SPANDIDOS PUBLICATIONS

ONCOLOGY LETTERS 16: 2453-2461, 2018

Identification of potential target genes in pancreatic ductal adenocarcinoma by bioinformatics analysis

YUCHEN TANG[1,2*], ZIXIANG ZHANG[1,2*], YAOCHENG TANG[1], XINYU CHEN[1] and JIAN ZHOU[1,2]

[1]Department of General Surgery and [2]Pancreatic Disease Research Center,
The First Affiliated Hospital of Soochow University, Suzhou, Jiangsu 215006, P.R. China

Received January 23, 2018; Accepted May 22, 2018

DOI: 10.3892/ol.2018.8912

Abstract. Pancreatic ductal adenocarcinoma (PDAC) is one of the most complicated and fatally pathogenic human malignancies. Therefore, there is an urgent need to improve our understanding of the underlying molecular mechanism that drives the initiation, progression, and metastasis of PDAC. The aim of the present study was to identify the key genes and signaling pathways associated with PDAC using bioinformatics analysis. Four transcriptome microarray datasets (GSE15471, GSE55643, GSE62165 and GSE91035) were acquired from Gene Expression Omnibus datasets, which included 226 PDAC samples and 65 normal pancreatic tissue samples. We screened differentially expressed genes (DEGs) with GEO2R and investigated their biological function by Gene Ontology (GO) and Kyoto Encyclopedia of Genes (KEGG) analysis. The overall survival data was obtained from UALCAN, which calculated the data shared with The Cancer Genome Atlas. In addition, a protein-protein interaction (PPI) network of the DEGs was constructed by STRING and Cytoscape software. The four sets of DEGs exhibited an intersection consisting of 205 genes (142 up-regulated and 63 down-regulated), which may be associated with PDAC. GO analysis showed that the 205 DEGs were significantly enriched in the plasma membrane, cell adhesion molecule activity and the Energy pathways, and glycine, serine, threonine metabolism were the most enriched pathways according to KEGG pathway analysis. Kaplan-Meier survival analysis revealed that 22 of 205 common genes were significantly associated with the overall survival of pancreatic cancer patients. In the PPI network and sub-network, DKK1 and HMGA2 were considered as hub genes with high connectivity degrees. DKK1 and HMGA2 are strongly associated with WNT3A and TP53 separately, which indicates that they may play an important role in the Wnt and P53 signaling pathways. Using integrated bioinformatics analysis, we identified DKK1 and HMGA2 as candidate genes in PDAC, which may improve our understanding of the mechanisms of the pathogenesis and integration; the two genes may be therapeutic targets and prognostic markers for PDAC.

Correspondence to: Dr Jian Zhou, Department of General Surgery, The First Affiliated Hospital of Soochow University, 188 Shizi Street, Suzhou, Jiangsu 215006, P.R. China
E-mail: zhoujian06@suda.edu.cn

*Contributed equally

Key words: bioinformatics analysis, PDAC, differentially expressed genes, DKK1, HMGA2

Introduction

Pancreatic ductal adenocarcinoma (PDAC) is a highly lethal disease among all types of cancer, with an average of 43,090 cases of mortality reported every 5 years (1). The relative 5-year survival rate is merely 6% (2). This low rate is due to several factors, of which the most important factor may be that that more than half of cases are diagnosed at a late stage, for which the 5-year survival is only 3% (1). According to the latest statistics, at least half of patients with PDAC are asymptomatic until the disease develops to a distant stage (3). Consequently, many patients miss the optimal period for effective systemic therapy without symptoms anesis or disease regression. Unfortunately, the biological characteristics of PDAC manifest not only as early recurrence and invasion, but also as chemoresistance and radioresistance (4,5). From this, it is urgent for us to figure out the pathophysiology and progression mechanism.

Over the past decades, plentiful clinical and experiment research of PDAC has led to the identification of more sensitive and effective biomarkers of PDAC. Mazarico *et al* reported that CHKα may be considered to be a therapeutic target in PDAC and has suggested the possibility of new underlying mechanisms (6). The present study reported that SULF2 expression was independently associated with poor survival and may be a therapeutic target for patients with PDAC (7). These findings provide a good foundation to analyze key genes associated with PDAC that may act as diagnostic, prognostic or therapeutic biomarkers; meanwhile, it must be acknowledged that experimental conditions differ from one another and the genes associated with PDAC are numerous. Therefore, it is necessary to unify experimental conditions, taking these aspects into consideration, only then can we screen additional key genes associated with PDAC.

Fortunately, many scientists have provided multiple genetic chips, second-generation sequencing and other forms of

high-throughput sequencing to public web platforms, which are freely available to academic and nonprofit cancer research communities. With the availability of data from large-scale omics data like Gene Expression Omnibus (GEO) (8), The Cancer Genome Atlas (TCGA) (9) and Oncomine (10) and others, it is possible for us to compare cancer profiles with normal profiles in multiple aspects. In this study, we used bioinformatics methods to analyze the mRNA expression data of PDAC to identify differentially expressed genes (DEGs), Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. A protein-protein interaction (PPI) network was also constructed to identify the key genes associated with PDAC, in attempt to provide valuable information for the investigation into the mechanism underlying the pathogenesis of PADC, and for the identification of diagnostic and therapeutic targets of PDAC.

Materials and methods

Microarray data information. PDAC datasets were obtained from the Gene Expression Omnibus (GEO, available online: https://www.ncbi.nlm.nih.gov/geo/) and The Cancer Genome Atlas (TCGA, available online: https://cancergenome.nih.gov/). The DEGs were identified using four independent PDAC microarray datasets, including GSE15471, GSE55643, GSE62165 and GSE91035, with 226 primary tumor samples and 65 normal control samples. The microarray data of GSE15471 was produced using the GPL570 Platform [(HG-U133_Plus_2) Affymetrix Human Genome U133 Plus 2.0 Array], including 36 matched tumor and normal samples. The GSE55643 dataset was based on the GPL6480 Platform (Agilent-014850 Whole Human Genome Microarray 4x44K G4112F) and composed of 45 PDAC and 8 normal samples. The GSE62165 dataset was based on GPL13667 Platform [(HG-U219) Affymetrix Human Genome U219 Array], which contained 118 surgically resected PDAC and 13 control samples. We used a validation dataset, GSE91035, which included 27 PDAC and 8 normal pancreatic tissues. Moreover, the gene dataset was based on GPL22763 (Agilent-039714 LincRNA SurePrint G3 Human GE 8x60K Microarray PVD 028004). GSE32676, an independent dataset for further confirmation, contained 7 normal tissues and 25 PDAC tissues, was based on Platform [(HG-U133_Plus_2) Affymetrix Human Genome U133 Plus 2.0 Array]. All of the clinical datasets included 92 normal and 252 tumor tissues, which were diagnosed as PDAC (Table I).

Data preprocessing. The raw probe-level data was normalized and converted to expression profiles using the Affy package of R (11). Background correction and quartile data normalization were applied for the five datasets. Annotations for the probe arrays were downloaded from GEO. If multiple probe sets corresponded to the same gene, the mean expression value was used as the expression value.

Screening of DEGs. GEO2R (available online: https://www.ncbi.nlm.nih.gov/geo/geo2r/), a web-portal for the identification of genes that exhibit differentially expressed according to experimental conditions, was used to identify the DEGs between PDAC and normal tissues. We identified that Genes with |log2 fold change (FC)|>1 and P<0.05 were differentially expressed between PDAC and normal controls. The volcano plot was achieved by using ggplot2. A Venn diagram was produced using VENNY (available online: http://bioinfogp.cnb.csic.es/tools/venny/index.html), a scientific service of the Spanish National Biotechnology Centre (CNB).

Functional annotation and pathway enrichment analysis. GO analysis is an extraordinary useful method for annotating genes and identifying biological characteristics, including biological process, cellular component and molecular function, based on high-throughput genomic or transcriptomic data (12). The KEGG pathway database is a synthetic database, which includes a variety of biochemical pathways (13). In addition, the annotation, data integration and visualization of the DEGs were processed by DAVID (available online: https://david.ncifcrf.gov/) and FunRich software (14). GO or KEGG analyses with P<0.05 considered to indicate a statistically significant difference..

Survival analysis of DEGs. For validation, UALCAN (available online: http://ualcan.path.uab.edu/index.html), an interactive web portal for the in-depth analysis of TCGA gene expression data, was used for survival analysis (15). Survival analysis with P<0.05 was considered statistically significant.

Integration of PPI network and module analysis. The Search Tool for the Retrieval of Interacting Genes (STRING, available online: http://www.stringdb.org/) database is a commonly used online tool designed to calculate information regarding PPIs (16). Moreover, PPI networks can help identity the key genes associated with PDAC development on the level of protein interactions. Then, PPI networks were constructed using Cytoscape software (v.3.51) (17). The CentiScape module was used to screen the nodes of the PPI network in Cytoscape and the degree value was set as a criterion for key genes. Genes with the highest degree scores were considered to be candidate key genes. P<0.05 was set as the filter criterion.

Statistical analysis. Numerical data were expressed as mean ± SD GraphPad Prism v.5.0 (GraphPad Software, Inc., La Jolla, CA, USA) and SPSS software v.18.0 (SPSS, Inc., Chicago, IL, USA) was used for statistical analysis. A Student's t-test was performed to compare two groups of gene expression. The correlation between mRNA expression and clinicopathological features was assessed using the χ^2 test and Student's t-test. Survival analysis was performed through the Kaplan-Meier method and the log-rank test was used to evaluate the statistical significance of the differences. P<0.05 was considered to indicate a statistically significant difference.

Results

Identification of DEGs. The gene expression profiles of GSE15471, GSE55643, GSE62165 and GSE91035 were obtained from NCBI-GEO and used for the identification of genes differentially expressed in PDAC. The gene datasets included 226 tumor tissues and 65 normal tissues. A total of 1791, 1641, 4061 and 3602 DEGs were identified from GSE15471, GSE55643, GSE62165, and GSE91035 datasets, respectively. In particular, there were 1558 upregulated

Table I. Clinical characteristics of participants.

	Tissues		Sex		Stages (AJCC)			
Dataset	Tumor	Normal	Male	Female	I	II	III	IV
GSE15471	39	39	-	-	-	-	-	-
GSE55643	45	8	32	21	-	-	-	-
GSE62165	118	13	-	-	8	92	5	13
GSE91035	25	23	-	-	-	-	-	-
GSE32676	25	7	-	-	2	23	0	0

AJCC, American Joint Committee on Cancer.

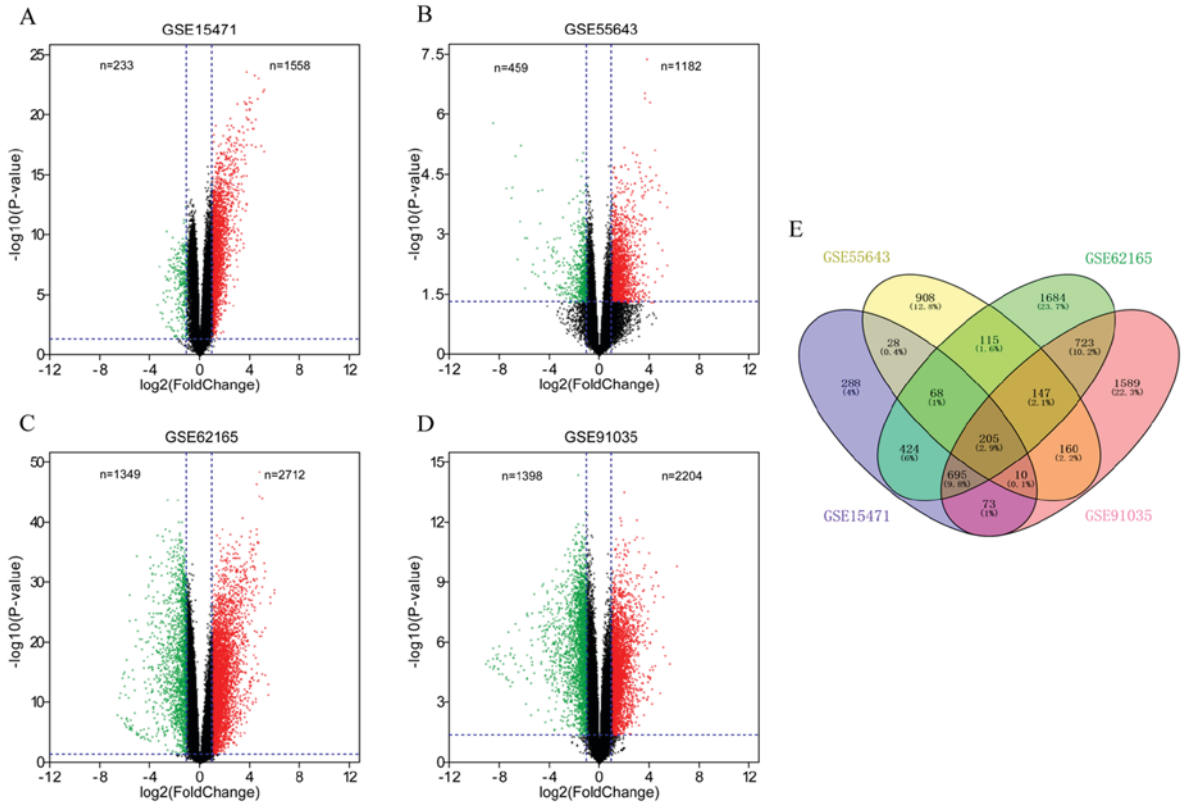

Figure 1. Identification of DEGs between PDAC and non-malignant tissues. (A-D) Volcano plots of differentially expressed genes. (A) There were 1558 genes upregulated and 233 downregulated genes in GSE15471, (B) 1182 genes upregulated and 459 downregulated genes in GSE55643, (C) 2712 genes upregulated and 1349 downregulated genes in GSE62165 and (D) 2204 genes upregulated and 1398 downregulated genes in GSE91035. (E) Venn diagram of the four sets of DEGs. There were 205 DEGs common to all DEGs sets. DEGs, differentially expressed genes; PDAC, pancreatic ductal adenocarcinoma.

and 233 downregulated genes in GSE15471, 1182 upregulated and 459 downregulated genes in GSE55643, 2712 upregulated and 1349 downregulated genes in GSE62165, 2204 upregulated and 1398 downregulated genes in GSE91035 (Fig. 1A-D). The intersection of the four sets of DEGs included 205 genes (142 upregulated and 63 downregulated), which were common to all PDAC samples analyzed (Fig. 1E) and were listed in Table II.

Functional enrichment analysis of DEGs. GO was analyzed and processed by FunRich software. The annotation of the 205 DEGs were mainly classified into three functional groups: Cellular component, molecular function and biological process groups (Fig. 2A-C). As shown in Fig. 2A, in cellular component group, GO analysis suggested that the common DEGs are significantly enriched in the plasma membrane, exosomes, extracellular, cytoplasmic vesicle, extracellular space and cell surface. In terms of molecular function, the enriched GO terms were mainly cell adhesion molecule activity, transaminase activity, metallopeptidase activity, calcium ion binding, inward rectifier channel and steroid binding (Fig. 2B). In addition, biological process analysis also revealed that the DEGs were significantly enriched in the energy pathways, transport, immune response, cell adhesion, regulation of cellular process and humoral immune response (Fig. 2C). Moreover, KEGG

Table II. Common DEGs identified in PDAC.

Regulation	DEGs (gene symbol)
Upregulated	COL1A1, CXCL5, S100P, CEACAM6, GABRP, SLC6A14, CLDN18, GPRC5A, ANO1, RUNX2, ANTXR1, KIF26B, PHLDA2, PCDH7, CEACAM5, LAMC2, KYNU, LEF1, SFN, SLC6A6, AHNAK2, SHISA2, SULF2, TFF1, CD109, PLAU, SERPINB5, SDR16C5, DKK1, LAMA3, IFI27, TGM2, OSBPL10, NQO1, ECT2, HOXB3, MMP11, TPM2, IL1RN, TMEM158, S100A11, S100A4, NHS, TRIM29, NPR3, CD55, TMPRSS4, FOXQ1, CAPN8, DKK3, PLAUR, CXCL3, SYTL2, CEACAM7, RSAD2, ADAM28, XAF1, BICD1, FXYD3, SCD, VSIG1, TSPAN1, FXYD5, IGF2BP3, KCNK1, TWIST1, TMPRSS3, NMU, MTMR11, GBP2, PI3, PLAC8, PGM2L1, INPP4B, GBP3, S100A2, CRIP1, DCBLD2, MGLL, TPBG, FERMT1, NRP2, BIK, OSBPL3, PHLDA1, MSLN, MPZL2, NT5E, LY6E, HK2, MBOAT2, HN1, ARNTL2, CDH3, MLPH, LAYN, ARHGAP26, LRRC15, ANXA2, MALL, TNFSF11, MUC4, CEACAM1, SLC2A1, KITLG, CXCL10, ITGA3, IER5L, HMGA2, LEMD1, EREG, ID1, OAS3, ASAP2, SCEL, STYK1, ST6GAL2, ITGB4, ULBP2, MMP14, ACSL5, RHBDL2, S100A16, LIF, ECM1, ZG16B, ZNF365, GALNT5, KCNN4, EGFL6, IFI6, ADAMTS6, AHNAK, YWHAZ, TRIM31, ASPHD2, HOXB6, CORO2A, UBE2S, CDH6, IL1RAP, CENPF
Downregulated	BTG2, SEC11C, SLC17A4, CYB5A, ABAT, LMO3, NUCB2, SLC4A4, FAM129A, AOX1, ACAT1, GMNN, CRAT, ARSE, AKR7A3, CTH, FAM46C, ECHDC3, SLC39A14, DMD, SLC39A8, GAMT, SERPINI1, FKBP11, SLC30A2, DDC, SLC25A15, SEL1L, SMPDL3B, GATM, TCEA3, SPAG4, EPB41L4B, NR5A2, GAS2, CCDC69, BNIP3, SYBU, DPEP1, MT1G, PDK4, GLS2, SLC25A45, MYRIP, KCNJ5, EPHX2, SLC39A5, CBS, TPST2, C5, PAIP2B, CCDC110, COCH, PSAT1, TEX11, RNF186, BHLHA15, ZG16, DPP10, TMEM52, ANPEP, DNASE1, TMED6

DEGs, differentially expressed genes; PDAC, pancreatic ductal adenocarcinoma.

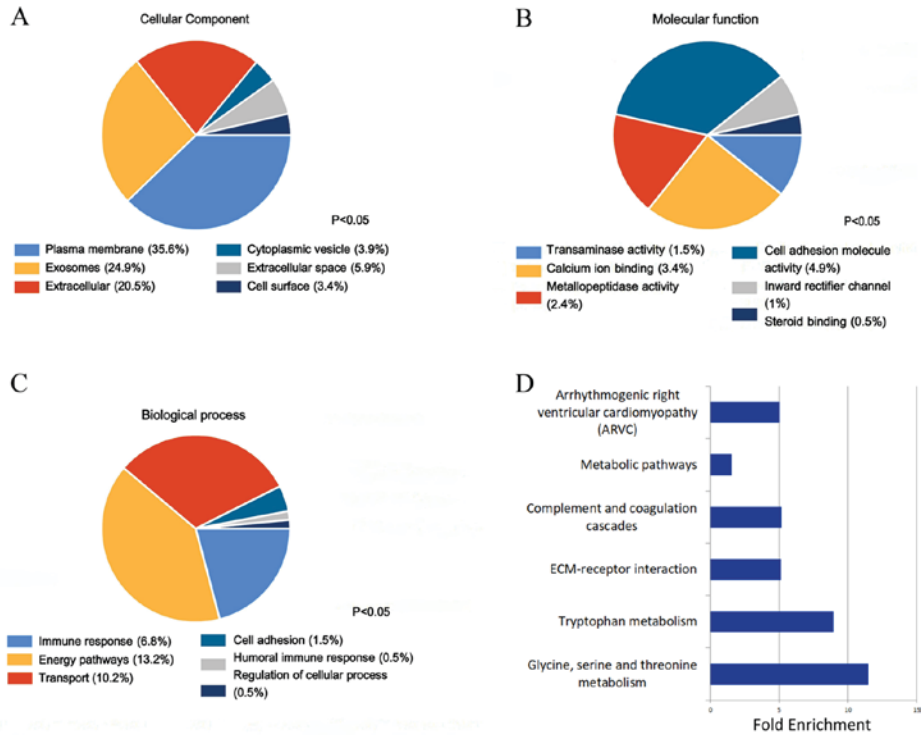

Figure 2. GO and KEGG pathway enrichment analysis of the 205 DEGs. (A-C) GO terms of significantly enriched genes by (A) cellular component, (B) molecular function and (C) biological process, P<0.05. (D) Significantly enriched KEGG pathways, P<0.05. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEGs, differentially expressed genes.

pathway analysis indicated that glycine, serine, threonine and tryptophan metabolism were the most significantly enriched pathways (Fig. 2D).

PPI network construction and modules selection. Based on the information obtained from the STRING database, we produced a network diagram. A total of six of the twenty-two

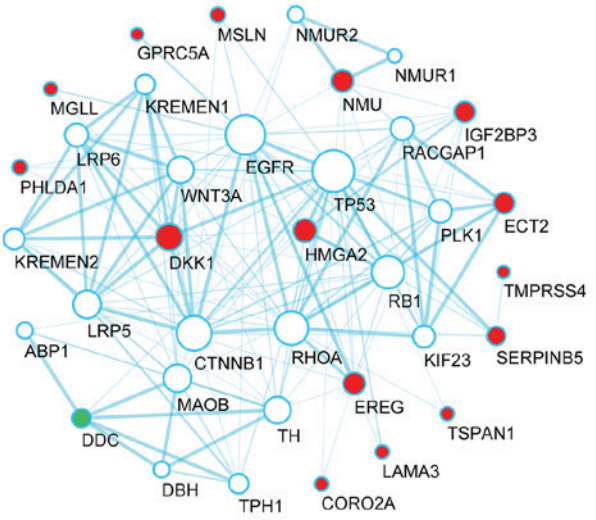

Figure 3. The protein-protein interaction network of the top 15 upregulated, and 1 downregulated, DEGs in PDAC. The red circular nodes represent downregulation DEGs in PDAC. The green circular nodes represent the downregulated DEGs in PDAC. The diameter of the nodes indicates degree score. Solid lines indicate interaction between DEGs and proteins, and the width of the solid lines indicates the combined score. DEGs, differentially expressed genes; PDAC, pancreatic ductal adenocarcinoma.

genes were on the margin and isolated, suggesting no association with other genes, and so were removed from the network diagram. Then, we constructed a PPI network diagram (Fig. 3) with the 16 associated genes by Cytoscape software. There were 36 nodes and 160 edges in the network; genes with higher degrees of association with PDAC were screened as hub genes, and were scored using the Centiscape module. Genes with high scores, including DKK1 (degree=13) and HMGA2 (degree=10), were selected as the hub genes. As shown in Fig. 3, the red circular nodes represent upregulated DEGs and green circular nodes represent downregulated DEGs, respectively.

Screening DEGs by survival analysis. By using UALCAN, we attempted to investigate the associations between gene expression and overall survival; associations that were significantly different between normal and pancreatic cancer tissues ($P<0.05$) were considered to be candidate genes. We obtained 22 genes from the 205 DEGs, including GABRP, EREG, PHLDA1, MGLL, DDC, CORO2A, TMPRSS4, SDR16C5, LAMA3, LEMD1, NMU, HMGA2, DKK1, FERMT1, MSLN, AHNAK2, S100A16, GPRC5A, IGF2BP3, ECT2, SERPINB5 and TSPAN1. As shown in Fig. 4, patients with high expression levels of DKK1 or HMGA2 had significantly poorer overall survival than those with low expression levels.

Validation of DEGs in independent PDACs. As DKK1 and HMGA2 were selected from the other DEGs, further confirmation of the altered expressions was necessary. GSE32676, an independent dataset that contained 7 normal tissues and 25 PDAC tissues, was used to validate the expressions of the two genes. As previously seen, the expressions levels of DKK1 and HMGA2 were significantly elevated in PDAC tissues than in normal tissues for each of the four datasets (Fig. 5). In addition, it had been acknowledged that DKK1 and HMGA2 had higher expression in tumor tissue than in normal tissue (Fig. 6A). Furthermore, we attempted to further investigate the relationship between the prognosis and the two genes in more samples. Attributed to GEO, from where we obtained the two datasets, GSE78229 and GSE57495, we found that the differential expressed DKK1 and HMGA2 genes were statistically different with the prognoses of patients (Fig. 6B).

Correlation between differentially expressed mRNAs and clinicopathological features of patients with PDAC. Next, to assess the association of the clinicopathological features and the two candidate genes, we compared DEGS, DKK1 and HMGA2 with the clinicopathological features of patients with TCGA datasets. The mean ± SD age for all 177 patients was 64.86±11.05 years, and the mean ± SD follow-up time was 566.47±472.79 days. As shown in Table III, significant differences in the distribution of vital status, AJCC pathological stage, tumor size and tumor grade between the differential expressed DKK1 groups were noted. In addition, analysis showed that there was a statistical difference between the differential expressed HMGA2 groups in vital status and tumor grade.

Discussion

Although numerous clinical and basic studies on PDAC have been conducted, the overall incidence and rates mortality have not markedly changed over the past decades. This may be due to the lack of reliable biomarkers for detection of early stage PDAC and of effective treatment for more advanced stages of PDAC. Therefore, comprehensive studies to improve understanding of the molecular mechanisms of the pathogenesis of PDAC are extremely important and necessary. The rapid development of microarray technology must be acknowledged as it is easier and probable that general genetic alterations involved progression of diseases may be detected, providing that more gene targets for diagnosis, therapy and prognosis of PDAC are discovered.

In the present study, four transcriptome microarray datasets from different groups were integrated and analyzed by bioinformatics methods; 205 DEGs were screened, consisting of 142 upregulated and 63 downregulated genes. GO term analysis indicated that the DEGs were mainly enriched in energy pathways, transport, immune response, plasma membrane and cell adhesion molecule activity. It was first reported in the 1920s by Otto Warburg, that most cancer cells mainly utilize aerobic glycolysis and lactic acid fermentation rather than mitochondrial oxidative phosphorylation for energy production (5). Recent research has shown that PDAC tumor cells are also highly dependent on glutamine metabolism to support pancreatic cancer growth (8). The overexpression of ATP in various cancers has been confirmed to be associated with more aggressive tumor progression, greater invasiveness and poorer prognosis (18). In addition, dysfunction of cell adhesion is associated with the invasiveness and metastasis of PDAC (19).

Then, survival analysis using the 205 DEGs revealed that 22 DEGs were significantly associated with overall survival of patients. Furthermore, we constructed the PPI network and identified two high-scoring genes, including DKK1 and HMGA2, inextricably linked with WNT3A and TP53, which are widely for their importance in the progression of cancer (10,20).

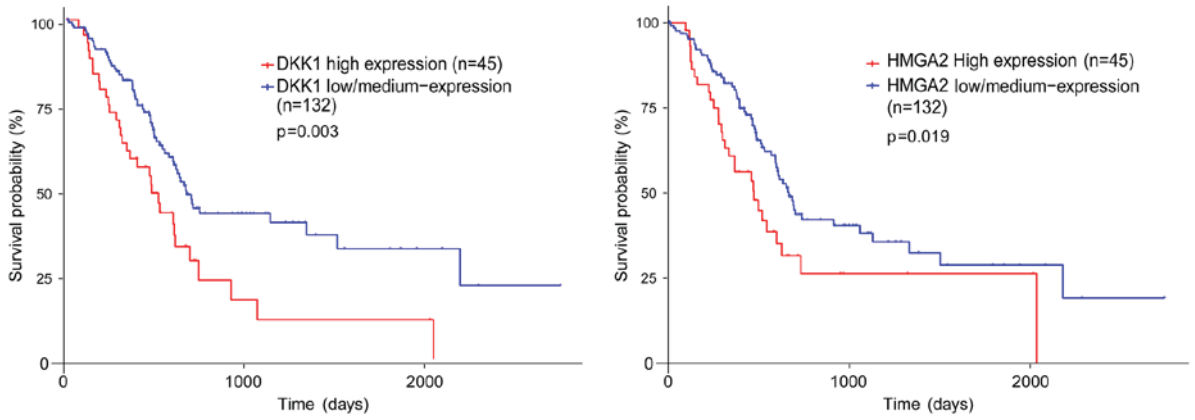

Figure 4. The overall survival curve of DKK1 and HMGA2 with different expressions. They were obtained from UALCAN. DKK1, dickkopf WNT signaling pathway inhibitor 1; HMGA2, high mobility group AT-hook 2.

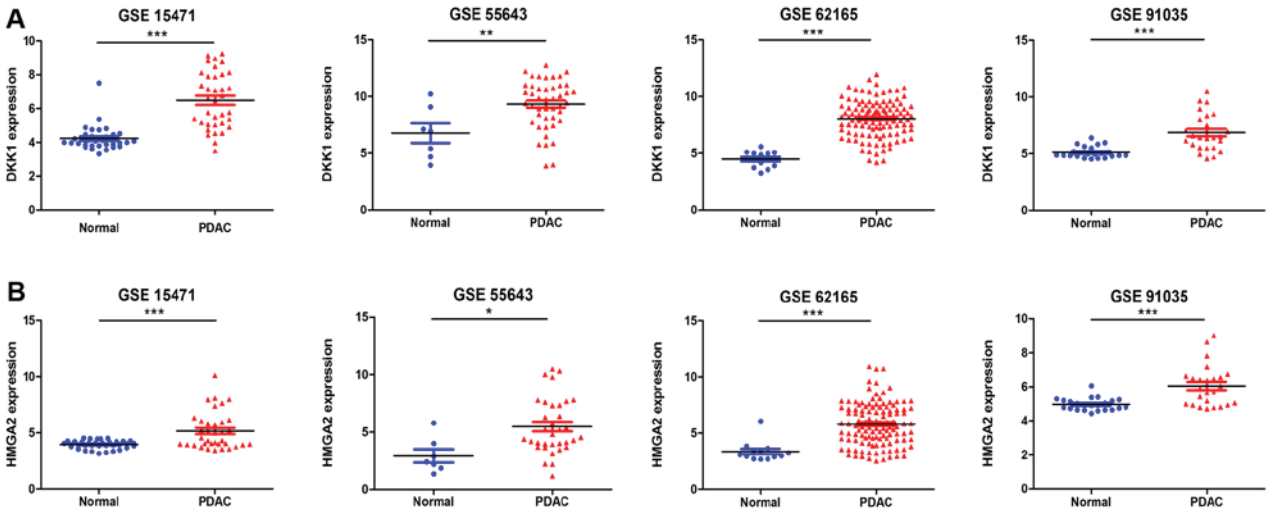

Figure 5. Differential expression of DKK1 and HMGA2 in the four datasets. (A) Expression of DKK1 was significantly increased in PDAC tissues than in normal tissues. (B) Expression of HMGA2 was significantly elevated in PDAC tissues than in normal tissues. *P<0.05, **P<0.01, ***P<0.001. DKK1, dickkopf WNT signaling pathway inhibitor 1; HMGA2, high mobility group AT-hook 2; PDAC, pancreatic ductal adenocarcinoma.

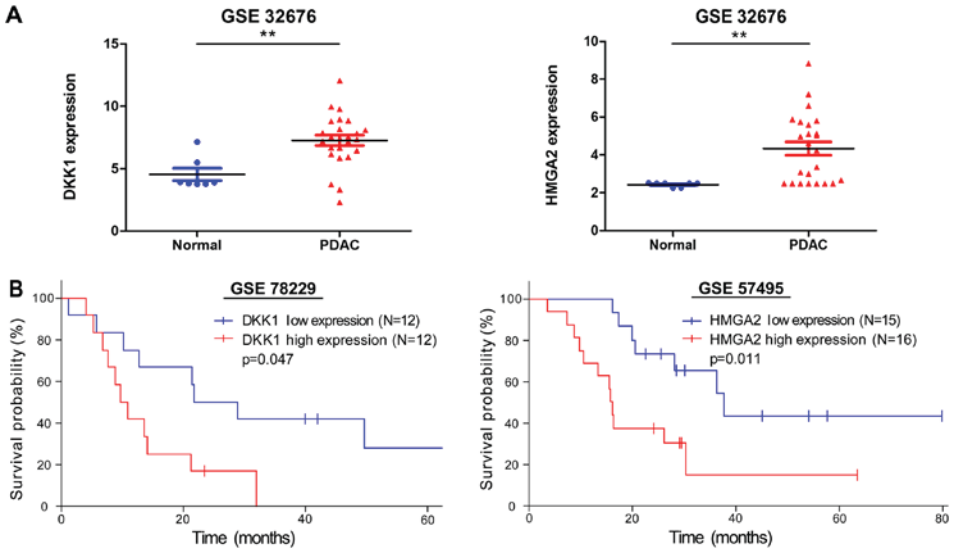

Figure 6. Validation of the altered expressions and Kaplan-Meier survival curves of DKK1 and HMGA2. (A) Elevated expression of DKK1 and HMGA2 in PDAC tissues in GSE32676. **means P<0.01. (B) Kaplan-Meier survival curves according to DKK1 and HMGA2 expression based on GSE 78229 and GSE 57495 respectively (P<0.05). DKK1, dickkopf WNT signaling pathway inhibitor 1; HMGA2, high mobility group AT-hook 2; PDAC, pancreatic ductal adenocarcinoma.

Table III. Association between DKK1, HMGA2 and clinicopathological features of patients with PDAC.

	DKK1 expression			HMGA2 expression		
Variables	Low	High	P-value	Low	High	P-value
Age, years	67.18±9.46	64.26±11.03	0.191	64.68±10.91	64.72±10.84	0.986
Sex			0.054			0.272
Male	20	77		18	79	
Female	8	72		10	70	
Vital status			0.002[b]			0.007[b]
Alive	21	64		20	65	
Succumbed	7	85		8	84	
AJCC stage			0.009[b]			0.958
Stage I	11	10		6	15	
Stage II	16	130		20	126	
Stage III	0	3		0	3	
Stage IV	0	4		0	4	
NA	1	2		2	1	
Tumor size			<0.001[b]			0.262
T1	3	4		1	6	
T2	12	12		8	16	
T3	12	129		17	124	
T4	0	3		0	3	
TX	1	0		1	0	
Lymph node involvement			0.377			0.469
N0	13	36		9	40	
N1	12	111		16	107	
NX	2	2		2	2	
NA	1	0		1	0	
Metastasis status			0.111			0.239
M0	9	70		10	69	
M1	0	4		0	4	
MX	19	75		18	76	
Tumor grade			0.024[a]			0.002[b]
G1	11	20		11	20	
G2	12	82		13	81	
G3	4	44		4	44	
G4	0	2		0	2	
NA	1	1		0	2	

[a]P<0.05; [b]P<0.01; NA, no data; AJCC, American Joint Committee on Cancer; PDAC, pancreatic ductal adenocarcinoma; DKK1, dickkopf WNT signaling pathway inhibitor 1; HMGA2, high mobility group AT-hook 2.

DKK1 is a member of the Dickkopf family and a secreted protein with two cysteine rich regions, which has already been defined as a direct inhibitor of the Wnt signaling pathway by binding to LRP5/6 coreceptors (21). New research shows that DKK1 may participate in the progression of cancer via epithelial-mesenchymal transition or the Wnt/β-catenin signaling pathway (22-24). Previous studies have confirmed that DKK1 was downregulated in human colon cancer and may have contributed to the suppression of colon cancer by inhibiting epithelial-mesenchymal transition (22,25). Nevertheless, the expression of DKK1 is upregulated in non-small cell lung cancer, esophageal squamous cell carcinoma and hepatocellular carcinoma, as well as in PDAC (23,24,26,27). An *in vitro* experiment showed that the knockdown of DKK1 may suppress the invasion and migration of the PDAC cell line, SUIT-2, indicating that DKK1 may play a positive role in the progression of PDAC (26). With serum mRNA detection, another study found that the detection of DKK1 was more sensitive than CA199 in the diagnosis of pancreatic cancer, especially in early stage. It has been confirmed that CKAP4, a DKK1 receptor, could promote pancreatic cancer cell proliferation through the PI3K/AKT signaling pathway (28), and may be a

potential therapeutic target. Moreover, in the group of patients with higher DKK1 expression levels, the poorer overall survival rate and median survival time were reported (29).

Another hub gene, HMGA2, is a member of the non-histone chromosomal high mobility group (HMG) protein family. HMG proteins function as architectural factors and are essential components of the enhanceosome, and may act as transcription regulating factors for that these proteins contain structural DNA-binding domains (30,31). HMGA2 interferes with E4F1 binding to the ATF/CRE site on the CCNA2 promoter, and so positively regulates CCNA2 expression and plays a part in the cell cycle (32). HMGA2 is deemed to maintain EMT in pancreatic cancer cells, via oncogenic RAS signaling (33). A recent study indicated that PDGFRβ-positive fibroblasts was closely linked to HMGA2 in malignant cells and that HMGA2 expression was enhanced by paracrine stroma epithelial signaling (34). In addition, HMGA2 was also correlated with the poor prognoses of patients with PDAC in a univariate survival analysis (35). This is consistent with the survival rate that we predicted using a vast amount of biological data. It has been reported that HMGA2 was associated with a p53 mutation in high-grade papillary serous carcinoma (36,37) and it may cause a malignant phenotype when coupled with MDM2, which participates in P53 negative feedback (38). In addition, this indicated that HMGA2 may have a potential impact on P53 signaling pathway.

Additionally, as early as 2004, it has been confirmed that there was an internal association between the P53 and Wnt signaling pathways (39). In a recent study, researchers identified the Wnt signaling pathway as one of the major targets of P53 in mouse embryonic stem cells (40). In addition, another study demonstrated that P53 plays an important role in suppressing the Wnt signaling pathway by inducing DKK1 (41). Overall, DKK1 and HMGA2 interacted with each other via the Wnt and P53 signaling pathway, which may pose potential targets for the investigation of the progression and pathogenesis of tumors.

In conclusion, we integrated and analyzed multiple gene data sets by bioinformatics to investigate the biological and clinical value genes. Finally, using a series of particular conditions we screened two hub genes from 205 DEGs. These findings may improve our understanding of the etiology, pathology, and the potential molecular mechanisms and gene targets of PDAC, which may be beneficial for the identification of diagnostic biomarkers and treatment methods for PDAC. In addition, future analysis may be segregated into various divisions, including the detection of different staging tumors and metastatic tumors, further exploration of genetic mutations and specific regulatory mechanisms. Of note, the two genes of interest in the present study were statistically significant in terms of prognosis, which may assist us in the determination of prognosis and may possibly change treatment current strategies to improve prognosis. Nevertheless, lacking of experimental verification is a limitation of this study. Further molecular biological experiments *in vivo* and *in vitro* are required to confirm the function of the identified genes in PDAC.

Acknowledgements

Not applicable.

Funding

The present study was supported by the Project of Key Research and Development of Jiangsu Province of P.R. China (grant no. BE2016673), the Six Major Talent Peak Project of Jiangsu Province (grant no. 2017-WSW-059), Jiangsu Provincial Medical Youth Talent (grant no. QNRC2016734), the Project of Suzhou People's Livelihood Science and Technology (grant nos. SS201531 and SS201632).

Availability of data and materials

The datasets used and analyzed during the current study are available from GEO, TCGA and UALCAN.

Authors' contributions

YuCT, ZXZ and JZ made substantial contributions to the study conception and design. YuCT, YaCT, XYC and JZ analyzed the data and wrote the manuscript. YuCT, ZXZ, YaCT and XYC assisted with revising the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

References

1. Siegel RL, Miller KD and Jemal A: Cancer statistics, 2017. CA Cancer J Clin 67: 7-30, 2017.
2. DeSantis CE, Lin CC, Mariotto AB, Siegel RL, Stein KD, Kramer JL, Alteri R, Robbins AS and Jemal A: Cancer treatment and survivorship statistics, 2014. CA Cancer J Clin 64: 252-271, 2014.
3. Kamisawa T, Wood LD, Itoi T and Takaori K: Pancreatic cancer. Lancet 388: 73-85, 2016.
4. Wang P, Zhang J, Zhang L, Zhu Z, Fan J, Chen L, Zhuang L, Luo J, Chen H, Liu L, *et al*: MicroRNA 23b regulates autophagy associated with radioresistance of pancreatic cancer cells. Gastroenterology 145: 1133-1143.e12, 2013.
5. Adamska A, Elaskalani O, Emmanouilidi A, Kim M, Abdol Razak NB, Metharom P and Falasca M: Molecular and cellular mechanisms of chemoresistance in pancreatic cancer. Adv Biol Regul: Nov 22, 2017 (Epub ahead of print).
6. Mazarico JM, Sánchez-Arévalo Lobo VJ, Favicchio R, Greenhalf W, Costello E, Carrillo-de Santa Pau E, Marqués M, Lacal JC, Aboagye E and Real FX: Choline Kinase Alpha (CHKα) as a therapeutic target in pancreatic ductal adenocarcinoma: Expression, predictive value, and sensitivity to inhibitors. Mol Cancer Ther 15: 323-333, 2016.
7. Alhasan SF, Haugk B, Ogle LF, Beale GS, Long A, Burt AD, Tiniakos D, Televantou D, Coxon F, Newell DR, *et al*: Sulfatase-2: A prognostic biomarker and candidate therapeutic target in patients with pancreatic ductal adenocarcinoma. Br J Cancer 115: 797-804, 2016.
8. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, *et al*: NCBI GEO: Archive for functional genomics data sets-update. Nucleic Acids Res 41 (Database Issue): D991-D995, 2013.

9. Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, Shmulevich I, Sander C and Stuart JM: The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet 45: 1113-1120, 2013.
10. Cancer Genome Atlas Research Network. Electronic address: andrew_aguirre@dfci.harvard.edu; Cancer Genome Atlas Research Network: Integrated genomic characterization of pancreatic ductal adenocarcinoma. Cancer Cell 32: 185-203.e13, 2017.
11. Gautier L, Cope L, Bolstad BM and Irizarry RA: affy-analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 20: 307-315, 2004.
12. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, *et al*: Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25: 25-29, 2000.
13. Kanehisa M, Goto S, Sato Y, Furumichi M and Tanabe M: KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res 40 (Database Issue): D109-D114, 2012.
14. Pathan M, Keerthikumar S, Ang CS, Gangoda L, Quek CY, Williamson NA, Mouradov D, Sieber OM, Simpson RJ, Salim A, *et al*: FunRich: An open access standalone functional enrichment and interaction network analysis tool. Proteomics 15: 2597-2601, 2015.
15. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK and Varambally S: UALCAN: A portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia 19: 649-658, 2017.
16. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, *et al*: STRING v10: Protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 43 (Database Issue): D447-D452, 2015.
17. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T: Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Res 13: 2498-2504, 2003.
18. Calvo MB, Figueroa A, Pulido EG, Campelo RG and Aparicio LA: Potential role of sugar transporters in cancer and their relationship with anticancer therapy. Int J Endocrinol 2010: pii: 205357, 2010.
19. Jiang H, Hegde S, Knolhoff BL, Zhu Y, Herndon JM, Meyer MA, Nywening TM, Hawkins WG, Shapiro IM, Weaver DT, *et al*: Targeting focal adhesion kinase renders pancreatic cancers responsive to checkpoint immunotherapy. Nat Med 22: 851-860, 2016.
20. Morris JP IV, Wang SC and Hebrok M: KRAS, Hedgehog, Wnt and the twisted developmental biology of pancreatic ductal adenocarcinoma. Nat Rev Cancer 10: 683-695, 2010.
21. Ahn VE, Chu ML, Choi HJ, Tran D, Abo A and Weis WI: Structural basis of Wnt signaling inhibition by Dickkopf binding to LRP5/6. Dev Cell 21: 862-873, 2011.
22. Qi L, Sun B, Liu Z, Li H, Gao J and Leng X: Dickkopf-1 inhibits epithelial-mesenchymal transition of colon cancer cells and contributes to colon cancer suppression. Cancer Sci 103: 828-835, 2012.
23. Sato N, Yamabuki T, Takano A, Koinuma J, Aragaki M, Masuda K, Ishikawa N, Kohno N, Ito H, Miyamoto M, *et al*: Wnt inhibitor Dickkopf-1 as a target for passive cancer immunotherapy. Cancer Res 70: 5326-5336, 2010.
24. Zhang J, Zhang X, Zhao X, Jiang M, Gu M, Wang Z and Yue W: DKK1 promotes migration and invasion of non-small cell lung cancer via β-catenin signaling pathway. Tumour Biol 39: 1010428317703820, 2017.
25. González-Sancho JM, Aguilera O, García JM, Pendás-Franco N, Peña C, Cal S, García de Herreros A, Bonilla F and Muñoz A: The Wnt antagonist DICKKOPF-1 gene is a downstream target of beta-catenin/TCF and is downregulated in human colon cancer. Oncogene 24: 1098-1103, 2005.
26. Takahashi N, Fukushima T, Yorita K, Tanaka H, Chijiiwa K and Kataoka H: Dickkopf-1 is overexpressed in human pancreatic ductal adenocarcinoma cells and is involved in invasive growth. Int J Cancer 126: 1611-1620, 2010.
27. Shen Q, Fan J, Yang XR, Tan Y, Zhao W, Xu Y, Wang N, Niu Y, Wu Z, Zhou J, *et al*: Serum DKK1 as a protein biomarker for the diagnosis of hepatocellular carcinoma: A large-scale, multicentre study. Lancet Oncol 13: 817-826, 2012.
28. Kimura H, Fumoto K, Shojima K, Nojima S, Osugi Y, Tomihara H, Eguchi H, Shintani Y, Endo H, Inoue M, *et al*: CKAP4 is a Dickkopf1 receptor and is involved in tumor progression. J Clin Invest 126: 2689-2705, 2016.
29. Hinton J, Callan R, Bodine C, Glasgow W, Brower S, Jiang SW and Li J: Potential epigenetic biomarkers for the diagnosis and prognosis of pancreatic ductal adenocarcinomas. Expert Rev Mol Diagn 13: 431-443, 2013.
30. Are C, Chowdhury S, Ahmad H, Ravipati A, Song T, Shrikandhe S and Smith L: Predictive global trends in the incidence and mortality of pancreatic cancer based on geographic location, socio-economic status, and demographic shift. J Surg Oncol 114: 736-742, 2016.
31. Perera RM and Bardeesy N: Ready, set, go: The EGF receptor at the pancreatic cancer starting line. Cancer Cell 22: 281-282, 2012.
32. Li Y, Peng L and Seto E: Histone deacetylase 10 regulates the cell cycle G2/M phase transition via a Novel Let-7-HMGA2-Cyclin A2 pathway. Mol Cell Biol 35: 3547-3565, 2015.
33. Watanabe S, Ueda Y, Akaboshi S, Hino Y, Sekita Y and Nakao M: HMGA2 maintains oncogenic RAS-induced epithelial-mesenchymal transition in human pancreatic cancer cells. Am J Pathol 174: 854-868, 2009.
34. Strell C, Norberg KJ, Mezheyeuski A, Schnittert J, Kuninty PR, Moro CF, Paulsson J, Schultz NA, Calatayud D, Löhr JM, *et al*: Stroma-regulated HMGA2 is an independent prognostic marker in PDAC and AAC. Br J Cancer 117: 65-77, 2017.
35. Haselmann V, Kurz A, Bertsch U, Hübner S, Olempska-Müller M, Fritsch J, Häsler R, Pickl A, Fritsche H, Annewanter F, *et al*: Nuclear death receptor TRAIL-R2 inhibits maturation of let-7 and promotes proliferation of pancreatic and other tumor cells. Gastroenterology 146: 278-290, 2014.
36. Wei JJ, Wu J, Luan C, Yeldandi A, Lee P, Keh P and Liu J: HMGA2: A potential biomarker complement to P53 for detection of early-stage high-grade papillary serous carcinoma in fallopian tubes. Am J Surg Pathol 34: 18-26, 2010.
37. Mahajan A, Liu Z, Gellert L, Zou X, Yang G, Lee P, Yang X and Wei JJ: HMGA2: A biomarker significantly overexpressed in high-grade ovarian serous carcinoma. Mod Pathol 23: 673-681, 2010.
38. Italiano A, Bianchini L, Keslair F, Bonnafous S, Cardot-Leccia N, Coindre JM, Dumollard JM, Hofman P, Leroux A, Mainguené C, *et al*: HMGA2 is the partner of MDM2 in well-differentiated and dedifferentiated liposarcomas whereas CDK4 belongs to a distinct inconsistent amplicon. Int J Cancer 122: 2233-2241, 2008.
39. Iwai A, Marusawa H, Matsuzawa S, Fukushima T, Hijikata M, Reed JC, Shimotohno K and Chiba T: Siah-1L, a novel transcript variant belonging to the human Siah family of proteins, regulates beta-catenin activity in a p53-dependent manner. Oncogene 23: 7593-7600, 2004.
40. Lee KH, Li M, Michalowski AM, Zhang X, Liao H, Chen L, Xu Y, Wu X and Huang J: A genomewide study identifies the Wnt signaling pathway as a major target of p53 in murine embryonic stem cells. Proc Natl Acad Sci USA 107: 69-74, 2010.
41. Wang J, Shou J and Chen X: Dickkopf-1, an inhibitor of the Wnt signaling pathway, is induced by p53. Oncogene 19: 1843-1848, 2000.