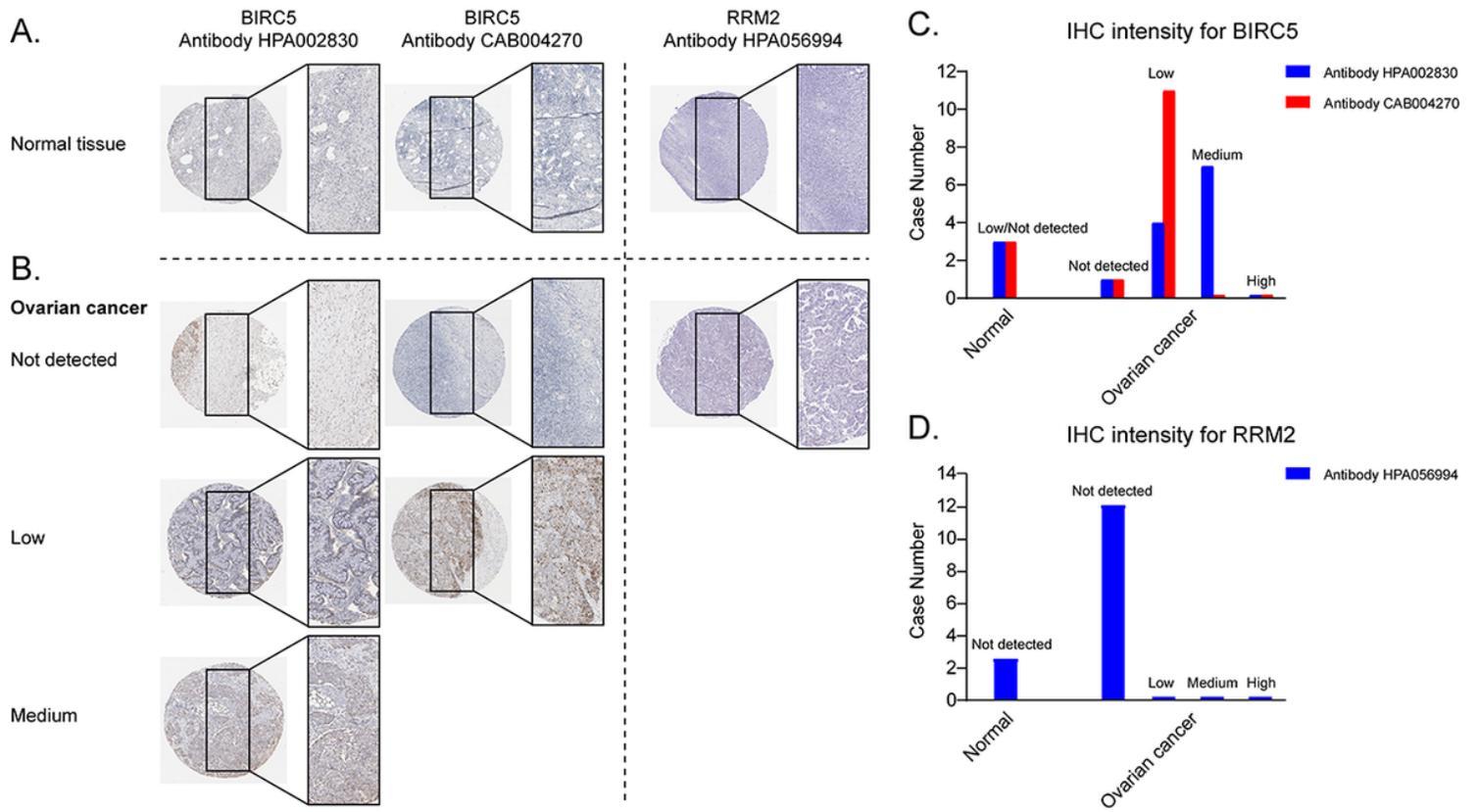
Figure 7

The stage, expression, and overall survival information of hub genes BIRC5 (A) and RRM2 (B) in ovarian cancer (GEPIA). Ovarian cancer sample number = 426 is in red, normal ovarian tissue number = 88 is in gray. Overall survival is predicted by Kaplan-Meier analysis. A log-rank p-value < 0.05 was considered statistically significant. The median seq expression level was set as the cutoff for the KM plot. Total patient number = 424. The red line indicates high expression, and the blue line indicates low expression. \*  $p < 0.05$ , statistically significant difference in expression. OV, ovarian cancer.



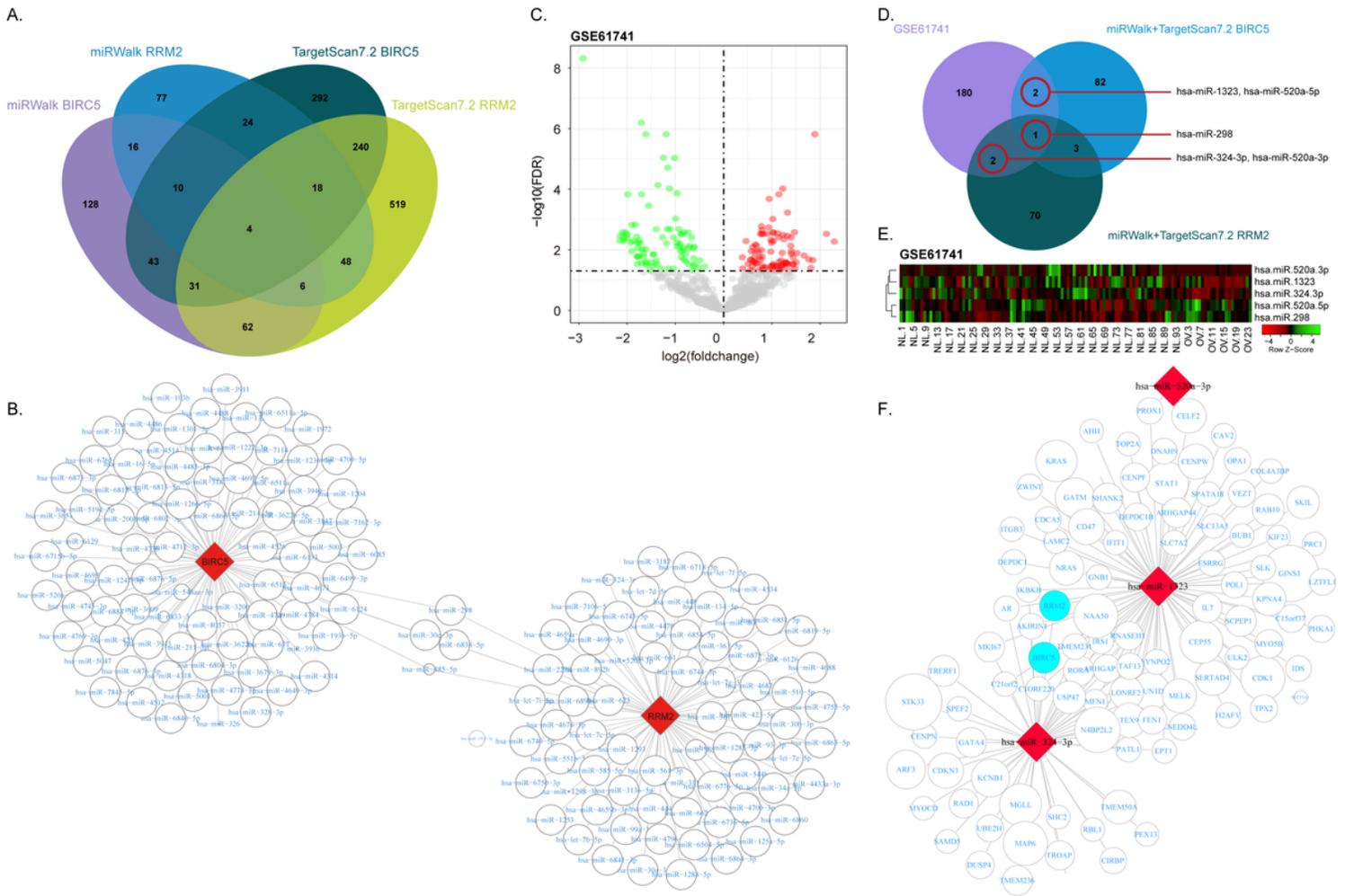
**Figure 8**

The mutation state of 10 hub genes in ovarian cancer (cBioportal). (A) Oncoprint. (B) Different mutation types of BIRC5 and RRM2 in ovarian cancer.



**Figure 9**

The IHC expression pattern of BIRC5 and RRM2 in normal ovarian tissue and ovarian cancer. (A) Normal ovarian tissue. (B) Representative IHC of ovarian cancer with not detected, low, and medium staining intensity. The black rectangle indicates a higher magnification of the indicated area in the picture. (C) Bar chart of IHC staining intensity of BIRC5 for ovarian cancer (total of 15 cases). (D) Bar chart of the IHC staining intensity of RRM2 for ovarian cancer (total of 15 cases).



**Figure 10**

The predicted upstream miRNAs of BIRC5 and RRM2 (miRWalk, TargetsCan7.2, and GSE61741). (A) Venn diagram of overlapping miRNAs of BIRC5 and RRM2 from miRWalk, TargetsCan7.2. (B) mRNA-miRNA map by Cytoscape. The size of the node reflects the strength of the relationship. mRNA of BIRC5 and RRM2 are in the red diamond box. (C) Volcano plot of differential expressed miRNAs between 24 OC blood samples and 94 normal blood samples in GSE61741. (D) Venn diagram of overlapping miRNAs of BIRC5 and RRM2 from miRWalk, TargetsCan7.2, and GSE61741. (E) Heatmap shows the distribution of differential expressed miRNAs data of each group extracted from GSE61741. (F) miRNA-mRNA map by Cytoscape. The size of the node reflects the strength of the relationship. hsa-miR-520a-3p, hsa-miR-1323, and hsa-miR-324-3p are in the red diamond box, while mRNA of BIRC5 and RRM2 are in the blue circle box.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFigure1.tif](#)

- [SupplementaryFigure2.tif](#)
- [SupplementaryFigure3.tif](#)
- [SupplementaryFigure4.tif](#)
- [SupplementaryFigure5.tif](#)
- [SupplementaryTable1.docx](#)
- [SupplementaryTable2.docx](#)