

# Identification of novel therapeutic target genes and pathway in pancreatic cancer by integrative analysis

Bangli Hu, MD, Cheng Shi, BD, Hai-xing Jiang, PhD, Shan-yu Qin, PhD\*

## Abstract

**Background:** Gene alterations are crucial to the molecular pathogenesis of pancreatic cancer. The present study was designed to identify the potential candidate genes in the pancreatic carcinogenesis.

**Methods:** Gene Expression Omnibus database (GEO) datasets of pancreatic cancer tissue were retrieval and the differentially expressed genes (DEGs) from individual microarray data were merged. Gene Ontology (GO) enrichment analysis, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, protein–protein interaction (PPI) networks, and gene coexpression analysis were performed.

**Results:** Three GEO datasets, including 74 pancreatic cancer samples and 55 controls samples were selected. A total of 2325 DEGs were identified, including 1383 upregulated and 942 downregulated genes. The GO terms for molecular functions, biological processes, and cellular component were protein binding, small molecule metabolic process, and integral to membrane, respectively. The most significant pathway in KEGG analysis was metabolic pathways. PPI network analysis indicated that the significant hub genes including cytochrome P450, family 2, subfamily E, polypeptide 1 (CYP2E1), mitogen-activated protein kinase 3 (MAPK3), and phospholipase C, gamma 1 (PLCG1). Gene coexpression network analysis identified 4 major modules, and the potassium channel tetramerization domain containing 10 (KCTD10), kin of IRRE like (KIRREL), dipeptidyl-peptidase 10 (DPP10), and unc-80 homolog (UNC80) were the hub gene of each modules, respectively.

**Conclusion:** Our integrative analysis provides a comprehensive view of gene expression patterns associated with the pancreatic carcinogenesis.

**Abbreviations:** DEG = differentially expressed gene, GO = Gene Ontology, KEGG = Kyoto Encyclopedia of Genes and Genomes, PPI = protein–protein interaction, GEO = Gene Expression Omnibus database, KCTD10 = potassium channel tetramerization domain containing 10, KIRREL = kin of IRRE like, DPP10 = dipeptidyl-peptidase 10, UNC80 = unc-80 homolog, MAPK3 = mitogen-activated protein kinase 3, PLCG1 = phospholipase C, gamma 1, CYP2E1 = cytochrome P450, family 2, subfamily E, polypeptide 1, IGF1R = insulin-like growth factor 1 receptor.

**Keywords:** differentially expressed genes, integrative analysis, microarray data, pancreatic cancer

Editor: Mariusz Adamek.

*Authorship:* BLH and HXJ conceived and designed the study. BLH and CS analyzed the data. HXJ and SYQ supervised the project. BLH and CS wrote the manuscript with inputs from all authors. All authors reviewed the manuscript.

*Funding/support:* This work was supported by grants from the National Natural Scientific Foundation of China (31360221; 31560257).

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

*The authors have no conflicts of interest to disclose.*

*Department of Gastroenterology, the First Affiliated Hospital of Guangxi Medical University, Nanning, China.*

\* Correspondence: Shan-yu Qin, Department of Gastroenterology, the First Affiliated Hospital of Guangxi Medical University, Shuangyong Road 6, Nanning 530021, Guangxi, China (e-mail: qsy0511@163.com).

Copyright © 2017 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the Creative Commons Attribution-NoDerivatives License 4.0, which allows for redistribution, commercial and non-commercial, as long as it is passed along unchanged and in whole, with credit to the author.

Medicine (2017) 96:42(e8261)

Received: 16 March 2017 / Received in final form: 12 September 2017 /

Accepted: 18 September 2017

<http://dx.doi.org/10.1097/MD.00000000000008261>

## 1. Introduction

Pancreatic cancer is the 4th leading cause of cancer death in the world, with a median survival time less than 6 months and a dismal 5 years survival rates of 3% to 5%.<sup>[1]</sup> Despite considerable progress has been achieved, the diagnosis and management of pancreatic cancer remains challenging.<sup>[2]</sup> To date, the exact mechanisms of pancreatic cancer have not been fully elucidated, but the genes alteration during the local and systemic tumor development of many cancers has been well known. Many studies have reported that abnormal alteration of some genes were crucial to the pancreatic carcinogenesis.<sup>[3,4]</sup>

During the last decade, gene expression profiling with microarrays has developed greatly and become an important technology for identifying the genes and biological pathways that associated with various diseases. This approach is appropriated for identifying potentially useful diagnostic and prognostic biomarkers. Recently, there are studies using microarrays to identify potentially useful genes or gene signatures that are associated with pancreatic cancer.<sup>[5–7]</sup>

Although these studies have successfully identified gene expression signatures that associated with pancreatic carcinogenesis, the expression signatures identified in these studies were

not consistent with each other.<sup>[5–7]</sup> Therefore, in order to unravel the gene expression signatures associated with the pancreatic carcinogenesis, we used the microarray datasets of human pancreatic tissue available in the databases, and performed integrated analysis on their differentially expressed gene (DEGs), our results will identify the possible biologically active molecules and the potential therapeutic targets for pancreatic cancer.

## 2. Materials and methods

### 2.1. Microarray data selection

The Gene Expression Omnibus database (GEO, <http://www.ncbi.nlm.nih.gov/geo>) of NCBI was searched to identify the relevant microarray datasets. The microarray datasets are selected according to the following rules: the samples must be human pancreatic cancer tissue; the sample size must large than 5 samples; the patients did not receive special treatment, including radiotherapy and chemotherapy; and the study type of dataset is expression profiles studies. We excluded nonhuman studies, pancreatic cancer cells studies, and integrated analysis of expression profiles studies. Data were extracted from the original studies by 2 independent reviewers. The following information was extracted from each identified study: GEO accession number, sample type, platform, number of cases and controls, references, and gene expression data. Any discrepancies between reviewers were resolved by consensus or a 3rd reviewer. The study was approved by the Review Boards of the First Affiliated Hospital of Guangxi Medical University.

### 2.2. Differentially expressed genes (DEGs) analysis

The raw data in CEL files were normalized using Robust Multiarray Analysis algorithm in R Affy package. Then the datasets were assigned into 2 groups: pancreatic cancer group and control group. The DEGs between pancreatic cancer tissue and normal tissue were estimated by *t* test using limma package in R statistical software.<sup>[8]</sup> Genes exhibiting at least 2-fold changes corresponding to a false discovery rate less than 0.05 were selected as the significantly DEGs.

### 2.3. Identification of the overlap DEGs from 3 microarray datasets

The DEGs from individual microarray data were merged and the overlap DEGs of the 3 microarray datasets were identified using the *robuRankAggreg* package<sup>[9]</sup> in R statistical software. Only the overlap DEGs were used for the integrated analysis. According to a nonparametric permutation test of the *robuRankAggreg* algorithm, a list of upregulated or downregulated genes were identified based on *P* value (where threshold < 0.05) and fold change (FC) level in a given number of replicates multiplied across different microarray datasets. The cutoff *P* value was adjusted by the Benjamin–Hochberg false discovery rate.

### 2.4. Functional and pathway enrichment analyses of DEGs

In order to screen the biological processes involved in the pathogenesis of pancreatic cancer, the online software Database for Annotation, Visualization and Integrated Discovery was used to perform Gene Ontology (GO) annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis for the DEGs. These analyses were using Hypergeometric

Distribution test, and the *P* value < .05 was set as the cut-off criteria. We also constructed pathway relation network for the DEGs and identified the relationship among the pathways.

### 2.5. Analysis of protein–protein interaction (PPI) network

To determine the function of the proteins that they encoded, DEGs were imported into the PPI network constructed by using the Biological General Repository for Interaction Datasets (BioGRID) (<http://thebiogrid.org/>) in Cytoscape software (<http://www.cytoscape.org/>). The PPI network identified for the DEGs was screened at a genome-wide scale, with both end nodes having DEGs. The network construction using methods based on genomic context and structure information.<sup>[10]</sup>

### 2.6. Gene coexpression network analysis

To further identify possible genes that crucial to the pancreatic cancer, we selected DEGs the both significant expression in GO annotation and KEGG pathway using GCBI online program (<https://www.gcbi.com.cn/gclib/html/index>). Next, DEGs from the intersection of GO annotation and KEGG pathway analysis were used to construct a gene coexpression network, we mapped the DEGs to the immense database of already-known networks and screened significant gene–gene interactions using GCBI online program. The correlation between the genes in the network was determined by GO term (biological process)-based weighting. To construct a coexpression network, a correlation matrix was built by calculating pairwise Spearman Rank correlations for all pairs of expression vectors, and the analysis for modularity using the Louvain method, with a *P* value < .05 as significant.<sup>[10]</sup>

## 3. Results

### 3.1. Selection of microarray datasets

From microarray datasets retrieved in GEO of NCBI, we selected 3 microarray datasets, including GSE28735,<sup>[11]</sup> GSE32676,<sup>[12]</sup> and GSE43288<sup>[13]</sup> that meet our criteria for DEGs analysis. These 3 microarray datasets provided the gene expression profiles on the human pancreatic cancer tissue. The dataset of GSE43288 included pancreatic cancer tissue, precursor lesions, and normal control tissue, we only chose data from pancreatic cancer tissue and normal control tissue. Therefore, a total of 74 pancreatic cancer and 55 normal control sample were included our study. With regard to the used GEO platforms, all these 3 datasets used Affymetrix Human Gene Array platforms. The details of the datasets are shown in Table 1.

### 3.2. Identification of DEGs for the 3 microarray datasets

The raw data of the 3 microarray datasets were log<sub>2</sub>-transformed and normalized in order that their mean and unit variance was zero. According to *robuRankAggreg* algorithm, we identify a total of 2325 DEGs from the microarray datasets depending on their *P*-value and FC level. There are 1383 upregulated and 942 downregulated DEGs.

### 3.3. Functional annotation and KEGG pathway for DEGs

To gain insights into the biological roles of the DEGs, we performed a GO categories enrichment analysis. The GO

**Table 1****Characteristic of included microarray data.**

GSE ID	Platform	No cancer	No control	Sample
GSE28735	Affymetrix Human Gene 1.0 ST Array	45	45	Tissue
GSE32676	Affymetrix Human Genome U133 Plus 2.0 Array	25	7	Tissue
GSE43288	Affymetrix Human Genome U133A Array	4	3	Tissue

function and KEGG pathway enrichment of the total DEGs are listed in Table 2. We found that the top enriched GO terms for biological processes were: small molecule metabolic process ( $P=5.38e-47$ ), signal transduction ( $P=7.33e-34$ ), and transmembrane transport ( $P=2.65e-27$ ); for molecular function were: protein binding ( $P=4.16e-97$ ), ATP binding ( $P=8.34e-52$ ), and metal ion binding ( $P=4.93e-37$ ); for cellular component were: integral to membrane ( $P=3.45e-99$ ), plasma membrane ( $P=1.07e-93$ ), and cytoplasm ( $P=1.11e-82$ ). In the KEGG pathway enrichment analysis for the DEGs, we found that the most significant pathways in the KEGG analysis were metabolic pathways ( $P=5.83e-34$ ). Furthermore, pancreatic secretion ( $P=1.98e-22$ ) and phagosome ( $P=2.64e-15$ ) pathways were also highly enriched.

To further assess the relation of the pathways, we constructed a pathway relation network for the DEGs. This pathway relation network composed of 30 nodes and 118 edges, and we found that the MAPK signaling pathway (degree=41), apoptosis (degree=31), and pathways in cancer (degree=27) ranked the top 3 largest of degree. We also found that the genes expression was upregulated only in apoptosis, cell cycle, and Jak-STAT signaling

pathway, while the rest pathway included up- and downregulated genes (Fig. 1).

**3.4. PPI network analysis of the DEGs**

The PPI network for the total of DEGs with significant interaction relation composed of 453 nodes and 1039 edges (Fig. 2A) and a subnetwork for the top 30 DEGs is shown in Fig. 2B. As shown in Fig. 2B, the top 5 list of DEGs were determined in the order of the interacting edges, and the upregulated DEGs were: mitogen-activated protein kinase 3 (MAPK3) (degree=31), phospholipase C, gamma 1 (PLCG1) (degree=27), and phospholipase A2, group IVA (PLA2G4A), the downregulated DEGs were: cytochrome P450, family 2, subfamily E, polypeptide 1 (CYP2E1), insulin-like growth factor 1 receptor (IGF1R). The top 20 list of DEGs in the PPI network are shown in Table 3.

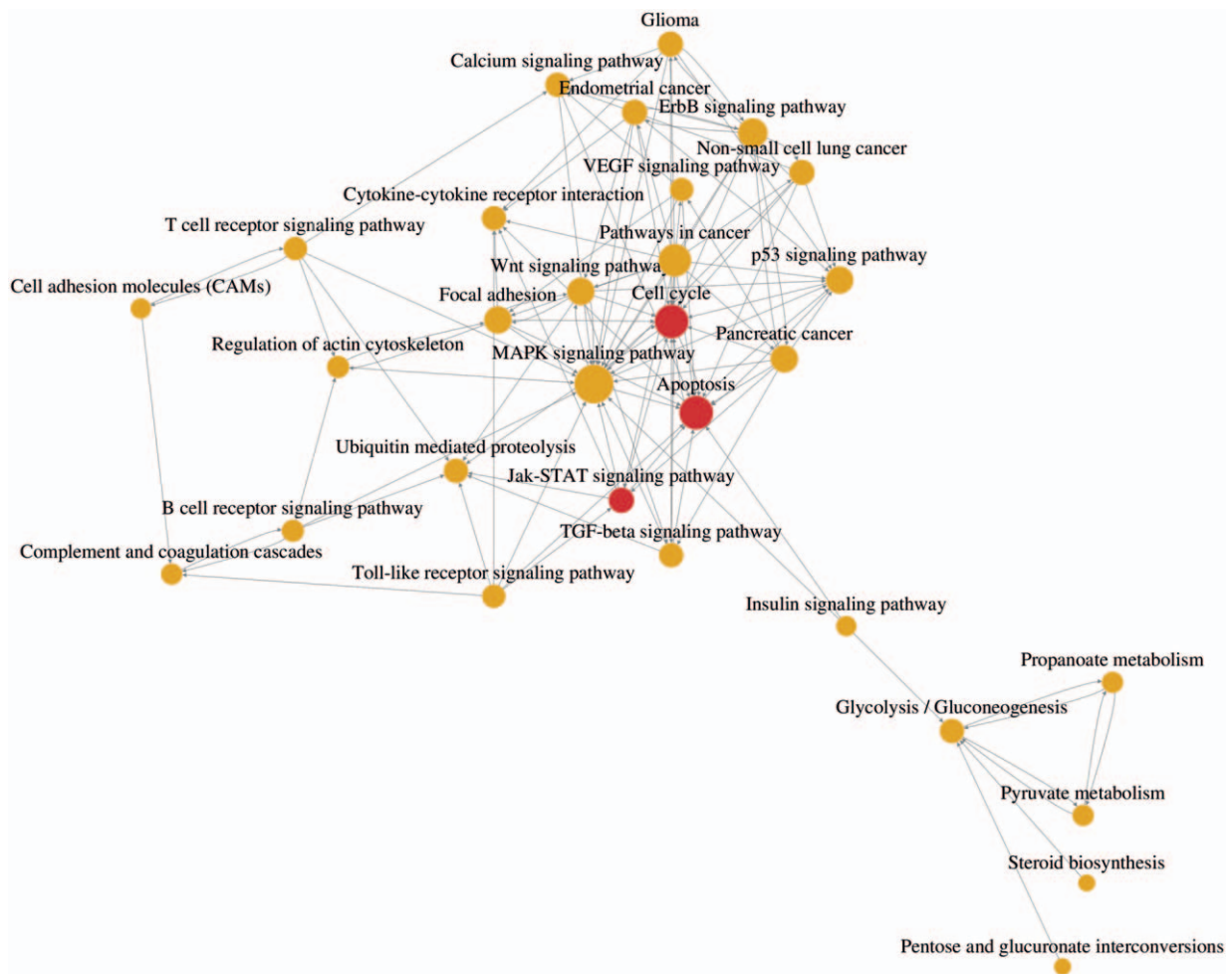
**3.5. Gene coexpression network analysis**

Gene coexpression network was constructed using the intersection DEGs from GO analysis and pathway analysis. The gene coexpression network composed of 80 nodes and 294 edges

**Table 2****Top 5 GO function and KEGG pathway enrichment of the total DEGs.**

GO ID	GO name	Diff gene counts in GO	Gene amount in GO	Enrichment score	P	FDR
	Biological processes					
GO:0044281	Small molecule metabolic process	205	1363	3.06	5.38e-47	2.13e-43
GO:0007165	Signal transduction	152	1030	3.00	7.33e-34	1.45e-30
GO:0055085	Transmembrane transport	95	538	3.59	2.65e-27	3.51e-24
GO:0000278	Mitotic cell cycle	75	363	4.21	3.81e-26	3.78e-23
GO:0045087	Innate immune response	93	554	3.42	4.469e-25	3.54e-22
	Molecular function					
GO:0005515	Protein binding	597	5116	2.37	4.16e-97	5.49e-94
GO:0005524	ATP binding	224	1482	3.08	8.34e-52	5.50e-49
GO:0046872	Metal ion binding	160	1056	3.08	4.93e-37	2.16e-34
GO:0004674	Protein serine/threonine kinase activity	61	361	3.44	5.89e-17	1.94e-14
GO:0008270	Zinc ion binding	181	1991	1.85	2.27e-15	5.99e-13
	Cellular component					
GO:0016021	Integral to membrane	536	4240	2.57	3.45e-99	2.0e-96
GO:0005886	Plasma membrane	468	3498	2.72	1.07e-93	3.14e-91
GO:0005737	Cytoplasm	506	4269	2.412	1.11e-82	2.18e-80
GO:0005829	Cytosol	359	2457	2.97	1.28e-80	1.88e-78
GO:0005634	Nucleus	486	4991	1.98	1.26e-51	1.48e-49
	KEGG					
Pathway ID	Pathway name					
01100	Metabolic pathways	166	1189	5.83e-34	1.45e-31	1.46e-31
04972	Pancreatic secretion	36	96	1.98e-22	2.47e-20	2.48e-20
04145	Phagosome	37	158	2.64e-15	2.20e-13	2.20e-13
04380	Osteoclast differentiation	34	135	3.58e-15	2.24e-13	2.24e-13
04151	PI3K-Akt Signaling pathway	55	347	3.48e-14	1.67e-12	1.67e-12

DEG = differentially expressed gene, FDR = false discovery rate, GO = Gene Ontology, KEGG = Kyoto Encyclopedia of Genes and Genomes.



**Figure 1.** Pathway relation network for the total differentially expressed genes (DEGs). Expression levels are represented by red (upregulated), green (downregulated), and yellow (up- and downregulated).

(Fig. 3A). The top 20 DEGs of intersection are listed in Table 4. In the coexpression network, we found that the network can be divided into 4 major modules, the potassium channel tetramerization domain containing 10 (KCTD10) (degree=31) and kin of IRRE like (KIRREL) (degree=26) were the hub upregulated gene of the modules, and the dipeptidyl-peptidase 10 (DPP10) (degree=33) and unc-80 homolog (UNC80) (degree=27) were the hub downregulated gene of the modules (Fig. 3B).

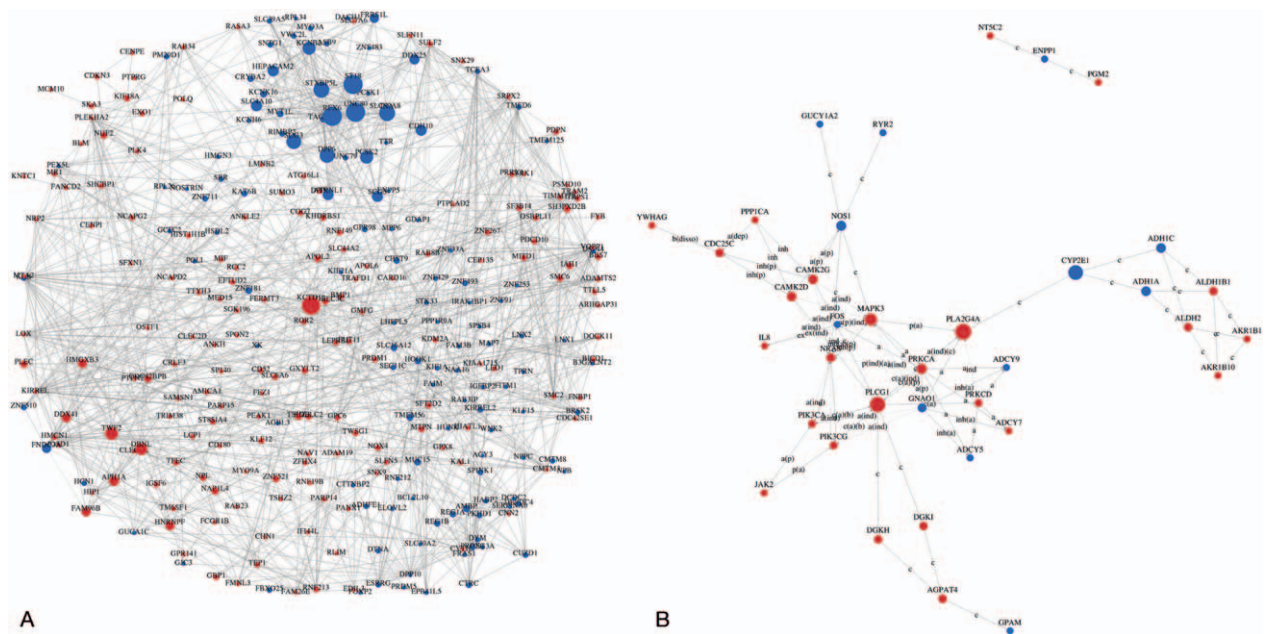
#### 4. Discussion

The development of pancreatic cancer is a complicated process and involving numerous alterations of genes. A comprehensive analysis of the molecular mechanism underlying pancreatic carcinogenesis is crucial for the management strategy of pancreatic cancer. To date, several studies reported the results of gene expression signatures in pancreatic cancer using microarray profiling, but some of them with less samples and the results were inconsistent. Therefore, combining data from multiple existing studies can increase the reliability and generalizability of results. Integrative analysis is an approach of gaining new insights on microarray data. In this study, we performed an integrative

analysis to understand the mechanism using 3 microarray datasets of pancreatic cancer tissue.

In the DEGs expression analysis, we found that there were 1383 upregulated and 942 downregulated DEGs after combing the data from the 3 datasets. Among these DEGs, MAPK3, CYP2E1, KCTD10, KIRREL, DPP10, and UNC80 were found to be the hub genes in the PPI analysis and coexpression. Other top DEGs, such as PLCG1, PLA2G4A, and IGF1R, were also implicated in the pathogenesis of several cancers. PLCG1 is a downstream effector signaling molecule for fibroblast growth factor receptor 1, and a study has shown that sonic hedgehog signaling promotes gastric cancer proliferation through induction of PLCG1-extracellular regulated protein kinases 1/2.<sup>[14]</sup> PLA2G4A is an enzyme that implicated in cancer cells proliferation. Targeting of PLA2G4A impedes cell cycle re-entry of quiescent prostate cancer cells.<sup>[15]</sup> Agents targeting the IGF1R have shown antitumor activity. Cyclin dependent kinases 4/6 and IGF1 inhibitors synergized to suppress the growth of p16<sup>INK4A</sup>-deficient pancreatic cancers.<sup>[16]</sup>

In the GO analysis, the significant GO terms were related to small molecule metabolic process, protein binding, and integral to membrane of the host cell. Like other cancers, it has been known that the pancreatic carcinogenesis involving

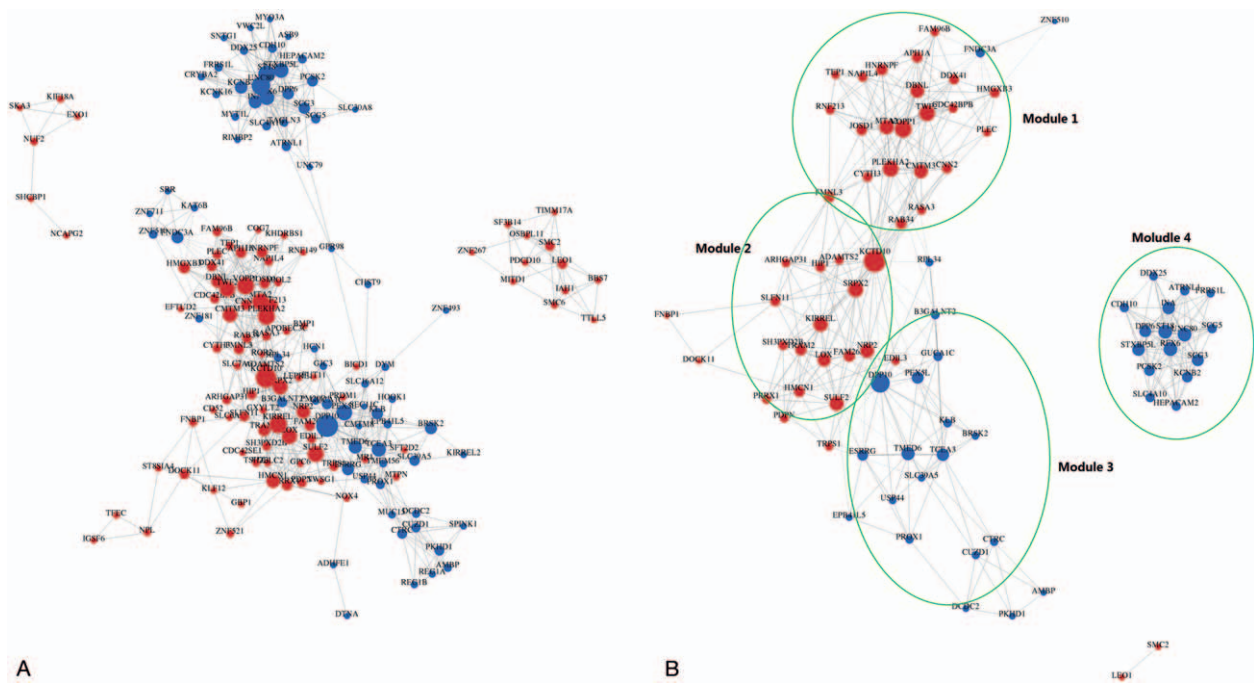


**Figure 2.** PPI network for the DEGs. (A) PPI network for the total DEGs; (B) PI network for the top 30 DEGs. The red nodes represented upregulated of genes, the blue nodes represented downregulated of genes. DEG=differentially expressed gene, PPI=protein–protein interaction.

**Table 3**  
**Top 20 DEGs from the PPI network.**

Gene symbol	Gene description	Gene feature	Betweenness	Degree
CYP2E1	"Homo sapiens cytochrome P450, family 2, subfamily E, polypeptide 1 (CYP2E1), mRNA."	Down	9115.763	40
MAPK3	"Homo sapiens mitogen-activated protein kinase 3 (MAPK3), transcript variant 1, mRNA."	UP	18793.82	31
PLCG1	"Homo sapiens phospholipase C, gamma 1 (PLCG1), transcript variant 2, mRNA."	Up	19583.64	27
PLA2G4A	"Homo sapiens phospholipase A2, group IVA (cytosolic, calcium-dependent) (PLA2G4A), mRNA."	Up	14072.87	27
IGF1R	"Homo sapiens insulin-like growth factor 1 receptor (IGF1R), mRNA."	Down	7643.3	26
PRKCA	"Homo sapiens protein kinase C, alpha (PRKCA), mRNA."	Up	17565.41	24
NRAS	"Homo sapiens neuroblastoma RAS viral (v-ras) oncogene homolog (NRAS), mRNA."	Up	9304.432	22
PLCE1	cdna:known chromosome:GRCh37:10:95790569:96088146:1 gene: ENSG00000138193 gene_biotype:protein_coding	Down	7107.665	17
DGKH	"Homo sapiens diacylglycerol kinase, eta (DGKH), transcript variant 3, mRNA."	Up	8866.62	16
DGKI	"Homo sapiens diacylglycerol kinase, iota (DGKI), mRNA."	Up	8866.62	16
ENPP1	"Homo sapiens ectonucleotide pyrophosphatase/phosphodiesterase 1 (ENPP1), mRNA."	Down	8961.871	14
ALDH1B1	"Homo sapiens aldehyde dehydrogenase 1 family, member B1 (ALDH1B1), nuclear gene encoding mitochondrial protein, mRNA."	Up	8038.822	14
ALDH2	"Homo sapiens aldehyde dehydrogenase 2 family (mitochondrial) (ALDH2), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA."	Up	8038.822	14
NOS1	"Homo sapiens nitric oxide synthase 1 (neuronal) (NOS1), transcript variant 1, mRNA."	Down	14193.11	12
AKR1B1	"Homo sapiens aldo-keto reductase family 1, member B1 (aldose reductase) (AKR1B1), mRNA."	Up	7231.052	10
AKR1B10	"Homo sapiens aldo-keto reductase family 1, member B10 (aldose reductase) (AKR1B10), mRNA."	Up	7231.052	10
AGPAT4	"Homo sapiens 1-acylglycerol-3-phosphate O-acyltransferase 4 (lysosphosphatidic acid acyltransferase, delta) (AGPAT4), mRNA."	Up	13756.43	8
GPAM	"Homo sapiens glycerol-3-phosphate acyltransferase, mitochondrial (GPAM), nuclear gene encoding mitochondrial protein, transcript variant 2, mRNA."	Down	13608.41	8
GK	"Homo sapiens glycerol kinase (GK), transcript variant 3, mRNA."	Up	12596.68	8
RYR2	"Homo sapiens ryanodine receptor 2 (cardiac) (RYR2), mRNA."	Down	7288.727	6
FIGF	"Homo sapiens c-fos induced growth factor (vascular endothelial growth factor D) (FIGF), mRNA."	Down	6783.978	5

DEG = differentially expressed gene, PPI = protein–protein interaction.



**Figure 3.** Gene coexpression network for the differentially expressed genes (DEGs). (A) Gene coexpression network for the total DEGs; (B) gene coexpression network for the top 20 DEGs. The red nodes represented upregulated genes, the blue nodes represented downregulated genes.

many small molecules changed, the genes alteration subsequently leads to the change of the corresponding proteins. Abnormal expression of proteins interacting selectively with other proteins or form protein complex, which is a vital mechanism of cancer development.<sup>[17,18]</sup> Besides, the abnormal expression protein integrals to membrane of the host cells, then results in the change of morphology and function of host cells, finally leads to the degeneration of normal cells.<sup>[19]</sup>

With regard to the signaling pathway, we found that metabolic pathways, pancreatic secretion, and phagosome pathways were highly enriched, indicating that the DEGs of the microarray data were largely involving in these signaling pathways. Besides, in the pathway relation analysis, we found that the downregulated pathways: MAPK signaling pathway, apoptosis, and pathways in cancer; and the upregulated pathways: apoptosis, cell cycle, and Jak-STAT signaling pathway showed the most relationship with

**Table 4**

**Top 20 differentially expressed genes (DEGs) for gene coexpression network.**

Gene symbol	Description	Degree	Gene feature
DPP10	"Homo sapiens dipeptidyl-peptidase 10 (non-functional) (DPP10), transcript variant 1, mRNA."	33	Down
KCTD10	"Homo sapiens potassium channel tetramerisation domain containing 10 (KCTD10), mRNA."	31	Up
UNC80	"Homo sapiens unc-80 homolog (C. elegans) (UNC80), transcript variant 2, mRNA."	27	Down
KIRREL	cdna:putative chromosome:GRCh37:1:158056470:158070052:1 gene:ENSG00000183853 gene_biotype:protein_coding transcript_biotype:protein_coding	26	Up
PLEKHA2	"Homo sapiens pleckstrin homology domain containing, family A (phosphoinositide binding specific) member 2 (PLEKHA2), mRNA."	25	Up
VOPP1	"Homo sapiens vesicular, overexpressed in cancer, prosurvival protein 1 (VOPP1), mRNA."	25	Up
MTA2	"Homo sapiens metastasis associated 1 family, member 2 (MTA2), mRNA."	24	Up
ST18	"Homo sapiens suppression of tumorigenicity 18 (breast carcinoma) (zinc finger protein) (ST18), mRNA."	24	Down
SULF2	"Homo sapiens sulfatase 2 (SULF2), transcript variant 2, mRNA."	23	Up
PEX5L	"Homo sapiens peroxisomal biogenesis factor 5-like (PEX5L), transcript variant 1, mRNA."	22	Down
TWF2	"Homo sapiens twinfilin, actin-binding protein, homolog 2 (Drosophila) (TWF2), mRNA."	22	Up
LOX	"Homo sapiens lysyl oxidase (LOX), transcript variant 1, mRNA."	21	Up
SRPX2	"Homo sapiens sushi-repeat containing protein, X-linked 2 (SRPX2), mRNA."	21	Up
CMTM3	"Homo sapiens CKLF-like MARVEL transmembrane domain containing 3 (CMTM3), transcript variant 1, mRNA."	20	Up
RFX6	"Homo sapiens regulatory factor X, 6 (RFX6), mRNA."	20	Down
STXBP5L	"Homo sapiens syntaxin binding protein 5-like (STXBP5L), mRNA."	20	Down
TCEA3	"Homo sapiens transcription elongation factor A (SII), 3 (TCEA3), mRNA."	19	Down
HMCN1	"Homo sapiens hemicentin 1 (HMCN1), mRNA."	18	Up
NRP2	"Homo sapiens neuropilin 2 (NRP2), transcript variant 2, mRNA."	18	Up
INA	"Homo sapiens internexin neuronal intermediate filament protein, alpha (INA), mRNA."	17	Down

other pathways. In addition, as shown in Fig. 1, we also found that most of the pathways were downregulated during the pancreatic carcinogenesis. The role of MAPK and Jak-STAT signaling pathways have been well defined in the pancreatic cancer,<sup>[20,21]</sup> and the cell apoptosis and cell cycle are the important process of normal cells develop into cancer cells.<sup>[22]</sup> These results provided more evidences of the important signaling pathways interaction during the process of pancreatic carcinogenesis.

In PPI network analysis, the results indicated that the significant hub gene (MAPK3 and CYP2E1) and their related pathways played key roles in the pancreatic carcinogenesis. MAPK directly interacts with scaffolding proteins, activators, and effectors, and these interactions also govern signaling specificity. The MAPK/ERK1/2 signaling pathway, which has been correlated with malignant carcinoma carcinogenesis, plays crucial roles in cell growth control, differentiation, proliferation, and apoptosis.<sup>[23]</sup> It has been suggested that a novel compound capable of inhibiting MAPK/ERK1/2 activity would be a good potential candidate for pancreatic cancer.<sup>[24]</sup> CYP2E1 is a member of the cytochrome P450 superfamily and is involved in the metabolic activation of many carcinogens, such as gastric cancer<sup>[25]</sup> and lung cancer.<sup>[26]</sup> These PPI network results demonstrated the crucial functional linkages between the DEGs in the pancreatic carcinogenesis.

Gene-gene interactions is an important process in the regulation of disease pathogenesis. In a bioinformatics study, the coexpression network analysis would detect several gene-gene interactions modules from the overall network, while these coexpression genes in the modules could not be easily detected and clustered by the differential expression analysis.<sup>[27]</sup> In general, the coexpressed genes tend to involve in the similar biological process, such as pathways, complexes, and those genes in a strong coexpression module (calculated by the degree value) present higher similarity of biological function than those from random gene pairs analysis.<sup>[28–30]</sup> In this study, we identified 4 major gene modules, and KCTD10, KIRREL, DPP10, and UNC80 were the hub gene of each modules, indicating that these genes are key to the pancreatic carcinogenesis.

KCTD10 belongs to the polymerase delta-interacting protein 1 gene family and is implicated in some tumors. Wang et al<sup>[31]</sup> observed that the expression of proliferating cell nuclear antigen was decreased by knockdown of KCTD10 in lung adenocarcinoma cells, and the cell proliferation was also inhibited. KCTD10 can be regulated by some transcription factors, including transcription factor specificity protein 1 and activating protein 2alpha, which regulated KCTD10 mRNA expression positively and negatively, respectively.<sup>[32]</sup> Recently, a study observed that KCTD10 was regulated by E twenty-six variant 1, and silencing of KCTD10 could increase gastrointestinal stromal tumor cell proliferation and invasion, suggesting that KCTD10 was vital to the gastrointestinal stromal tumor pathogenesis.<sup>[33]</sup>

KIRREL (also known as NEPH1) is a member of the nephrin-like protein family. KIRREL induces actin polymerization by transduction outside-in signals and recruitment of Grb2.<sup>[34]</sup> KIRREL and nephrin can form a functional receptor complex during podocyte development and following podocyte injury,<sup>[35]</sup> indicating that KIRREL is associated with these physiological or pathologic process.

DPP10 is an inactive peptidase that modulates the electrophysiological properties, cell-surface expression, and subcellular localization of voltage-gated potassium channels. DPP10 is restricted expression, including the brain, adrenal gland, and

pancreas, may serve as a marker in certain malignant states such as colorectal cancer and could have a prognostic significance.<sup>[36]</sup> DPP10 has been linked to asthma susceptibility by several genome-wide association studies<sup>[37]</sup> and recently impaired expressions of DPP10 gene in malignant mesothelioma<sup>[38]</sup> has been reported.

UNC80 is a component of the NALCN Na<sup>+</sup> leak cation channel, which is a voltage-independent ion-channel complex. In heterologous expression systems, UNC80 and NALCN have been shown to interact.<sup>[39]</sup> UNC79 and UNC80 also associate with each other, and UNC79 requires the presence of UNC80 to associate with NALCN. NALCN associates with UNC79 via UNC80 in the brain, and UNC79 influences UNC80 protein levels. Mutations in Nalcn, Unc79, or Unc80 lead to severe phenotypes that include neonatal lethality and disruption in rhythmic behaviors.<sup>[40]</sup>

In summary, the present integrated analysis provided a comprehensive perspective to understand the molecular mechanism underlying pancreatic carcinogenesis, and identified some hub genes and pathways. The hub genes and pathways may be potential targets of treatment for pancreatic cancer. However, further investigations are remained necessary for unraveling the mechanism in the pathogenesis of pancreatic cancer.

## References

- [1] Jemal A, Siegel R, Xu J, et al. Cancer statistics, 2010. *CA Cancer J Clin* 2010;60:277–300.
- [2] Goggins M. Molecular markers of early pancreatic cancer. *J Clin Oncol* 2005;23:4524–31.
- [3] Lu P, Weaver VM, Werb Z. The extracellular matrix: a dynamic niche in cancer progression. *J Cell Biol* 2012;196:395–406.
- [4] Muller PA, Vousden KH, Norman JC. p53 and its mutants in tumor cell migration and invasion. *J Cell Biol* 2011;192:209–18.
- [5] Campagna D, Cope L, Lakkur SS, et al. Gene expression profiles associated with advanced pancreatic cancer. *Int J Clin Exp Pathol* 2008;1:32–43.
- [6] Goggins M. Identifying molecular markers for the early detection of pancreatic neoplasia. *Semin Oncol* 2007;34:303–10.
- [7] Tan ZJ, Hu XG, Cao GS, et al. Analysis of gene expression profile of pancreatic carcinoma using cDNA microarray. *World J Gastroenterol* 2003;9:818–23.
- [8] Diboun I, Wernisch L, Orengo CA, et al. Microarray analysis after RNA amplification can detect pronounced differences in gene expression using limma. *BMC Genomics* 2006;7:252.
- [9] Kolde R, Laur S, Adler P, et al. Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics* 2012;28:573–80.
- [10] Reichardt J, Bornholdt S. Partitioning and modularity of graphs with arbitrary degree distribution. *Phys Rev E Stat Nonlin Soft Matter Phys* 2007;76(1 Pt 2):015102.
- [11] Zhang G, Schetter A, He P, et al. DPEP1 inhibits tumor cell invasiveness, enhances chemosensitivity and predicts clinical outcome in pancreatic ductal adenocarcinoma. *PLoS One* 2012;7:e31507.
- [12] Donahue TR, Tran LM, Hill R, et al. Integrative survival-based molecular profiling of human pancreatic cancer. *Clin Cancer Res* 2012;18:1352–63.
- [13] Crnogorac-Jurcevic T, Chelala C, Barry S, et al. Molecular analysis of precursor lesions in familial pancreatic cancer. *PLoS One* 2013;8:e54830.
- [14] Ertao Z, Jianhui C, Chuangqi C, et al. Autocrine Sonic hedgehog signaling promotes gastric cancer proliferation through induction of phospholipase Cgamma1 and the ERK1/2 pathway. *J Exp Clin Cancer Res* 2016;35:63.
- [15] Yao M, Xie C, Kiang MY, et al. Targeting of cytosolic phospholipase A2alpha impedes cell cycle re-entry of quiescent prostate cancer cells. *Oncotarget* 2015;6:34458–74.
- [16] Heilmann AM, Perera RM, Ecker V, et al. CDK4/6 and IGF1 receptor inhibitors synergize to suppress the growth of p16INK4A-deficient pancreatic cancers. *Cancer Res* 2014;74:3947–58.
- [17] Ding NH, Li JJ, Sun LQ. Molecular mechanisms and treatment of radiation-induced lung fibrosis. *Curr Drug Targets* 2013;14:1347–56.

- [18] Han Y, Zhang Y, Jia T, et al. Molecular mechanism underlying the tumor-promoting functions of carcinoma-associated fibroblasts. *Tumour Biol* 2015;36:1385–94.
- [19] Fukui Y. Mechanisms behind signet ring cell carcinoma formation. *Biochem Biophys Res Commun* 2014;450:1231–3.
- [20] Furukawa T. Impacts of activation of the mitogen-activated protein kinase pathway in pancreatic cancer. *Front Oncol* 2015;5:23.
- [21] Quintas-Cardama A, Verstovsek S. Molecular pathways: Jak/STAT pathway: mutations, inhibitors, and resistance. *Clin Cancer Res* 2013;19:1933–40.
- [22] Arlt A, Muerkoster SS, Schafer H. Targeting apoptosis pathways in pancreatic cancer. *Cancer Lett* 2013;332:346–58.
- [23] Sosa MS, Avivar-Valderas A, Bragado P, et al. ERK1/2 and p38alpha/beta signaling in tumor cell quiescence: opportunities to control dormant residual disease. *Clin Cancer Res* 2011;17:5850–7.
- [24] Hu Y, Yang H, Lu XQ, et al. ARHI suppresses pancreatic cancer by regulating MAPK/ERK 1/2 pathway. *Pancreas* 2015;44:342–3.
- [25] Elingarami S, Liu H, Kalinjuma AV, et al. Polymorphisms in NEIL-2, APE-1 CYP2E1 and MDM2 Genes are Independent Predictors of Gastric Cancer Risk in a Northern Jiangsu Population (China). *J Nanosci Nanotechnol* 2015;15:4815–28.
- [26] Zhai W, Feng R, Yang H, et al. Note of clarification of data on the association between CYP2E1 RsaI polymorphism and lung cancer risk. *Eur J Cancer* 2015;51:1650–5.
- [27] Elo LL, Jarvenpaa H, Oresic M, et al. Systematic construction of gene coexpression networks with applications to human T helper cell differentiation process. *Bioinformatics* 2007;23:2096–103.
- [28] Fuller TF, Ghazalpour A, Aten JE, et al. Weighted gene coexpression network analysis strategies applied to mouse weight. *Mamm Genome* 2007;18:463–72.
- [29] Ghazalpour A, Doss S, Zhang B, et al. Integrating genetic and network analysis to characterize genes related to mouse weight. *PLoS Genet* 2006;2:e130.
- [30] Huang Y, Li H, Hu H, et al. Systematic discovery of functional modules and context-specific functional annotation of human genome. *Bioinformatics* 2007;23:i222–9.
- [31] Wang Y, Zheng Y, Luo F, et al. KCTD10 interacts with proliferating cell nuclear antigen and its down-regulation could inhibit cell proliferation. *J Cell Biochem* 2009;106:409–13.
- [32] Liu R, Zhou A, Ren D, et al. Transcription factor specificity protein 1 (SP1) and activating protein 2alpha (AP-2alpha) regulate expression of human KCTD10 gene by binding to proximal region of promoter. *FEBS J* 2009;276:1114–24.
- [33] Kubota D, Yoshida A, Tsuda H, et al. Gene expression network analysis of ETV1 reveals KCTD10 as a novel prognostic biomarker in gastrointestinal stromal tumor. *PLoS One* 2013;8:e73896.
- [34] Harita Y, Kurihara H, Kosako H, et al. Neph1, a component of the kidney slit diaphragm, is tyrosine-phosphorylated by the Src family tyrosine kinase and modulates intracellular signaling by binding to Grb2. *J Biol Chem* 2008;283:9177–86.
- [35] Garg P, Verma R, Nihalani D, et al. Neph1 cooperates with nephrin to transduce a signal that induces actin polymerization. *Mol Cell Biol* 2007;27:8698–712.
- [36] Park HS, Yeo HY, Chang HJ, et al. Dipeptidyl peptidase 10, a novel prognostic marker in colorectal cancer. *Yonsei Med J* 2013;54:1362–9.
- [37] Michel S, Liang L, Depner M, et al. Unifying candidate gene and GWAS Approaches in Asthma. *PLoS One* 2010;5:e13894.
- [38] Bueno R, De Rienzo A, Dong L, et al. Second generation sequencing of the mesothelioma tumor genome. *PLoS One* 2010;5:e10612.
- [39] Lu B, Zhang Q, Wang H, et al. Extracellular calcium controls background current and neuronal excitability via an UNC79-UNC80-NALCN cation channel complex. *Neuron* 2010;68:488–99.
- [40] Ren D. Sodium leak channels in neuronal excitability and rhythmic behaviors. *Neuron* 2011;72:899–911.