RESEARCH ARTICLE

Identification of Key Candidate Genes and Pathways for Relationship between Ovarian Cancer and Diabetes Mellitus Using Bioinformatical Analysis

Yi Sun[1], Huang Xiaoyan[1], Liu Yun[2], Liu Chaoqun[3], Wen Jialing[4], Yang Liu[1], Zhao Yingqi[1], Yi Peipei[1], Peng Junjun[1], Lu Yuanming[1]*

Abstract

  Ovarian cancer is one of the three major gynecologic cancers in the world. The aim of this study is to find the relationship between ovarian cancer and diabetes mellitus by using the genetic screening technique. By GEO database query and related online tools of analysis, we analyzed 185 cases of ovarian cancer and 10 control samples from GSE26712, and a total of 379 different genes were identified, including 104 up-regulated genes and 275 down-regulated genes. The up-regulated genes were mainly enriched in biological processes, including cell adhesion, transcription of nucleic acid and biosynthesis, and negative regulation of cell metabolism. The down-regulated genes were enriched in cell proliferation, migration, angiogenesis and macromolecular metabolism. Protein-protein interaction was analyzed by network diagram and module synthesis analysis. The top ten hub genes (CDC20, H2AFX, ENO1, ACTB, ISG15, KAT2B, HNRNPD, YWHAE, GJA1 and CAV1) were identified, which play important roles in critical signaling pathways that regulate the process of oxidation-reduction reaction and carboxylic acid metabolism. CTD analysis showed that the hub genes were involved in 1,128 distinct diseases (bonferroni-corrected P<0.05). Further analysis by drawing the Kaplan-Meier survival curve indicated that CDC20 and ISG15 were statistically significant (P<0.05). In conclusion, glycometabolism was related to ovarian cancer and genes and proteins in glycometabolism could serve as potential targets in ovarian cancer treatment.

Keywords: Bioinformatics analysis- microarray- ovarian cancer- diabetes mellitus

Introduction

  Ovarian cancer (OV) is one of the most common malignant tumors of the female genital organ. The morbidity of OV ranks second after the cervical cancer and uterine body carcinoma, and moreover, the third in the list of malignance. Globally about 200,000 people are diagnosed with ovarian cancer and 125,000 die from the disease each year (Torreet al., 2012). As the symptoms of ovarian cancer is vague and the lack of effective screening techniques and early diagnosis methods, 70%~80% of patients are diagnosed with advanced disease, which contributes to poor treatment effect that 5-year survival rate is less than 20%(Chou et al., 2010; Ledermann et al., 2016). From a medical point of view, the embryo development of ovary, the histological anatomy and the endocrine function are rather complex, and the early symptoms are not typical, as a result it is difficult to identify the origin and malignancy degree before surgery.

  Epidemiological studies have demonstrated that the risk of OV is associated with obesity, diabetes, and metabolic syndrome. More evidences show that persons with type I diabetes mellitus (DM) have an approximately 20-25% higher cancer incidence compared with persons without diabetes. For instance, women with diabetes have been shown to have an elevated incidence of liver, pancreatic, kidney, and endometrial cancer.

  Recently a meta-analysis of epidemiological studies on the correlation between diabetes and OV indicated that diabetes was associated with an increased risk of OV. The incidence of OV risks related approximately to 1.17 (95% confidences interval (CI), 1.02-1.33)](Donal and Lamkin, 2009). Besides, a variety of proteins in glycometabolism have been attracting increasing attention as potential targets for OV treatment (Birrer et al., 2011). The clinical results are still controversial though, it suggests that

*[1]Department of Toxicology, Guilin Medical University School of Public Health, Guilin, [2]Department of Gynaecology and Obstetrics, Kailuan General Hospital, Tangshan, Hebei, [3]Department of Nutrition, School of Medicine, Jinan University, Guangzhou, Guangdong, [4]Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Diseases, The Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou, China. *For Correspondence: bobojpn2002@163.com. Yi Sun, Xiaoyan Huang and Yun Liu have equal contribution in this study.*

targeting glycometabolism could be a novel approach to treat OV.

High throughput platform for genetic analysis is becoming more and more important, and the proper application of microarray is one of the most promising technology in medical oncology. Gene microarray technology in cancer gene research, cancer gene expression profile analysis, molecular classification of the tumor and tumor gene mutations is widely used due to the great advantage in tumor biology. For the past few years, gene microarray technology has been used in the study of malignant OV to analyze the expression difference of genes, detect the molecular markers of early diagnosis, analyze the characteristics of drug resistance genes, judge the sensitivity to chemothearpy, analyze characteristics of gene expression and predict the early tendency of recurrence. All of these are the critical methods to study the gene expression profile of OV.

In our research, the initial data were downloaded from Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo). GSE26712 genetic data from the patients with OV were compared with normal controls, then GO and pathway enrichment analysis was performed. We want to integrate the publicly available microarray datasets in the database, analyze the biological functions of genes and pathways, explore whether diabetes is the risk factor of OV, which provides meaningful clues for the relationship of diabetes and OV.

Materials and Methods

Microarray data information and DEGs identification

GSE26712 gene expression profiles were downloaded from the GEO database. GSE26712 was submitted by Michael Birrer based on the agilent GPL96 platform ([hg-u133a] Affymetrix Human Genome U133A Array). The database consisted of 195 samples in total, including185 tissues from ovarian cancer patients and 10 from normal human. We downloaded the raw data in text form (agilent platform). After analyzing the top 250 genes with GEO online tools 2R, we got all genes with differences expression. And the genetic variations were extracted in Excel spreadsheets, with the conditions of $|\log FC| \geq 2$, P-value< 0.05. Finally 379 eligible difference genes were found. The results showed that the data was statistically significant.

Genetic ontology and pathway enrichment analysis of differential genes

A variety of online databases were utilized to analyze candidate DEGs function and access enrichment. GO constitutes one of a variety of biological ontology languages, providing a three-tier system definition method for describing the function of gene products, which is suitable for all kinds of species, gene and protein function for qualified and description]Gene Ontology C 2006). KEGG (http://www.kegg.jp/) is a database to know the advanced features and biological systems (such as cells, biological and ecological systems) from the molecular information, especially large data sets generated genome sequencing and other high-throughput experimental technologies (Ashburner et al., 2000). Now David (https://david.ncifcrf.gov/) provides researchers with a set of complete functional annotation tools for investigators to understand the biological significance behind the numerous genes (Kanehisa and Goto, 2000). GO analysis and KEGG analysis of the selected different genes were performed using David online tools. The biological process, molecular mechanism and the cell composition of the GO term were obtained by the annotation of the altered genes, and the correlation pathway of KEGG was analyzed. The differential genes of David online tool analysis (P-value < 0.05) was considered to be statistically significant.

The network diagram and module synthesis analysis of protein-protein interaction

Search Tool for the Retrieval of Interaction Genes-STRING- is an online tool used to identify protein-protein interaction information.

The STRING (version10.5) contains 9,643,763 proteins from 2031 species; 1,380,838,440 interactions. To evaluate the interaction of different genes, we mapped all the differences of all the genes into a STRING and used score > 0.4 as the cut-off criteria. Then, the PPI network diagram was constructed by using the cytoscape, and the three modules of the expression were selected by the plug-in MCODE of cytoscape, and the module analysis was carried out. The corresponding proteins in the central nodes might be the core proteins and key candidate genes with significant physiological regulatory functions.

2.4 Disease predictions of candidate genes

The Comparative Toxicogenomics Database CTD, http://ctdbase.org./ was used for describing the relationship between chemicals, genes and diseases. We put ten central genes into CTD website, and the CTD tools was used to express the abnormal expression of OV in these ten hub genes and to analyze the relationship between metabolic and cancer related diseases.

Analysis of ROC in candidate gene ovarian cancer

Receiver operating characteristic (ROC) curves for the hub genes were analyzed using web-based tools (http://www.proteinatlas.org). Oncolnc is a website that incorporates various RNA data and patient clinical data from TCGA and provides survival analysis. We mapped ten central genes into this site, the use of the functions and characteristics of the Oncolnc hub genes in OV survival situation is analyzed, and the survival curves of the ten central genes, is used to evaluate 10 hub genes diagnosis effectiveness of abnormal expression in OV.

Results

Identification of differential genes in ovarian cancer

A total of 185 cases of ovarian cancer and 10 control samples were analyzed. The series of microarray data were analyzed with GEO 2R tool, and the different genes were obtained with P-value < 0.05 and $|\log FC| \geq 2.0$ as the cut-off criterion. In the analysis of GSE26712, a total of 379 DEGs were identified, including 104 up-regulated genes and 275 down-regulated genes (Table 1). The heat

Table 1. Identification of the Up-Regulated and Down-Regulated Differential Genes

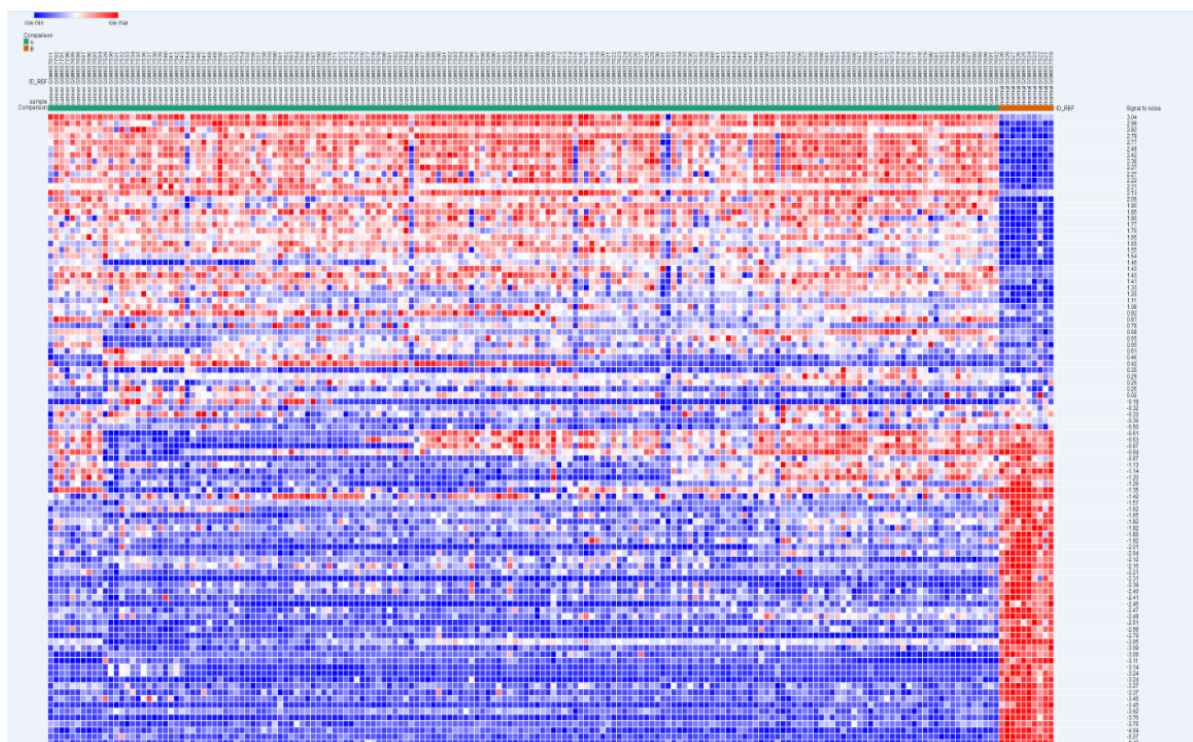| DEGs	Genes Name
Up-regulated	CD24 CD24 SNORA68 RPL18A CD24 CLDN3 IGFBP2 LOC101929219 LOC100505650 C1orf186 HIST2H2AA4 HIST2H2AA3 CALR WFDC2 SCGB2A1 FOLR1 CD24 PFN1 CRABP2
	PKM COL1A1 RPL37ASOX17 RPS7 CD24 PPDPF MUC1 MECOM C9orf16 NOTCH3 CLDN3 ACTB RHOB S100A13 ELF3 LCN2 SLC52A2 IGLC1 KLK6 S100A11 DEFB1
	INS-IGF2 IGF2 HMGA1 ISG15 FOLR1 GSTP1 TAGLN S100A2 H2BFS IGLC1 CALR COL1A1 CKS2 CP IGHA2 IGHA1 IGH ATP5H H2AFX YWHAE PEA15 ADGRG1 PRMT1
	S100A6 LSM4 SOX9 MMP11 SETSIP SETP4 SET CYAT1 IGLV1-44 IGLC1 PRKCSH MIR7113 MIR4691 NDUFS8 CL2L2-PABPN1 PABPN1 C9orf16 GRINA LOC100506248
	MIR1244-2 MIR1244-3 MIR1244-1 LOC728026 PTMA MMP7 ACTB NT5DC2 BCAM COL6A2 FXYD3 JUNB MUC1 PAX8 UBE2C CLDN4 LYPD1 UQCRQ HSPG2 MRPS12
	ZNF593 CDC20 IGLV1-44 C1orf106 IDH2 IFI6 ENO1 COX7B SLC39A4 IFI27 KLK8 KDELR1 MDK HIST1H1C UCP2 MIR1282 HYPK SERF2 TMED2 PRSS8 ATP2A2 BCAM
	CHMP5 PAXBP1 EZR MEF2C SCP2 RARRES1 PLIN2 HSD17B11 PRNP DCUN1D4 HAS1 DLG5 EPS15 PNRC2 PDZD2 ATF1 ADGRG6 CPE BAMBI PRKAR1A GJA1 NR2F2
	RSRP1 PLCB1 CHMP2B VPS13C TJP2 ADH5 NEFH CAV2 ANK3 SMARCA2 CHGB ZNF45 RARRES1 RHOT1 LIPA EFHC1 ATMIN PURA GLS WT1 SFRP1 C21orf62
	LOC101929500 CRIM1 KAT2B BICC1 SERPINB9 ADAMTS3 PRKAR2B GALNT12 PDPN HNRNPD GNG11 TOB1 SEL1L3 DCN GPR137B ISOC1 MAP3K8 TSPYL1 SNX7 UPK3B
	CIRBP PREPL PTPRC NT5C1B-RDH14 RDH14 CLK1 PODXL SNORD45C SNORD45A SNORD45B RABGGTB HPR HP CELF2 ANOS1 LGALS2 PCDH9 PKD2 ARAP2 CELF2
	CAV2 KLF4 PCOLCE2 PLCE1 LUC7L3 ARHGAP44 LAMB1 FZD7 SLC16A1 PRKAR1A ARHGAP6 GFPT2 DAPK1 FZD7 CLDN15 S100PBP ACSL1 ZDHHC17 DSC3 ADAM9
	FLRT3 SEMA5A ITGAM CFH PGRMC2 SOBP GABARAPL3 GABARAPL1 PSD3 TSPAN8 ATP10D ME1 ZBTB20 IGFBP6 RBM25 WSB1 RAP2C TCF21 PTGER3 GPRASP1
	CSGALNACT1 SLK GCOM1 POLR2M ATP8A1 GALC GCA TSPAN13 ATXN1 CFHR1 CFH DOCK4MIR22 MIR22HG OGN UFSP2 TLE4 RECK SARAF FEZ2 SMARCA2
Down-regulated	TPD52L1 PDGFC PDE8B GSAP GSAP HBA2 HBA1 FGF9 PHACTR2 PNISR CREBL2 CPVL SLC46A3 OLFML1 HBG2 HBG1 MTUS1 CLEC4M TRPC1 LXN CAST CLEC4M HBA2
	BA1 GHR CPE SEMA3C SULF1 HBA2 HBA1 DPYSL2 TACC1 SLC16A1 NSG1 METTL7A ECM2 QKI HBA2 HBA1 NKX3-1 CFI CLIC5 PTGIS LOC101926921 DAB2 RARRES1
	DMD HOXC6 WNT2B ZNF330 PTGER4 NR2F1 GREB1 SNCAIP MARCO GAS1 LAMA4 FRMD4B AQP9 IL6ST BTAF1 HP SCG5 ZFPM2 FGF13 FAM13B RSRP1 DIRAS3
	ALDH1A1 WNT5A TMX4 AMIGO2 ANXA3 CREBL2 ID4 N4BP2L2 GATA6 RGS4 CAV1 HBA2 HBA1 DDX17 LOC101926921 DAB2 TFPI2 PLPP1 RNASE4 RTN1 PRG4 HBB
	REEP1 PDGFD MEIS2 TFPI2 DSE FRY KDR HEG1 HBB PEG3 SFRP1 MAF SPOCK1 MET LGALS8 NAP1L3 PSD3 TMEM255A DFNA5 PROS1 LGALS8 PLSCR4 ANXA8L1
	ANXA8 ARMCX1 PROCR ADH1B HBA2 HBA1 CHRDL1 SLC39A8 ADH1B SLC4A4NPY1R NELL2 DPYD BNC1 LHX2 GPM6A EFEMP1 CALB2 REEP1 STK26 OC100506718
	FLRT2 TCEAL2 HB ABCA8 EFEMP1 AOX1 LOC101928635 ALDH1A2 BCHE MNDA LOC101930363 LOC101928349 LOC100507387 FAM153C FAM153A FAM153B

Figure 1. Heat Map of the Top Differentially Expressed Genes (50 up- and 50 down-regulated genes). Red, up-regulation; Blue, down-regulation

Table 2.Gene Ontology Analysis of Differentially Expressed Genes Associated with Ovarian Cancer3.3 Signal Path Enrichment Analyses

Expression	Category	Term	Count %	P- Value
Up-reulated				
GOTERM_BP_FAT	GO:0007399~nervous system development	25	27.47252747	5.33E-05
GOTERM_BP_FAT	GO:0007155~cell adhesion	21	23.07692308	1.16E-04
GOTERM_BP_FAT	GO:0022610~biological adhesion	21	23.07692308	1.22E-04
GOTERM_BP_FAT	GO:1901700~response to oxygen-containing compound	19	20.87912088	1.64E-04
GOTERM_BP_FAT	GO:0033993~response to lipid	14	15.38461538	2.15E-04
GOTERM_BP_FAT	GO:0045892~negative regulation of transcription, DNA-templated	16	17.58241758	2.77E-04
GOTERM_BP_FAT	GO:0010558~negative regulation of macromolecule biosynthetic process	18	19.78021978	3.42E-04
GOTERM_BP_FAT	GO:1903507~negative regulation of nucleic acid-templated transcription	16	17.58241758	4.28E-04
GOTERM_BP_FAT	GO:0010605~negative regulation of macromolecule metabolic process	24	26.37362637	4.34E-04
GOTERM_BP_FAT	GO:0031324~negative regulation of cellular metabolic process	24	26.37362637	4.88E-04
GOTERM_BP_FAT	GO:0009890~negative regulation of biosynthetic process	18	19.78021978	6.74E-04
GOTERM_BP_FAT	GO:0010033~response to organic substance	26	28.57142857	0.0011882
GOTERM_BP_FAT	GO:0009892~negative regulation of metabolic process	24	26.37362637	0.001356973
GOTERM_BP_FAT	GO:0023051~regulation of signaling	27	29.67032967	0.00178692
GOTERM_BP_FAT	GO:0010604~positive regulation of macromolecule metabolic process	25	27.47252747	0.003192651
Down-regulated				
GOTERM_BP_FAT	GO:0040011~locomotion	48	21.33333333	5.62E-09
GOTERM_BP_FAT	GO:0008283~cell proliferation	53	23.55555556	1.77E-08
GOTERM_BP_FAT	GO:0006928~movement of cell or subcellular component	51	22.66666667	4.02E-08
GOTERM_BP_FAT	GO:0016477~cell migration	38	16.88888889	2.14E-07
GOTERM_BP_FAT	GO:0032268~regulation of cellular protein metabolic process	56	24.88888889	2.11E-06
GOTERM_BP_FAT	GO:0007399~nervous system development	52	23.11111111	5.25E-06
GOTERM_BP_FAT	GO:0010604~positive regulation of macromolecule metabolic process	62	27.55555556	1.04E-05
GOTERM_BP_FAT	GO:0010605~negative regulation of macromolecule metabolic process	53	23.55555556	1.72E-05
GOTERM_BP_FAT	GO:0009605~response to external stimulus	48	21.33333333	3.47E-05
GOTERM_BP_FAT	GO:0050790~regulation of catalytic activity	52	23.11111111	5.13E-05
GOTERM_BP_FAT	GO:0048514~blood vessel morphogenesis	19	8.444444444	4.92E-05
GOTERM_BP_FAT	GO:0072358~cardiovascular system development	33	14.66666667	2.77E-07
GOTERM_BP_FAT	GO:0008285~negative regulation of cell proliferation	25	11.11111111	2.34E-06
GOTERM_BP_FAT	GO:0060537~muscle tissue development	18	8	4.03E-06
GOTERM_BP_FAT	GO:0007155~cell adhesion	43	19.11111111	1.33E-05

map of differential gene expression (the genes of the first 50 and down-regulated genes) was presented in Figure 1.

The ontology analysis of differential genes in ovarian cancer

The results of GO analysis indicated that the different expressions of the top 20 DEGs were mainly associated with cell migration, proliferation and adhesion. The identified DEGs also participated in the metabolism process of cells and polymer, the response to lipid, and organ and tissue development. The up-regulated genes

Table 3. Significantly Enriched Pathway Terms of DEGs in Ovarian Cancer

Significant Enriched Pathway terms		P-Value
PI3K-Akt signaling pathway	|HAS-04151	0.037083032
Wnt signaling pathway	|HAS-04310	0.035010312
Ras signaling pathway	|HAS-04014	0.02893891
Rap1 signaling pathway	|HAS-04015	0.001684852

were mainly concentrated in biological processes, including cell adhesion, transcription of nucleic acid and biosynthesis, and negative regulation of cell metabolism. The down-regulated genes were predominantly concentrated in cell proliferation, migration, angiogenesis, and macromolecular metabolism. The molecular function enrichment analysis showed that the up-regulated genes were mainly focused on enzyme binding, RNA binding and protein dimerization activity, and the down-regulated genes were predominantly concentrated in receptor binding, molecular function regulatory mechanism, enzyme binding and enzyme regulation activity. In addition, the analysis of GO cell components indicated that the increase of genes is mainly concentrated in remote body, extracellular region and membrane binding domain. And the down-regulated genes were predominantly concentrated in extracellular regions, membrane-bounded vesicls, extracellular exosomes and cell joints (Figure 2 and Table 2). The signaling pathways and functions identified by KEGG PATHWAY enrichment analysis were

Table 4. KEGG Pathway Analysis of Differentially Expressed Genes Associated with Ovarian Cancer

Pathway	Name	Gene	%	P Value	Genes ID
Up-reulated	DEGs				
hsa01230:Biosynthesis of amino acid		3	3.296703297	0.074028877	IDH2, ENO1, PKM
hsa05010:Alzheimer's disease		4	4.395604396		ATP2A2, ATP5H , UQCRQ , COX7B
hsa04512:ECM-receptor interaction		3	3.296703297	0.097658172	COL1A1, COL1A1, HSPG2, COL6A2
Down-reulated	DEGS				
hsa05205:Proteoglycans in cancer		11	4.888888889	2.75E-04	WNT5A, CAV2, CAV1, PLCE1, EZR, ANK3, MET, DCN, FZD7, KDR, WNT2B
hsa04015:Rap1 signaling pathway		10	4.444444444	0.001684852	PLCE1, FGF9, MET, FGF13, PDGFC, PDGFD, PLCB1, ITGAM, KDR, DOCK4
hsa05200:Pathways in cancer		14	6.222222222	0.001775521	WNT5A, PTGER3, PTGER4, FGF9, MET, GNG11, FGF13, FZD7, DAPK1, WNT2B,LAMA4, KX3-1, LAMB1, PLCB1

Rap1, Ras, Wnt, and PI3K-AKT, as well as the metabolism of retinol and amino acids. The up-regulated genes were mainly enriched in the biosynthesis of amino acid and the ECM receptor interactions, while the down-regulated genes were mainly enriched in cancer-related signaling pathways, especially Rap1 signaling pathway (Figure 3 and Table 3-4). According to the information in STRING database and Cytoscape software, there were 379 DEGs (104 up- and 275 down-regulated) in the total of 4,271 genes, and the PPI network was constructed with 218 nodes
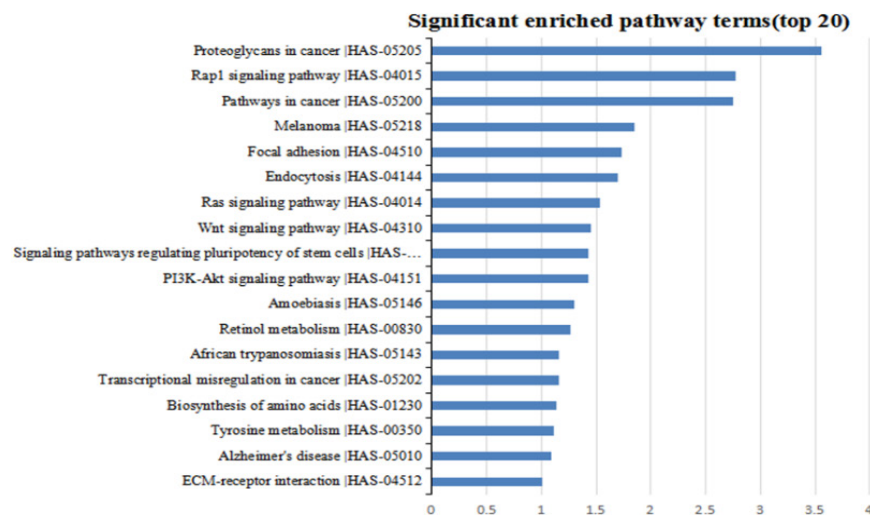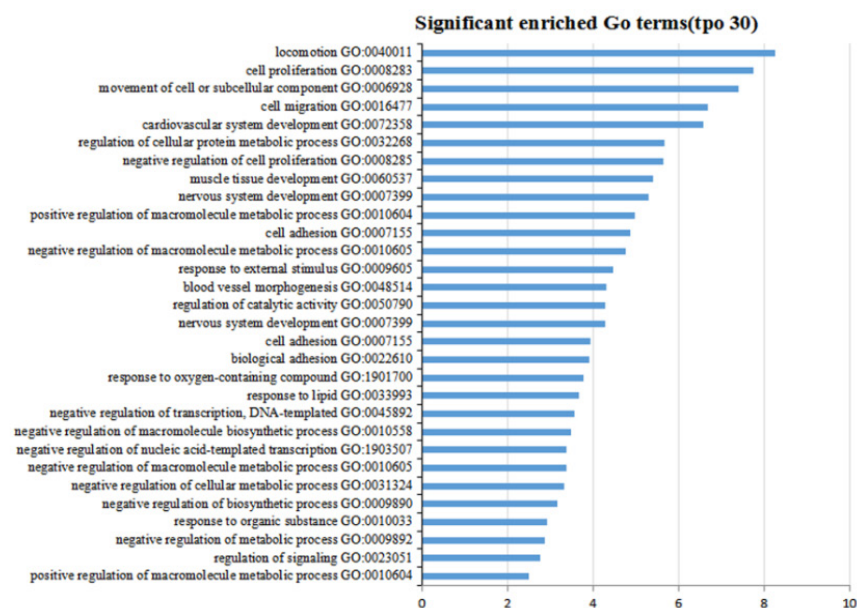

Figure 2. Gene Ontology Anlysis; A, Significantly enriched GO terms of DEGs in ovarian cancer; B, Significantly enriched pathway terms of DEGs in ovarian cancer

Table 5. Pathway Enrichment Analysis of Module Agenes Function

Gene Set	FDR	Node
positive regulation of biological process	0.00364	ACTB,ALDH1A1,BUB1,CALR,CCNB1,CDC16,CDC23,CDC27,CDK1,DD X17,H2AFX,HNRNPD,ISG15,KAT2B,MAD2L1,PFN1,YWHAE ALDH1A1,BUB1,BUB1B,CALR,CDC16,CDC20,CDC23,CDC27,CDK1,D
regulation of metabolic process	0.016	DX17,ENO1,EZR,GJA1,H2AFX,ISG15,KAT2B,MAD2L1,YWHAE ACTB,ALDH1A1,BUB1,CALR,CDC16,CDC20,CDC23,CDC27,CDK1,DD
cellular metabolic process	0.0487	X17,DPYD,ENO1,EZR,H2AFX,HAS1,HNRNPD,IDH2,KAT2B,MAD2L1, YWHAE

Table 6. Pathway Enrichment Analysis of Module B Genes Function

Gene Set	FDR	Node
oxidation-reduction process	5.82E-10	ALDH1A1,DPYD,ENO1,ENO2,ENO3,GAPDH,IDH2,PGK1,PKLR,PKM,TPI1
Carboxylic acid metabolic process	2.87E-09	CAD,DPYD,DPYS,ENO1,ENO2,ENO3,GAPDH,IDH2,PGK1,TPI1
single-organism catabolic process	7.08E-09	DPYD,DPYS,ENO1,ENO2,ENO3,GAPDH,PGK1,PKLR,PKM,TPI1

and 434 edges. The top ten principal nodes of node degree were screened. The hub genes were Cell division cycle 20 (CDC20), histone family, member X (H2AFX), Enolase 1 (ENO1), Actin beta (ACTB), interferon (IFN)-stimulated gene 15 (ISG15), Klysine acetyltransferase 2 b (KAT2B), Heterogeneous nuclear ribonucleoprotein D (HNRNPD), tryptophan 5-monooxygenase activation protein, epsilon (YWHAE), Gap junction protein alpha 1 (GJA1), Caveolin 1 (CAV1), respectively (Figure 3).

We used the plug-in MCODE analysis to further analyze the 218 nodes and 434 edges, and the top three most significant modules were selected. We also analyzed the function annotation of the genes in the modules, respectively. The consequences of enrichment analysis showed that the genes of these three modules were mainly related to the positive regulation of biological process, regulation of metabolic process, oxidation-reduction process and carboxylic acid metabolic process (Table 5-7).Ten hub genes related to cancers and metabolic diseases were analyzed by CTD. The results of the analysis indicated 1,128 distinct diseases with statistical significance (bonferroni-corrected P<0.05). It involved

Table 7. Pathway Enrichment Analysis of Module C Genes Function

Gene Set	FDR	Node
oxidation-reduction process	7.86E-12	ALDH1A1,DPYD,ENO1,ENO2,ENO3,GAPDH,IDH2,PGK1,PKLR,PKM,TPI1
generation of precursor metabolites and energy	2.45E-11	ENO1,ENO2,ENO3,GAPDH,IDH2,PGK1,PKLR,PKM,TPI1
nucleoside metabolic process	2.75E-10	DPYD,ENO1,ENO2,ENO3,GAPDH,PGK1,TPI1,UMPS

Table 8. The Relationship between Hub Gene with Ovarian Cancer or Metabolic Disease of the Comparative Toxicogenomics Database Prediction

| Hube Gene	Disease cantegories (p-value)			
	OC	Diabetes Mellitus, Type 1	Diabetes Mellitus, Type 2	Hypertriglyceridemia
CDC20	2.22E-09	1.56E-38	3.37E-98	5.93E-11
H2AFX	2.14E-09	1.32E-38	2.17E-98	5.67E-11
ENO1	2.24E-09	1.62E-38	3.69E-98	5.98E-11
ACTB	-	4.35E-35	1.21E-92	5.07E-11
ISG15	2.20E-09	1.49E-38	2.97E-98	5.85E-11
KAT2B	2.12E-09	1.27E-38	1.95E-98	5.60E-11
HNRNPD	2.15E-09	1.35E-38	2.30E-98	5.70E-11
YWHAE	1.88E-09	7.51E-39	4.88E-99	3.35E-09
GJA1	2.10E-09	1.23E-38	1.78E-98	5.55E-11
CAV1	2.12E-09	1.27E-38	1.95E-98	5.60E-11

-, means No available data
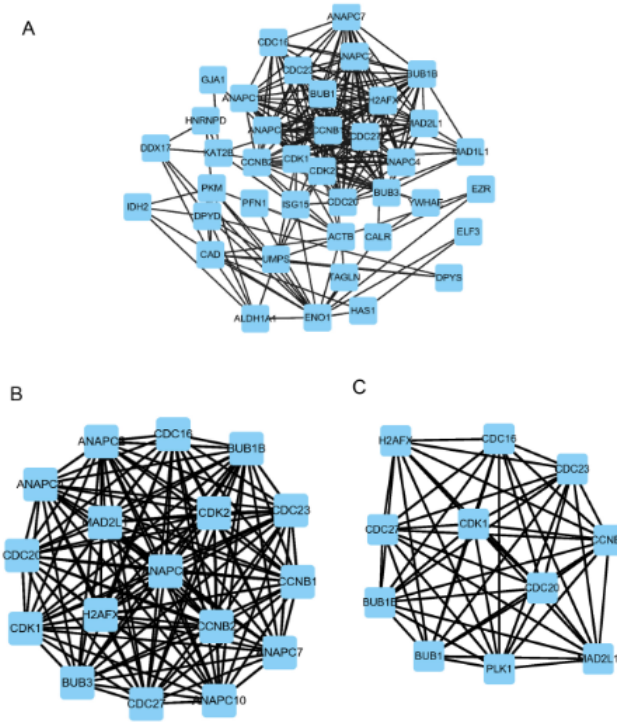
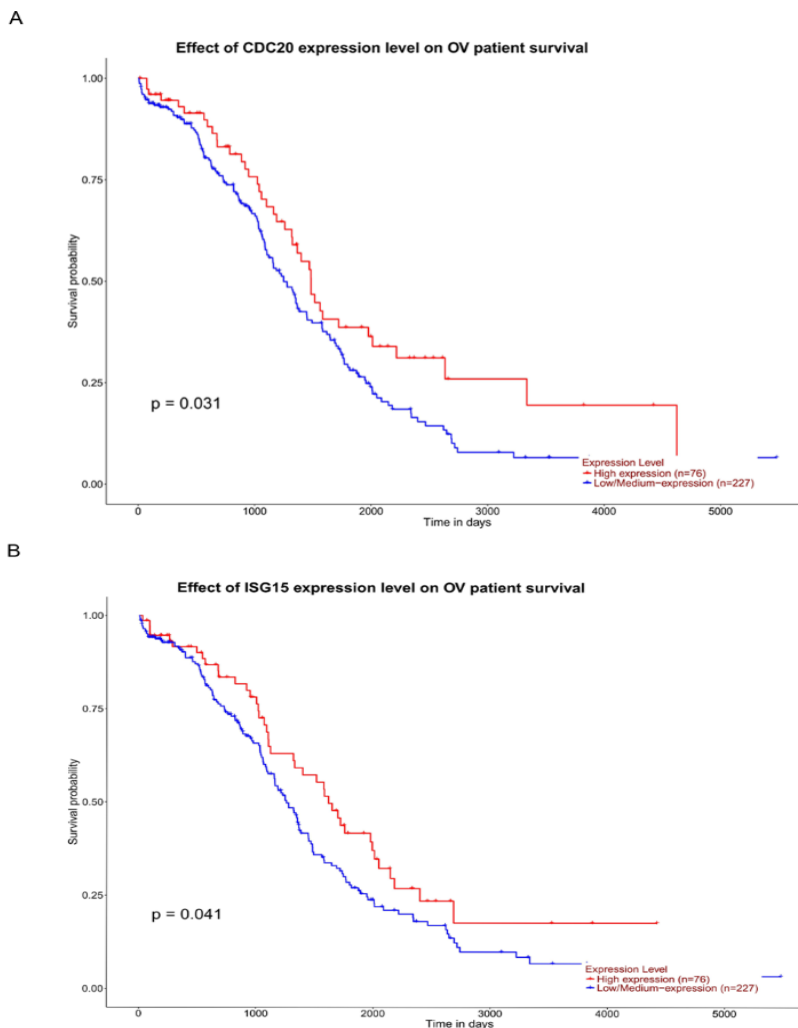Figure 3. DEGs Protein-Protein Interaction (PPI) Network Complex and Modular Analysis

Figure 4. K-M Cures of the Top 10 Hub Genes in OC. Red line represents high level of a hub gene, and green line represents low level. These genes were statistically significant. The X axis indicates overall survival time (day), and the Y axis indicates the survival rate.

185 kinds of cancer, including ovarian cancer. In addition, we found that the ten hub genes could cause many ovarian diseases and disorders, for example, ovarian cysts, ovarian neoplasms, ovarian epithelial cancer, and primary ovarian insufficiency. Interestingly, we also found that these genes were closely related to diabetes, weight change, hypertriglyceridemia and other metabolic disorders, particularly to type I and II diabetes. For the relationship between OV and diabetes, the results were displayed in Table 8.

The clinical significance of ten abnormally expressed hub genes

Further analysis was conducted on the above ten hub genes in abnormal expressions. The Kaplan-Meier survival curve showed that the survival time of CDC20, H2AFX, HNRNPD and ISG15 is significantly correlated with the survival time of OC (P < 0.05), as showed in Fig.4. Preliminary results showed that high expression of CDC20, H2AFX, HNRNPD and ISG15 in ovarian cancer prognosis has an obvious effect, compared with lower expression.

Discussion

It is extremely important to explore the risk factors and prognostic factors for the early diagnosis of OV. We received 185 OV samples and 10 normal samples from GSE26712 in GEO database. At last we identified 379 different genes, 104 up- and 275 down-regulated genes. The results of GO analysis showed that DEGs were mainly related to cell migration, cell proliferation and cell adhesion, the process of cell and macromolecular metabolism, and lipid response. Results of KEGG pathway analysis showed that DEGs were mainly related to signaling pathways such as Rap1, Ras, Wnt and PI3K-Akt as well as the metabolism of retinol and amino acids. We constructed PPI network diagrams genetic variations and obtained the most significant modules. Moreover the top ten hub genes were found from DEGs, which are: CDC20, H2AFX, ENO1, ACTB, ISG15, KAT2B, HNRNPD, YWHAE, GJA1 and CAV1.

Module analysis of PPI network demonstrated that Module A was focused on positive regulation of biological process, regulation of metabolic process, cellular metabolic process, that Module B was mainly enriched in the oxidation-reduction process, Carboxylic acid metabolic process, and that Module C was mostly concentrated in the oxidation-reduction process and generation of precursor metabolites and energy, nucleoside of the metabolic process. All of the three modules had statistical significance with MCODE score > 10, FDR < 0.05.

The functions of each hub gene were different. CDC20 exists in normal ovarian tissue and ovarian cancer tissue. High expression of CDC20 was associated with High tumor grade in the ovarian (P = 0.044). Studies using weighted gene co-expression network analysis (WGC) to analyze serous ovarian cancer (SOC) identified one stage-associated module and one grade-associated module.CDC20 was found to be one of the top hub genes

related to grade. In addition, studies showed that glucose and the activated Ras2 (Val19) protein sympathetically inhibited APC/C function via the cAMP/PKA pathway in yeast, at the same time CDC20 was involved in the APC/C regulation by the cAMP/PKA pathway (Bolte et al., 2003; Gayyed et al., 2016; Sun et al., 2017). H2AFX: In the treatment of ovarian cancer, the P38 lightning/ H2AX shaft was one of the molecular mechanisms of drug resistance in ovarian cancer. In addition, H2AX and SEI1 have co-localization in the nucleus, and the high expression of SHI1 in ovarian cancer plays a major role in the deterioration of ovarian cancer. At the same time, it is found that histone H2AX phosphorylation (gamma H2AX foci) expression was significantly increased in patients with type 1 diabetes, especially in women (Giovannini et al., 2014; Mo et al., 2016; You et al., 2017). ENO1 interferes with follicles in ovarian granulose cell by inducing the mRAN expression of hormone receptor (FSHR) and reducing the mRNA expression of luteinizing hormone receptor (LHR). Besides, ENO1 is a glycolysis enzyme that can reduce glycolysis in the cells of gastric cancer. The expression level of ENO1 increases with the shortening of the survival time in patients with gastric cancer. ENO1 is closely linked to glucose metabolism enzyme PGK1, and is up-regulated in the study of glucose uptake (Kim et al., 2013; Zhonghua et al., 2015; Zhao Y et al., 2016). ACTB is closely related to various cancers, including liver, melanoma, lung cancer, breast cancer, prostate cancer, uncontrolled ovarian cancer, and its expression rises in most tumor cells and tissues. Furthermore, abnormal expression of ACTB is associated with invasive and metastatic potential of cancer (Guo et al., 2013). In the chip analysis of tissues from 128 cases of ovarian serous carcinoma patients with surgery and chemotherapy treatment, it was found that ISG15 protein expression was significantly elevated in relapsed carcinomas as compared to primary tumors (P = 0.027) and ISG15-positive carcinomas had a significantly longer overall survival in university analysis (P = 0.002). The study also found the molecular basis for ubiquity and ISG15 cross-reactivity in viral ovarian tumor domains. IFNs: Using the Type I IFN receptors (NOD) IFNAR1 (- / -)) immunodeficient NOD mice model, it was found that ISG15 expression was significantly increased in mice at one week of age, and reached peak after 3-4 weeks. The results suggested that ISG15 is closely related to the function of pancreas in young mice (Akutsu et al., 2011; Darb-Esfahani et al., 2014). KAT2B: Whole genome analysis in the Han nationalities and ethnic minorities with high uric acid, Type 2 diabetes and obesity in China were carried on and it was found that KAT2B was closely related to HbA1c. The study also found that KAT2B and WDR5 stimulated gluconeogenesis through self-reinforcing cycle, and the small molecule inhibitors of KAT2B decreased the blood sugar levels, indicating KAT2B one of the effective targets in diabetes treatment. But no evidence showed that KAT2B concerned with ovarian carcinoma yet (Ravnskjaer et al.,2013; Wei et al., 2015). HNRNPD: It was found that AUF1/ HNRNPD over-expression could lead to tumor occurrence in the research of transgenic mice. Furthermore, in Akita mice

with type I diabetes, insulin caused the effects on Nrf2 and angiotensinogen (Agt) gene expression of the kidneys and unregulated heterogeneous nuclear ribonucleoprotein F and K (hnRNP F and hnRNP K). Insulin curbed Nrf^2 promoter activity via a specific DNA-responsive element that binds hnRNP F/K and hnRNP F/K that over expression curtailed Nrf2 promoter activity and hnRNP K. These findings identified hnRNP F/K and Nrf2 as potential therapeutic targets in diabetes (Gouble et al., 2002; Abdo et al., 2013; Lo et al., 2015; Singh 2016., Ghosh et al., 2017; Lo et al., 2017). YWHAE: We evaluated the expression of approximately 21,000 genes using DNA microarray screening of paired tumor samples taken prior to and after CT treatment from 6 patients with predominantly advanced stage, high-grade epithelial ovarian cancer. Up-regulated genes in post-CT tumors included research on diseases associated with diabetes or metabolic disorders. Substantial number of genes with previously known implication of mechanisms of tumorigenesis (L'Espérance et al., 2006). YWHAE: There is no research on the relationship between diabetes or dysmetabolic syndrome and YWHAE. GJA1: It was demonstrated that TGF-β up-regulateed GJA1/Cx43 in two human ovarian cancer cell lines, SKOV3 and OVCAR4. In the study of diabetes, it was discovered that intercellular coupling via gap junctions was decreased after insulin administration in diabetic and non-diabetic mice. This decrease in coupling was associated with a concomitant increase in the phosphorylation of GJA1/Cx43 at serine 368. Insulin regulated both gap junction-mediated intercellular communication and injury propagation in type I diabetic mouse heart. GJA1/Cx43 expression and cell-to-cell communication increased in response to elevated glucose and may protect the collecting duct from renal damage associated with established diabetic nephropathy (Hills et al., 2006; Qiu et al., 2015; Liu et al., 2015;Palatinuset al., 2015; Qiu et al., 2016). CAV1: In SKOV3 and A278 cells, it was reported that Cav1 promoted the chemoresistance of ovarian cancer by targeting apoptosis through Notch-1/Akt/ and NF-kappa B pathway. The other microarray analysis on ovarian tissues and SKOV- 3 and ES-2 cell lines proved that CAV1 gene was likely to act as a tumor suppressor gene in human ovarian epithelium. In addition, there was a significant change of Cav1 and mir-375 in the beta-cells (insulin secretory cells) in the islets of mice and people with different degrees of fusion. Fasudil was found in the diabetic rats of fasudil by blocking VEGFR2/Src/caveolin-1 signaling pathway to protect the diabetic rats (Wiechen et al., 2001; Zou et al., 2015; Ofori et al., 2017).

We used comparative toxicogenomics database to predict the onset of these ten hub genes with ovarian and metabolic diseases, and the results showed that the ten genes were associated with ovarian cancer or diabetes and metabolic disorders. Meanwhile, the analysis results of the ten survival curves illustrated that high expression of CDC20, H2AFX, HNRNPD and ISG15 had good prognosis in ovarian cancer. The studies of the above literatures indicated that five genes, GJA1, CAV1, ENO1, H2AFX and ISG15 were related to both ovarian cancer and diabetes, CDC20, however, was only involved in diabetes, and KAT2B, HNRNPD and YWHAE only in

ovarian cancer.

With the improvement of people's living conditions, the incidence of diabetes has increased year by year (Clery et al., 2017). Epidemiological studies have demonstrated that diabetes reduced the survival time and median survival of ovarian cancer patients.It was reported 23.3% newly diagnosed cervical cancer patients, almost twice as many as the previous Swiss census. Insulin resistance is common in type 2 diabetics, and insulin resistance often leads to hyperinsulinemia. The vitro studies have demonstrated that hyperinsulinemia affected the binding protein level of sex hormones, leading to elevated estradiol and testosterone levels, thus, affected the prognosis of patients with ovarian malignancy (Ruge et al., 2012). Besides, insulin inhibited cell apoptosis by affecting the PI3K/AKT pathway and mitotic kinase pathway, thereby inducing tumor cell proliferation (Gryko et al., 2014). Meanwhile, oxidative stress is closely related to the development of type 1 diabetes mellitus, which leads to genome damage, especially DNA double chain fracture. Previous studies have shown apparent metabolic changes that occur in cancer tissues, and low oxygen would result in higher HIF-1 alpha level, which increases the sugar intake in tumor cells. . Moreover, sugar increased within the tumor cells, not only leads to enhanced glycolysis pathway, but increases the biosynthesis at the same time. And studies in vitro have shown that P53 pathway influences the reprogramming of glucose metabolism in ovarian cancer cells. (Semczuk et al., 2017). It's found that metformin produced anti-cancer effects via AMPK dependent or non-dependent pathways, and the study provided evidence of diabetes and ovarian cancer. (Rattan et al., 2011). By using molecular bioinformatics tools we found out the possible common target genes in ovarian cancer and diabetes, provided evidence from the molecular level supporting that diabetes is a risk factor for ovarian cancer, which provides researchers a new way of thinking when facing the big challenges in cancer detection, early diagnosis and medicinal treatment of ovarian malignant tumors.

All in all, we studied the differences in ovarian cancer genes with bioinformatics analysis, aiming at the population based case-control study to collect information on relationship between diabetes and ovarian cancer. However, the study does have limitations and more experiments are needed to be done in order to further validate our observation, for example, to conduct experiments to validate the expression levels of these DEGs and to increase the sample size to confirm out findings.

Competing financial Interests

References

Abdo S, Lo CS, Chenier I, et al (2013). Heterogeneous nuclear ribonucleoproteins F and K mediate insulin inhibition of renal angiotensinogen gene expression and prevention of hypertension and kidney injury in diabetic mice. *Diabetologia*, **56**, 1649-60.

Akutsu M, Ye Y, Virdee S,et al (2011). Molecular basis for ubiquitin and ISG15 cross-reactivity in viral ovarian tumor domains. *Proc Natl Acad Sci U S A*, **108**, 2228-33.

Ashburner M, Ball CA, Blake JA, et al (2000). Gene ontology: tool for the unification of biology.The Gene Ontology Consortium. *Nat Genet*, **25**, 25-9

Birrer MJ, Bonome T, Levine DA, et al (2011).A gene signature predicting for survival in suboptimally debulked patients with ovarian cancery. Public on Jan 20.

Bolte M, Dieckhoff P, Krause C, et al (2003). Synergistic inhibition of APC/C by glucose and activated Ras proteins can be mediated by each of the Tpk1-3 proteins in Saccharomyces cerevisiae. *Microbiology*, **149**, 1205-16.

Chou JL, Su HY, Chen LY, et al (2010). Promoter hypermethylation of FBX032a novel TGF-p/SMAIM target gene and tumor suppressor is associated with poor prognosis in human ovarian cancer. *Lab Invest*, **90**, 414-25.

Clery P, Stahl D, lsmail K, et al (2017). Systematic review and meta-analysis of the efficacy of interventions for people with Type 1 diabetes mellitus and disordered eating. *Diabet Med*, **34**, 1667-75.

Darb-Esfahani S, Sinn BV, Rudl M, et al (2014). Interferon-stimulated gene, 15 kDa (ISG15) in ovarian high-grade serous carcinoma: prognostic impact and link to NF-κB pathway. *Int J Gynecol Pathol*, **33**, 16-22.

Donal ML (2009). Glucose as a prognostic factor in ovarian carcinoma. *Cancer*, **115**, 1021-6.

Gayyed MF, El-Maqsoud NM, Tawfiek ER, et al (2016). Comprehensive analysis of CDC20 overexpression in common malignant tumors from multiple organs: its correlation with tumor grade and stage. *Tumour Biol*, **37**, 749-62.

Ghosh A, Abdo S, Zhao S, et al (2017). Insulin inhibits Nrf2 gene expression via heterogeneous nuclear ribonucleoprotein F/K in diabetic Mice. *Endocrinology*, **158**, 903-19.

Giovannini C, Piaggi S, Federico G, et a l(2014). High levels of γ-H2AX foci and cell membrane oxidation in adolescents with type 1 diabetes. *Mutat Res*, **770**, 128-35.

Gouble A, Grazide S, Meggetto F, et al (2002). New player in oncogenesis: AUF1/hnRNPD overexpression leads to tumorigenesis in transgenic mice. *Cancer Res*, **62**, 1489-95.

Guo C, Liu S, Wang J, et al (2013). Elsevier B.V ACTB in cancer. *Clin Chim Acta*, **417**, 39-44.

Gryko M, Kiśluk J, Cepowicz D, et al (2014). Expression of insulin-like growth factor receptor type 1 correlate with lymphatic metastases in human gastric cancer. *Pol J Pathol*, **65**, 135-40.

Hills CE, Bland R, Wheelans DC, et al (2006). Glucose-evoked alterations in connexin43-mediated cell-to-cell communication in human collecting duct: a possible role in diabetic nephropathy. *Am J Physiol Renal Physiol*, **291**, 1045-51.

Kanehisa M, Goto S (2000). KEGG:kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*, **28**, 27-30.

Kim HJ, Lee HR, Chang Sup K, et al (2013). Investigation of protein expression profiles of erythritol-producing Candida magnoliae in response to glucose perturbation. *Enzyme Microb Technol*, **53**, 174-80.

Ledermann JA, Embleton AC, Raja F, et al (2016). Cediranib in patients with relapsed platinum-sensitive ovarian cancer (ICON6): a randomised, double-blind, placebo-controlled phase 3 trial. *Lancet*, **387**, 1066-74.

L'Espérance S, Popa I, Bachvarova M, et al(2006). Gene expression profiling of paired ovarian tumors obtained prior to and following adjuvant chemotherapy: molecular signatures of chemoresistant tumors. *Int J Oncol*, **29**, 5-24.

Liu W, Hua S, Dai Y, et al (2015). Roles of Cx43 and AKAP95 in ovarian cancer tissues in G1/S phase. *Int J Clin Exp Pathol*, **8**, 14315-24.

Lo CS, Shi Y, Chang SY, et al (2015). Overexpression of heterogeneous nuclear ribonucleoprotein F stimulates renal Ace-2 gene expression and prevents TGF-β1-induced kidney injury in a mouse model of diabetes. *Diabetologia*, **58**, 2443-54.

Lo CS, Shi Y, Chenier I, et al (2017). Heterogeneous nuclear ribonucleoprotein F stimulates Sirtuin-1 gene expression and attenuates nephropathy progression in diabetic Mice. *Diabetes*, **66**, 1964-78.

Mo Q, Zhang Y, Jin X, et al (2016). Geldanamycin, an inhibitor of Hsp90, increases paclitaxel-mediated toxicity in ovarian cancer cells through sustained activation of thep38/H2AX axis. *Tumour Biol*, **37**, 14745-55.

Ofori JK, Malm HA, Mollet IG, et al (2017). Confluence does not affect the expression of miR-375 and its directtargets in rat and human insulin-secreting cell lines. *Peer J*, **28**, e3503.

Palatinus JA, Gourdie RGJ (2016). Diabetes increases Cryoinjury size with associated Effects on Cx43 gap junction function and phosphorylation in the mouse heart. *Diabetes Res*, **10**, 8789617

Qiu X, Cheng JC, Klausen C, et al (2016). EGF-induced Connexin43 negatively regulates cell proliferation in human ovarian cancer. *J Cell physiol*, **231**, 111-9.

Qiu X, Cheng JC, Zhao J, et al (2015). Transforming growth factor-β stimulates human ovarian cancer cell migration by up-regulating connexin43 expression via Smad2/3 signaling. *Cell Signal*, **27**, 1956-62.

Quah HS, Miranda-Hernandez S, Khoo A, et al (2014). Deficiency in type I interferon signaling prevents the early interferon-induced gene signature in pancreatic islets but not type 1 diabetes in NOD mice. *Diabetes*, **63**, 1032-40.

Rattan R, Giri S, Hartmann LC, et al (2011). Metformin attenuates ovarian cancer cell growth in an AMP-kinase dispensable manner. *J Cell Mol Med*, **15**, 166-78.

Ravnskjaer K, Hogan MF, Lackey D, et al (2013). Glucagon regulates gluconeogenesis through KAT2B- and WDR5-mediated epigenetic effects., Montminy M. *Endocr J Clin Invest*, **123**, 4318-28.

Ruge T, Sukonina V, Kroupa O, et al (2012). Effects of hyperinsulinemia on lipoprotein lipase, angiopoietin-like protein 4, and glycosylphosphatidylinositol-anchored high-density lipoprotein binding protein 1 in subjects with and without type 2 diabetes mellitus. *Metabolism*, **61**, 652-60.

Semczuk A, Gogacz M, Semczuk-Sikora A, et al (2017). The putative role of TP53 alterations and p53 expression in borderline ovarian tumors - correlation with clinicopathological features and prognosis: A mini-review. *J Cancer*, **8**, 2684-91.

Singh AK, Lakhotia SC (2016). The hnRNP A1 homolog Hrp36 is essential for normal development, female fecundity, omega speckle formation and stress tolerance in Drosophila melanogaster. *Biochem Biophys Res Commun*, **471**, 260-5.

Sun Q, Zhao H, Zhang C, et al (2017). Gene co-expression network reveals shared modules predictive of stage and grade in serous ovarian cancers. *Oncotarget*, **8**, 42983-96.

Torre LA, Bray F, Siegel RL, et al (2012). Global cancer statistics. *CA Cancer J Clin*, **65**, 87-108.

Wei FJ, Cai CY, Yu P (2015). Quantitative candidate gene association studies of metabolic traits in Han Chinese type 2diabetes patients. *Genet Mol Res*, **14**, 15471-81.

Wei FJ, Cai CY, Yu P, et al (2015). Quantitative candidate gene association studies of metabolic traits in Han Chinese type 2 diabetes patients. *Genet Mol Res*, **14**, 15471-81.

Wiechen K, Diatchenko L, Agoulnik A, et al (2001). Caveolin-1 is down-regulated in human ovarian carcinom.a and acts as a candidate tumor suppressor gene. *Am J Pathol*, **159**, 1635-43.

You J, Liu J, Bao Y, et al (2017). Induces genomic instability by inhibiting DNA damage response in ovarian cancer. *Cancer Lett*, **385**, 271-9.

Zhao Y, Yan X, Li X, et al (2016). PGK1, a glucose metabolism enzyme, may play an important role in rheumatoid arthritis. *Inflamm Res*, **65**, 815-25.

Zhonghua Z, Liu Za, Zhi H, et al (2015). Screening of citrullinated proteins in ten tumor cell lines. *Zhonghua Zhong Liu Za Zhi*, **37**, 165.

Zou W, Ma X, Hua W, et al (2015). Caveolin-1 mediates chemoresistance in cisplatin-resistant ovarian cancer cells by targeting apoptosis through the Notch-1/Akt/NF-κB pathway. *Oncol Rep*, **34**, 3256-63.