Key genes associated with pancreatic cancer and their association with outcomes: A bioinformatics analysis

JIAJIA WU[1*], ZEDONG LI[2*], KAI ZENG[1], KANGJIAN WU[1], DONG XU[3], JUN ZHOU[2] and LIJIAN XU[1]

[1]Department of General Surgery, The Second Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu 210000; [2]Department of Minimally Invasive Surgery, The Second Xiangya Hospital, Central South University, Changsha, Hunan 410011; [3]Department of General Surgery, Gaochun People's Hospital, Nanjing, Jiangsu 211300, P.R. China

Received October 4, 2018; Accepted April 9, 2019

DOI: 10.3892/mmr.2019.10321

Abstract. Pancreatic cancer is a highly malignant neoplastic disease of the digestive system. In the present study, the dataset GSE62165 was downloaded from the Gene Expression Omnibus (GEO) database. GSE62165 contained the data of 118 pancreatic ductal adenocarcinoma samples (38 early-stage tumors, 62 lymph node metastases and 18 advanced tumors) and 13 control samples. Differences in the expression levels of genes between normal tissues and early-stage tumors were investigated. A total of 240 differentially expressed genes (DEGs) were identified using R software 3.5 (137 upregulated genes and 103 downregulated genes). Then, the differentially expressed genes were subjected to Gene Ontology and Kyoto Encyclopedia of Genes and Genomes analysis. The following 18 core genes were identified using Cytoscape, based on the protein-interaction network of DEGs determined using the online tool STRING: *EGF, ALB, COL17A1, FN1, TIMP1, PLAU, PLA2G1B, IGFBP3, PLAUR, VCAN, COL1A1, PNLIP, CTRL, PRSS3, COMP, CPB1, ITGA2* and *CEL*. The pathways of the core genes were primarily associated with pancreatic secretion, protein digestion and absorption, and focal adhesion. Finally, survival analyses of core genes in pancreatic cancer were conducted using the UALCAN online database. It was revealed that *PLAU* and *COL17A1* were significantly associated with poor prognosis ($P<0.05$). The expression levels of genes in primary pancreatic cancer tissues were then compared; only one gene, *COL17A1*, was identified to be significantly differentially expressed. Finally, another dataset from GEO, GSE28735, was analyzed to verify the upregulated expression of *COL17A1*. Taken together, the results of the present study have indicated that the expression of *COL17A1* gene may be associated with the occurrence and development of pancreatic cancer.

Correspondence to: Professor Lijian Xu, Department of General Surgery, The Second Affiliated Hospital of Nanjing Medical University, 121 Jiangjiayuan Road, Gulou, Nanjing, Jiangsu 210000, P.R. China
E-mail: xulijian134@126.com

*Contributed equally

Key words: pancreatic cancer, bioinformatics analysis, prognosis

Introduction

Pancreatic cancer is a highly malignant neoplasm of the digestive system that accounts for >200,000 deaths/year globally (1). The incidence of pancreatic cancer is low compared with that of lung, breast, colorectal and gastric cancers; however, it is associated with a very high mortality rate. It has been reported that the incidence of pancreatic cancer is very similar to the associated mortality rate; the reported 5-year survival rate of patients with pancreatic cancer is <6% (2). The mortality rate of patients with pancreatic cancer ranks fourth among common cancers, and is predicted to rise to second within a decade (3). A number of factors have been identified as contributing to the etiopathogenesis of pancreatic cancer, including heredity, smoking, high-fat diet, chronic pancreatitis and consumption of nitrous acid compounds (4). Due to the latency of pancreatic cancer, the majority of patients are diagnosed at an advanced stage, when tumor tissue has already infiltrated the surrounding tissues and has formed distant metastases, decreasing the usefulness of surgical interventions (5). As a result of drug resistance, the efficacy of postoperative adjuvant therapy has also been very unsatisfactory (6). Carbohydrate antigen 19-9 (CA19-9) is the most frequently used marker for the clinical diagnosis of cancer; the reported sensitivity and specificity of CA19-9 for the diagnosis of pancreatic cancer is 69-93 and 46-98%, respectively (7). Therefore, early diagnosis and treatment are important to improve the prognosis and survival of patients with pancreatic cancer.

At present, high-throughput sequencing is employed in a variety of contexts, such as the discovery of gene mutations and chromosomal translocations that are closely associated with the occurrence and development of tumors (8-10). High-throughput sequencing may be useful for the diagnosis of cancer and development of targeted therapies. These analyses may provide novel insights to guide subsequent research.

Materials and methods

Microarray data. The gene expression profile of GSE62165 (11) was downloaded from the GEO database (12). The data were

created using the GPL13667 Affymetrix® Human Genome U219 array (Affymetrix; Thermo Fisher Scientific, Inc.). GSE62165 contained data on 118 pancreatic ductal adenocarcinoma (PDAC) samples and 13 control samples. Data were standardized using the robust multi-array average (RMA) algorithm using limma package (version 3.38.3) (13). In addition, a separate dataset, GSE28735 (14,15), was used to verify the results. The expression profiles included 45 matched pairs of pancreatic tumor and adjacent non-tumor tissues from 45 patients with PDAC. The Cancer Genome Atlas (TCGA; http://cancergenome.nih.gov/) contains genomic sequencing data involving 33 species of cancer.

Identification of differentially expressed genes (DEGs). The limma package (version 3.38.3) (13) was used to identify DEGs between pancreatic cancer tissue and normal pancreatic tissue samples in R software (version 3.5; https://www.R-project.org). |log2 Fold Change (FC)|>3.0 and adjusted P-value <0.05 were considered to be the threshold for differential gene identification.

Gene Ontology (GO) and Kyoto Encyclopedia of genes and genomes (KEGG) pathway analysis of DEGs. GO (http://www.geneontology.org/) and the KEGG (https://www.kegg.jp/) (16-19) were used to analyze the function of DEGs using the cluster Profiler R package (20). P<0.05 was considered to indicate a statistically significant difference in functional enrichment analysis.

Core genes screening from the protein-protein interaction (PPI) network. A PPI network for the DEGs was generated using the STRING database (https://string-db.org/). Then, Cytoscape (version 3.6.1) (21) was employed, and a plug-in termed cytohubba (22) was integrated into the software. The plug-in provides 12 types of topological analysis methods [Maximal Clique Centrality, Maximum Neighborhood Component (MNC), Density of MNC, Degree, Edge Percolated Component, Bottleneck, EcCentricity, Closeness, Radiality, Betweenness, Stress and Clustering Coefficient). Using 12 analysis methods, we identified the top 18 genes as core genes.

Expression levels and survival analysis of core genes in pancreatic cancer. UALCAN (http://ualcan.path.uab.edu/index.html) (23) was employed to perform survival analysis based on the information of TCGA database. Survival analysis was performed via the Kaplan-Meier method using 18 identified core genes, based on their core gene expression levels in pancreatic adenocarcinoma (PAAD). P<0.05 was considered to be statistically significant. The P-value was calculated using log-rank test. The 'scaled_estimate' column provided the potential transcripts produced by each gene. The 'scaled_estimate' was multiplied by 10^6 to obtain a transcripts per million (TPM) expression value (24). Gene expression levels in tumor tissues exhibited notable inter-individual variability. High expression indicated that the TPM value was above the upper quartile value. Low expression indicated that the TPM value was equal or below the upper quartile value.

Verification of results. The findings from the bioinformatics analyses were validated using the dataset GSE28735 from the GEO database. The expression profiles included 45 matched pairs of pancreatic tumor and adjacent non-tumor tissues from 45 patients with PDAC. The online analysis tool

Table I. Top 20 differentially expressed genes in early-stage pancreatic cancer tissues based on Log2FC.

A, Upregulated genes

Gene symbol	Log2FC	Adjusted P-value
COL1A1	5.9240335	$3.17x10^{-16}$
KRT17	5.5334643	$4.10x10^{-12}$
CEACAM5	5.3990511	$1.66x10^{-09}$
S100P	5.2665291	$1.66x10^{-13}$
COL10A1	5.2258147	$1.17x10^{-17}$
SERPINB5	5.1642588	$5.50x10^{-15}$
GJB2	5.0747799	$7.98x10^{-18}$
COL17A1	5.0501325	$1.80x10^{-11}$
CXCL5	5.0384043	$1.28x10^{-11}$
TMPRSS4	5.0203823	$2.37x10^{-16}$
SDR16C5	4.9961337	$3.16x10^{-16}$
CTHRC1	4.9626426	$7.77x10^{-20}$
COL11A1	4.9350078	$6.41x10^{-17}$
SLC6A14	4.8841916	$2.47x10^{-15}$
MMP11	4.8824426	$3.14x10^{-16}$
SULF1	4.721966	$2.96x10^{-17}$
FN1	4.6424864	$2.99x10^{-16}$
POSTN	4.6415794	$1.33x10^{-16}$
CCL18	4.5489901	$1.41x10^{-11}$
MUC4	4.5022059	$1.25x10^{-09}$

B, Downregulated genes

Gene symbol	Log2FC	Adjusted P-value
SYCN	-6.6535946	$2.74x10^{-07}$
SERPINI2	-6.2894352	$8.73x10^{-09}$
AQP8	-6.2356139	$2.11x10^{-10}$
AMY1A	-6.1790263	$4.38x10^{-10}$
ALB	-6.1165814	$2.16x10^{-08}$
CELA2A	-6.0845313	$2.83x10^{-06}$
PNLIPRP1	-6.0676353	$6.52x10^{-10}$
CTRL	-5.9224624	$1.73x10^{-06}$
PDIA2	-5.9185276	$4.65x10^{-09}$
CPA1	-5.804844	$5.21x10^{-06}$
TMED6	-5.7967792	$2.37x10^{-10}$
CELP	-5.7183603	$1.15x10^{-10}$
AQP12A	-5.6598636	$3.88x10^{-14}$
CUZD1	-5.5766969	$1.68x10^{-06}$
CELA2B	-5.5649112	$2.23x10^{-05}$
CPA2	-5.55513	$3.43x10^{-06}$
CELA3A	-5.5508171	$1.84x10^{-05}$
GP2	-5.5087922	$1.20x10^{-06}$
ERP27	-5.4765153	$7.39x10^{-09}$
CPA1	-5.4417426	$1.81x10^{-05}$

Log2FC, log2 fold change.

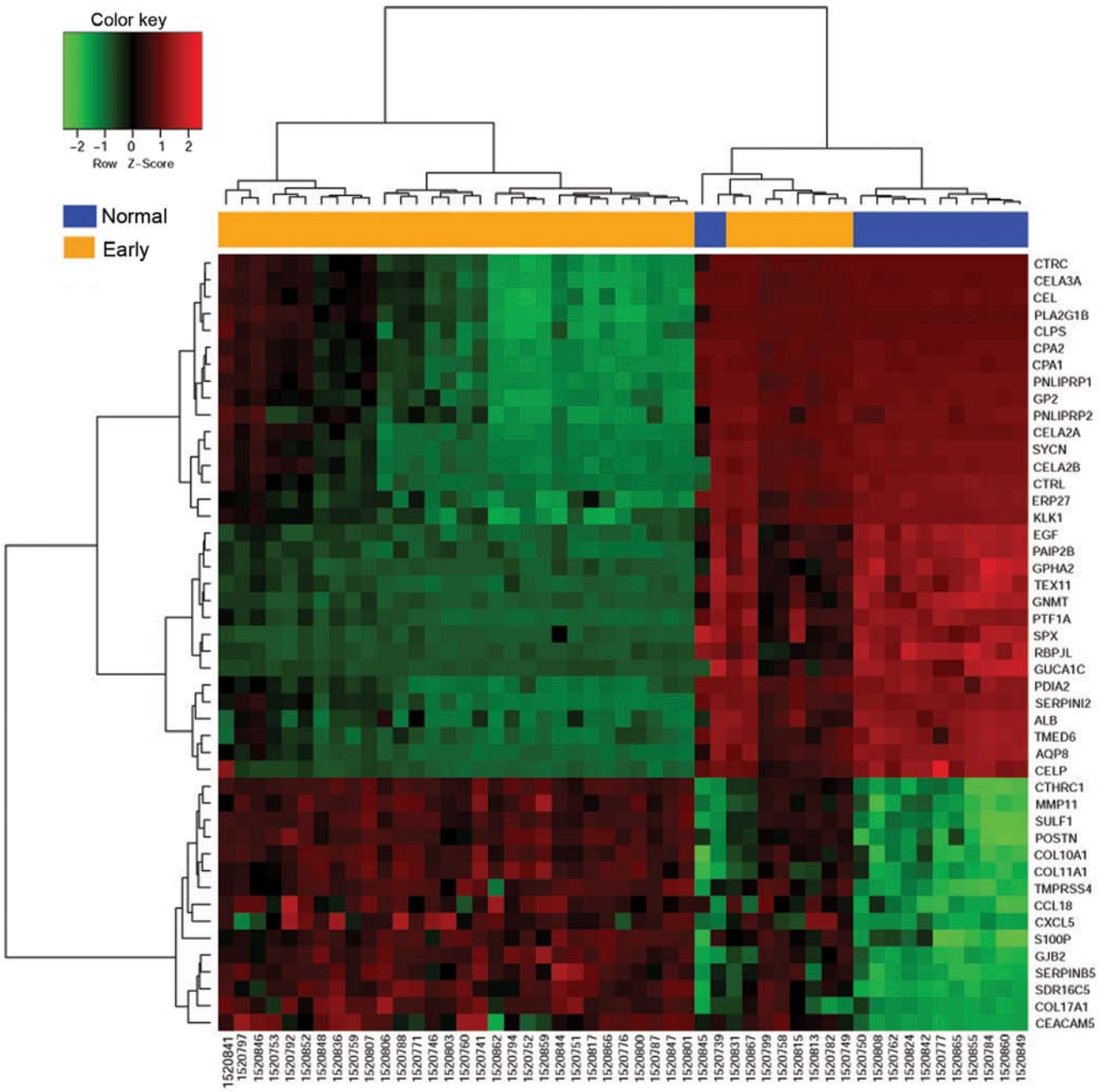

Figure 1. Heat map of gene expression in pancreatic ductal adenocarcinoma tissues and healthy controls. The expression levels of various genes in 51 samples (38 early-stage tumor samples and 13 controls) are presented. Green indicates downregulated expression; red indicates upregulated expression; black indicates no significant difference in expression.

GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r/) was used to determine the expression of DEGs. We further verified the expression of COL17A1.

Results

Analysis of DEGs. The selected chipset GSE62165 included 118 PDAC samples and 13 control samples. Differences in gene expression profiles were analyzed using 38 early-stage tumors and 13 normal tissues. A total of 240 DEGs (adjusted P-value <0.05; $|\log_2 FC| \geq 3.0$) were identified using R version 3.5 software, including 137 upregulated genes and 103 downregulated genes (Table I). The heat map of genes with upregulated expression is presented in Fig. 1. A volcanic map of all genes is presented in Fig. 2.

Enrichment analysis of DEGs. To investigated the distribution of DEGs, GO and KEGG analysis of upregulated and downregulated genes was conducted. GO analysis revealed that the 'biological processes' (BPs) of upregulated genes mainly included extracellular matrix organization, extracellular structure organization and collagen catabolic process. 'Molecular functions' (MFs) of upregulated DEGs primarily included extracellular matrix structural constituents, glycosaminoglycan binding and cytokine activity. For the 'cell components' (CCs) identified by GO analysis, proteinaceous

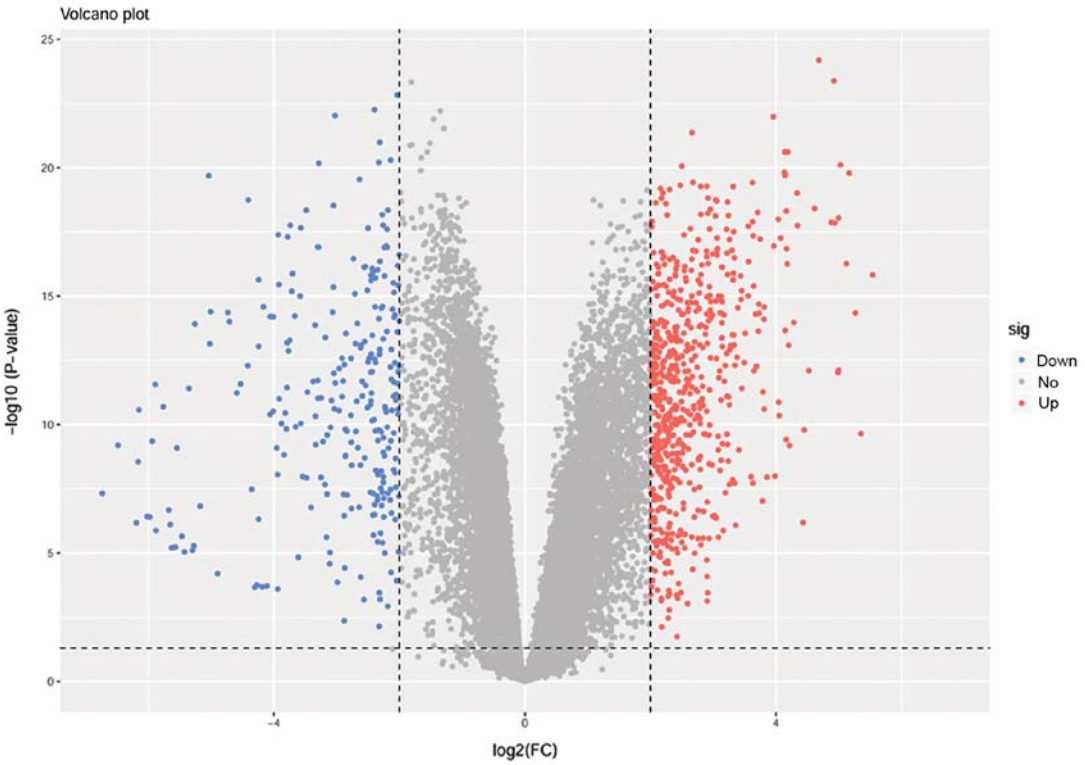

Figure 2. Volcano plot of the expression of genes in patients with early-stage pancreatic ductal adenocarcinoma. The expression of all identified genes in tumor tissues compared with in healthy control samples is presented. Blue indicates downregulated genes; red indicates upregulated genes; grey indicates genes that were not significantly differentially expressed. FC, fold change.

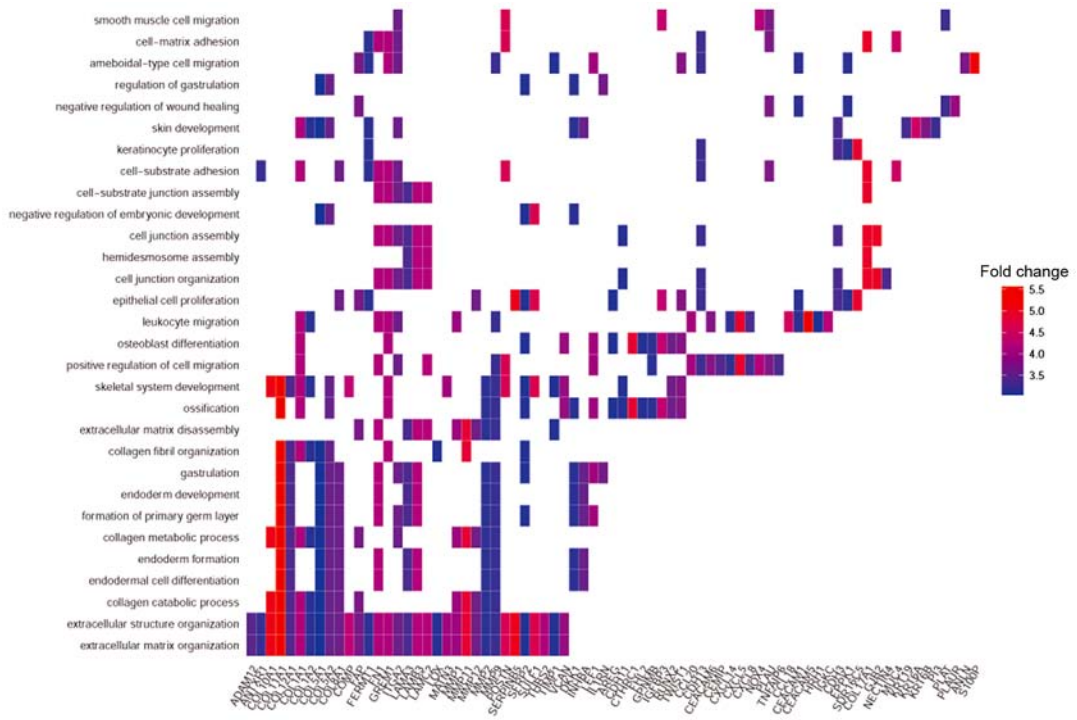

Figure 3. Functions of genes upregulated in pancreatic cancer tissues. Heat plot of the cell components, molecular functions and biological processes of upregulated genes in pancreatic cancer tissues, as identified by Gene Ontology analysis.

extracellular matrix, extracellular matrix component and endoplasmic reticulum lumen were the most prominent (Table II). For downregulated DEGs, the main enriched BPs were digestion, lipid digestion and sulfur amino acid metabolic process, whereas the primary MFs were exopeptidase activity, serine-type endopeptidase activity and serine-type

Table II. GO analysis of differentially expressed genes in pancreatic cancer.

A, Upregulated genes

Category	ID	Description	Count	P-value
GOBP	GO:0030198	Extracellular matrix organization	22	9.83×10^{-21}
GOBP	GO:0043062	Extracellular structure organization	22	1.05×10^{-20}
GOBP	GO:0030574	Collagen catabolic process	10	2.16×10^{-13}
GOBP	GO:0044243	Multicellular organismal catabolic process	10	6.89×10^{-13}
GOBP	GO:0032963	Collagen metabolic process	11	4.16×10^{-12}
GOMF	GO:0005201	Extracellular matrix structural constituent	8	1.09×10^{-09}
GOMF	GO:0005539	Glycosaminoglycan binding	7	2.85×10^{-05}
GOMF	GO:0005125	Cytokine activity	7	3.50×10^{-05}
GOMF	GO:0008009	Chemokine activity	4	4.62×10^{-05}
GOMF	GO:1901681	Sulfur compound binding	7	4.91×10^{-05}
GOCC	GO:0005578	Proteinaceous extracellular matrix	20	1.05×10^{-17}
GOCC	GO:0044420	Extracellular matrix component	10	4.54×10^{-11}
GOCC	GO:0005788	Endoplasmic reticulum lumen	13	1.70×10^{-10}
GOCC	GO:0005581	Collagen trimer	8	2.20×10^{-09}
GOCC	GO:0098644	Complex of collagen trimers	5	1.32×10^{-08}

B, Downregulated genes

Category	ID	Description	Count	P-value
GOBP	GO:0007586	Digestion	10	7.20×10^{-10}
GOBP	GO:0044241	Lipid digestion	5	1.00×10^{-08}
GOBP	GO:0000096	Sulfur amino acid metabolic process	4	1.27×10^{-05}
GOBP	GO:0009235	Cobalamin metabolic process	3	6.62×10^{-05}
GOBP	GO:1901605	α-Amino acid metabolic process	6	1.76×10^{-4}
GOMF	GO:0008238	Exopeptidase activity	8	7.37×10^{-09}
GOMF	GO:0004252	Serine-type endopeptidase activity	10	2.66×10^{-08}
GOMF	GO:0008236	Serine-type peptidase activity	10	7.39×10^{-08}
GOMF	GO:0008235	Metalloexopeptidase activity	6	8.31×10^{-08}
GOMF	GO:0017171	Serine hydrolase activity	10	8.76×10^{-08}

GO, Gene Ontology; MF, molecular function; CC, cell component; BP, biological process.

peptidase activity (Table II). Figs. 3 and 4 present the associations between genes and enrichment results, indicating the genes that were highly changed between the two conditions.

Table III presents KEGG pathway analysis of the DEGs, revealing that the upregulated genes were mainly located in extracellular matrix (ECM)-receptor interaction, protein digestion and absorption, and focal adhesion pathways. Conversely, downregulated genes were primarily located in pancreatic secretion, protein digestion and absorption, and fat digestion and absorption pathways. Figs. 5 and 6 present the distribution of the major KEGG pathways generated using clusterProfiler. It was observed that ECM-receptor interactions (Fig. 5) and pancreatic secretion (Fig. 6) were the pathways most enriched with up- and downregulated DEGs, respectively.

Screening of core genes in the PPI. Based on the information in the STRING database and using 12 types of calculation methods in Cytoscape, the following 18 core genes were identified: *EGF, ALB, COL17A1, FN1, TIMP1, PLAU, PLA2G1B, IGFBP3, PLAUR, VCAN, COL1A1, PNLIP, CTRL, PRSS3, COMP, CPB1, ITGA2* and *CEL*. These core genes were associated with each other and may exhibit synergistic effects in the development of pancreatic cancer (Fig. 7). According to the previous enrichment analysis, the core genes, were mainly located in pancreatic secretion, protein digestion and absorption, and focal adhesion pathways.

Gene expression level and survival analysis. Notably, COL17A1 and PLAU genes were the only genes associated with survival. Following the identification of core genes, survival analysis for PAAD was performed using UALCAN. *PLAU* [which encodes the serine protease urokinase-type plasminogen activator (uPA); Fig. 8] and *COL17A1* [which encodes collagen type XVII α1 chain (COL17A1); Fig. 9]

Table III. KEGG pathway analysis of differentially expressed genes in pancreatic cancer.

A, Upregulated genes

ID	Description	Count	P-value
hsa04512	Extracellular matrix-receptor interaction	7	2.23×10^{-07}
hsa04974	Protein digestion and absorption	7	4.25×10^{-07}
hsa04510	Focal adhesion	7	8.15×10^{-05}
hsa04657	Interleukin-17 signaling pathway	5	1.32×10^{-4}
hsa05146	Amebiasis	5	1.53×10^{-4}

B, Downregulated genes

ID	Description	Count	P-value
hsa04972	Pancreatic secretion	13	3.63×10^{-16}
hsa04974	Protein digestion and absorption	8	1.03×10^{-08}
hsa04975	Fat digestion and absorption	6	4.10×10^{-08}
hsa00561	Glycerolipid metabolism	4	2.27×10^{-4}
hsa00260	Glycine, serine and threonine metabolism	3	1.01×10^{-3}

KEGG, Kyoto Encyclopedia of Genes and Genomes.

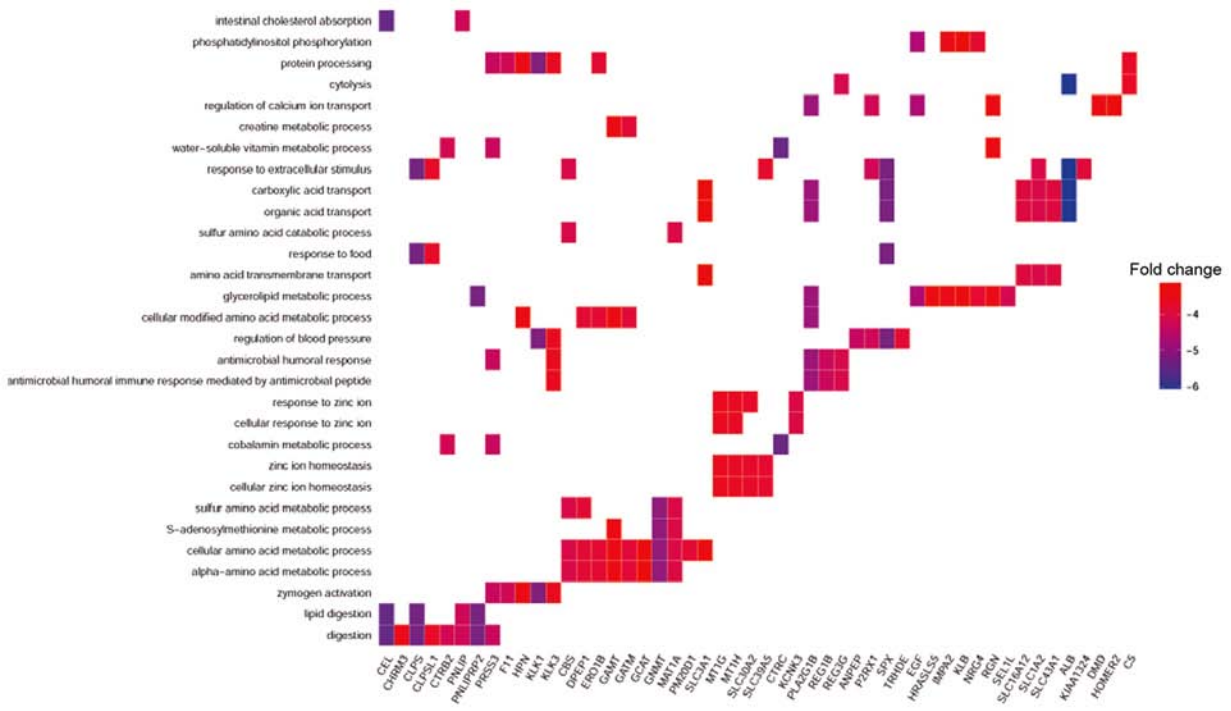

Figure 4. Functions of genes downregulated in pancreatic cancer tissues. Heat plot of the molecular functions and biological processes of downregulated genes in pancreatic cancer tissues, as identified by Gene Ontology analysis.

were demonstrated to be significantly associated with survival ($P<0.05$). Subsequently, the expression levels of genes in primary pancreatic cancer were compared; only one gene was identified to be significantly differentially expressed, *COL17A1*, whereas PLAU was not significantly differentially expressed. The expression levels of *COL17A1* were analyzed in TCGA database, and the results were consistent with those of the aforementioned differential gene analysis; *COL17A1* was significantly upregulated in PAAD tumor tissues compared with normal tissues ($P=1.62 \times 10^{-12}$; Fig. 10).

Verification of COL17A1. Differences in gene expression between 45 pancreatic cancer patients and 45 normal pancreatic tissues were analyzed. In particular, the expression level

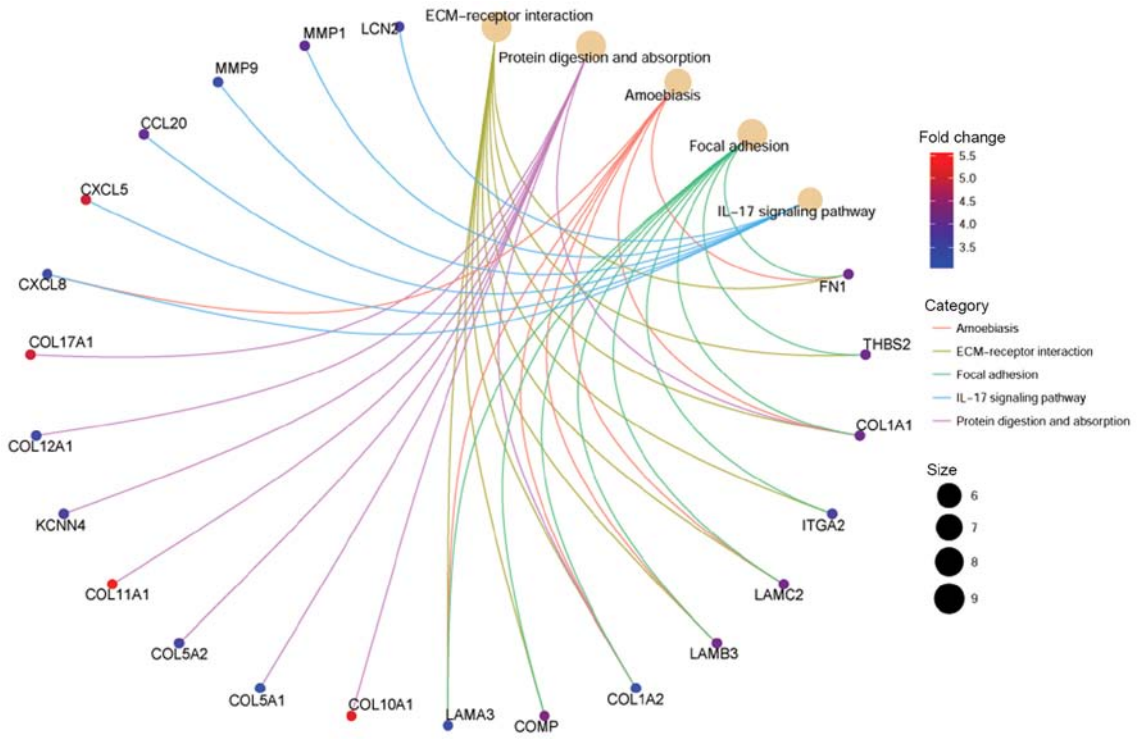

Figure 5. Pathways enriched with genes upregulated in pancreatic cancer. Net plot of the pathways enriched with genes identified as upregulated in pancreatic cancer tissues, as identified by Kyoto Encyclopedia of Genes and Genomes pathway analysis. ECM, extracellular matrix; IL-17, interleukin-17.

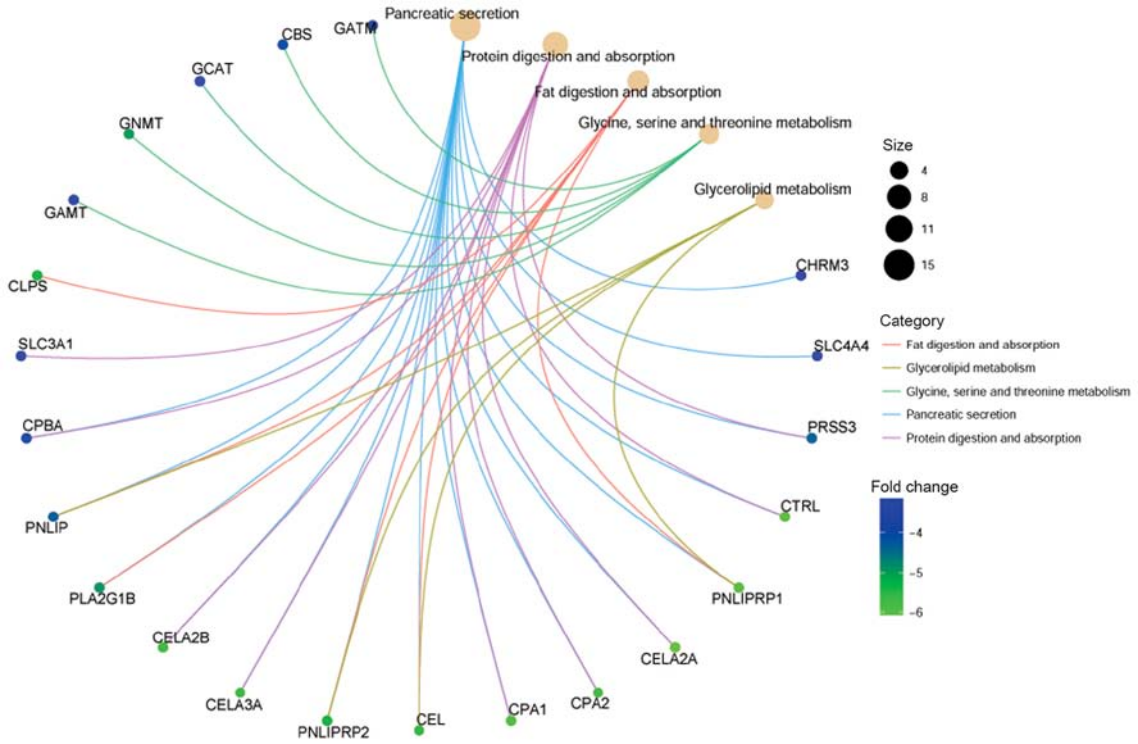

Figure 6. Pathways enriched with genes downregulated in pancreatic cancer. Net plot of the pathways enriched with genes identified as downregulated in pancreatic cancer tissues, as identified by Kyoto Encyclopedia of Genes and Genomes pathway analysis.

of COL17A1 was investigated. The results of the analysis to verify the importance of *COL17A1* are presented in Table IV; it was observed that *COL17A1* was significantly upregulated in pancreatic tumor tissue in the two GEO databases.

Discussion

The incidence of pancreatic cancer and the associated mortality rates have exhibited an increasing trend in previous years (3).

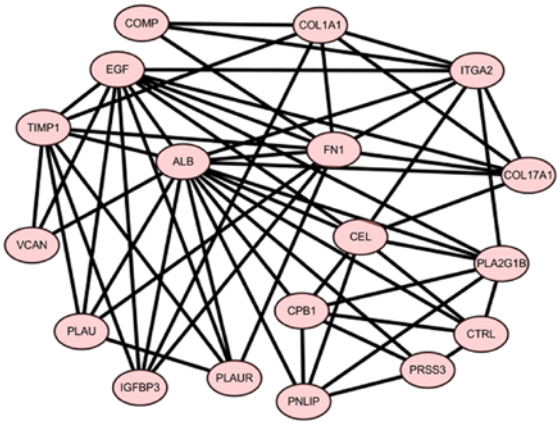

Figure 7. Protein-protein interaction network of the 18 identified core genes.

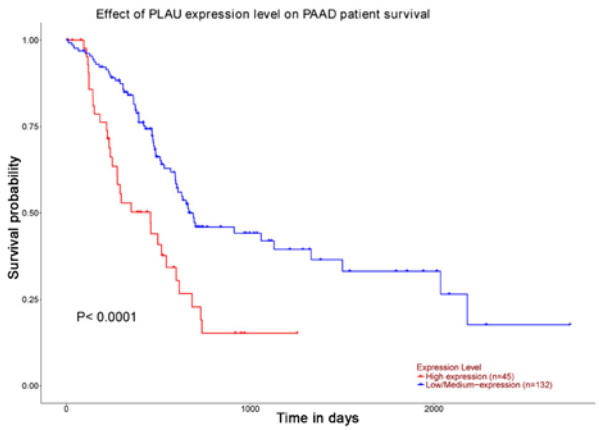

Figure 8. Survival analysis of *PLAU* in PAAD. Kaplan-Meier analysis of the association between the expression of *PLAU* and the overall survival of patients with PAAD. PAAD, pancreatic adenocarcinoma; *PLAU*, gene encoding urokinase-type plasminogen activator.

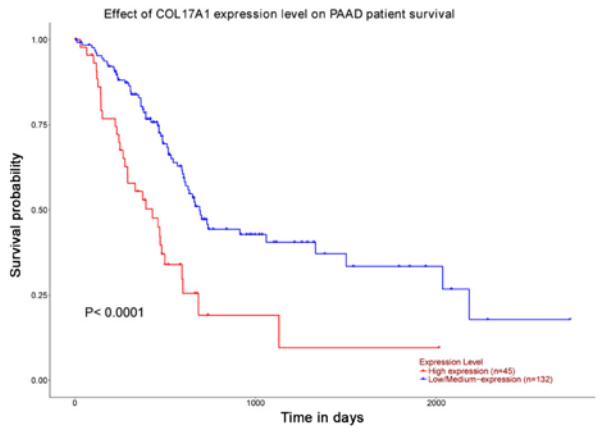

Figure 9. Survival analysis of *COL17A1* in PAAD. Kaplan-Meier analysis of the association between the expression of *COL17A1* and the overall survival of patients with PAAD. PAAD, pancreatic adenocarcinoma; *COL17A1*, gene encoding collagen type XVII α1 chain.

Studies have reported that patients with pancreatic cancer survive for only 4 months on average without treatment; even in patients who undergo treatment, the survival is not significantly extended (25). Therefore, accurate early diagnosis of pancreatic cancer and the development of effective targeted therapy is of major importance.

Table IV. Differential expression of COL17A1 in pancreatic cancer tissues in two databases.

Database	Gene	Log2FC	Adjusted P-value
GSE62165	COL17A1	5.0501325	$1.8x10^{-11}$
GSE28735	COL17A1	1.893626	$6.56x10^{-13}$

Log2FC, Log2 fold change; COL17A1, collagen type XVII α1 chain.

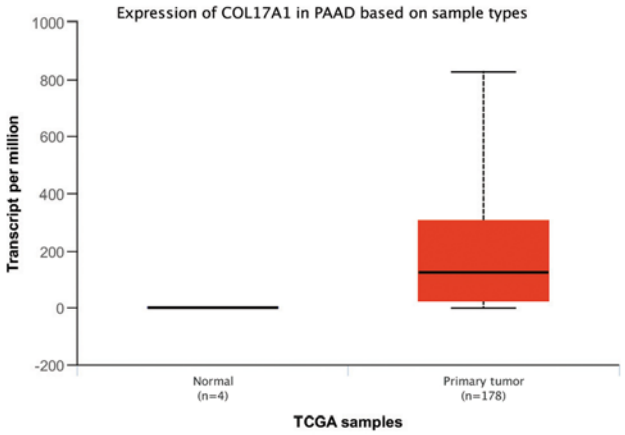

Figure 10. Expression levels of *COL17A1* in PAAD and normal tissues. The expression of *COL17A1* was compared between PAAD primary tumor and normal control tissues, based on data from TCGA. PAAD, pancreatic adenocarcinoma; *COL17A1*, gene encoding collagen type XVII α1 chain; TCGA, The Cancer Genome Atlas.

A previous study identified core genes in pancreatic cancer that were reported to be of diagnostic relevance (26). In the present study, the chipset GSE62165 from the GEO was analyzed, containing data of 118 PDAC and 13 normal pancreatic tissues (11). Differences in gene expression levels were only compared between normal tissues and early-stage tumor tissue. A total of 240 DEGs (137 upregulated and 103 downregulated) were identified using R, and GO (27) and KEGG pathway analyses of DEGs revealed the locations and functions of DEGs. Upregulated genes were mainly located in the ECM and collagen trimers, and were involved in ECM organization and ECM-receptor interactions, focal adhesion, and protein digestion and absorption. Conversely, downregulated genes were mainly enriched in digestion and exopeptidase activity pathways. A PPI network was built, and 18 core genes were identified; the prognostic value of these genes for patients with pancreatic cancer was analyzed using UACLAN. *PLAU* and *COL17A1* were significantly associated with poorer survival; it was then revealed using data from TCGA that *COL17A1* was significantly upregulated in pancreatic cancer tissues compared with control tissues, consistent with the results of the differential gene analysis. It was predicted that these two genes may be associated with the proliferation, invasion and metastasis of pancreatic cancer.

PLAU encodes a serine protease, uPA (28). Following GO and KEGG analyses, the functional enrichment of PLAU was

investigated. *PLAU* is mainly involved in the regulation of cell motility, cellular component movement and locomotion (29). It is primarily expressed in the endoplasmic reticulum lumen and invadopodium (30). PLAU plays a key role in regulating cell migration and adhesion during tissue regeneration and intracellular signaling (31). Increased expression of COL17A1 leads to tumor cell invasion and metastasis of tumor cells to surrounding tissues (32). PLAU is involved in predicting the survival rate of patients with gastric cancer (33). It may serve an important role in the invasion and metastasis of pancreatic cancer cells (34); however, the specific pathways involved are yet to be determined. It is hypothesized that *PLAU* may serve an important role in the diagnosis and treatment of pancreatic cancer in the future.

COL17A1 is mainly located in the extracellular matrix and collagen trimmers (35). Extracellular matrix molecules, including proteoglycan and fibrin, have been reported to affect the growth, migration and differentiation of cells (36). A study showed that COL17A1 can inhibit the migration and invasion of breast cancer cells, acting as a p53 transcriptional target gene (37). A previous study has reported that the extracellular matrix is closely associated with the metastasis of breast cancer (38). High levels of collagen in breast and colorectal cancers have been associated with tumor invasion (39,40). A previous study that employed the minimum-redundancy-maximum-relevance method also identified *COL17A1* as a core gene of pancreatic cancer (26); however, in the present study, the upregulated expression of *COL17A1* in pancreatic cancer was verified in multiple datasets, and its effects on patient survival were determined. Survival analysis using UACLAN based on data from TCGA revealed that the expression levels of *CLO17A1* were closely associated with the survival of patients with pancreatic cancer, and that *CLO17A1* was highly expressed in primary pancreatic tumor tissues. The present findings suggested that the expression of *COL17A1* is associated with the occurrence and development of pancreatic cancer. Therefore, this bioinformatics analysis may provide novel insight for future studies investigating the pathogenesis of pancreatic cancer.

However, the present study presented certain limitations. In examining the expression level of COL17A1, only four normal samples were investigated, and further studies examining a high number of control samples are required to confirm the present results.

Acknowledgements

Not applicable.

Funding

The present work was supported by The 'Six Talents Summit' Project in Jiangsu Province, miR-203 targets Survivin to upregulate the expression of Caspase-3 and promote the apoptosis of pancreatic cancer cells (grant no. WAW-008).

Availability of data and materials

The datasets used and/or analyzed in the present study are available in the GEO (http://www.ncbi.nlm.nih.gov/geo) and UALCAN (http://ualcan.path.uab.edu) repositories.

Authors' contributions

JZ and LX conceived the study. JW, ZL, KW, KZ and DX analyzed the data and drafted the manuscript. All authors reviewed and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Kamisawa T, Wood LD, Itoi T and Takaori K: Pancreatic cancer. Lancet 388: 73-85, 2016.
2. Michaud D: Epidemiology of pancreatic cancer. Minerva Chir 59: 99-111, 2004.
3. Siegel R, Ma J, Zou Z and Jemal A: Cancer statistics, 2014. CA Cancer J Clin 64: 9-29, 2014.
4. Risch HA: Etiology of pancreatic cancer, with a hypothesis concerning the role of N-nitroso compounds and excess gastric acidity. J Natl Cancer Inst 95: 948-960, 2003.
5. Kern SE, Shi C and Hruban RH: The complexity of pancreatic ductal cancers and multidimensional strategies for therapeutic targeting. J Pathol 223: 295-306, 2011.
6. Grasso C, Jansen G and Giovannetti E: Drug resistance in pancreatic cancer: Impact of altered energy metabolism. Crit Rev Oncol Hematol 114: 139-152, 2017.
7. Eskelinen M and Haglund U: Developments in serologic detection of human pancreatic adenocarcinoma. Scand J Gastroenterol 34: 833-844, 1999.
8. Bass AJ, Lawrence MS, Brace LE, Ramos AH, Drier Y, Cibulskis K, Sougnez C, Voet D, Saksena G, Sivachenko A, *et al*: Genomic sequencing of colorectal adenocarcinomas identifies a recurrent VTI1A-TCF7L2 fusion. Nat Genet 43: 964-968, 2011.
9. Sjöblom T, Jones S, Wood LD, Parsons DW, Lin J, Barber TD, Mandelker D, Leary RJ, Ptak J, Silliman N, *et al*: The consensus coding sequences of human breast and colorectal cancers. Science 314: 268-274, 2006.
10. Wood LD, Parsons DW, Jones S, Lin J, Sjöblom T, Leary RJ, Shen D, Boca SM, Barber T, Ptak J, *et al*: The genomic landscapes of human breast and colorectal cancers. Science 318: 1108-1113, 2007.
11. Janky R, Binda MM, Allemeersch J, Van den Broeck A, Govaere O, Swinnen JV, Roskams T, Aerts S and Topal B: Prognostic relevance of molecular subtypes and master regulators in pancreatic ductal adenocarcinoma. BMC Cancer 16: 632, 2016.
12. 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.
13. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W and Smyth GK: Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43: e47, 2015.
14. Zhang G, Schetter A, He P, Funamizu N, Gaedcke J, Ghadimi BM, Ried T, Hassan R, Yfantis HG, Lee DH, *et al*: DPEP1 inhibits tumor cell invasiveness, enhances chemosensitivity and predicts clinical outcome in pancreatic ductal adenocarcinoma. PLoS One 7: e31507, 2012.
15. Zhang G, He P, Tan H, Budhu A, Gaedcke J, Ghadimi BM, Ried T, Yfantis HG, Lee DH, Maitra A, *et al*: Integration of metabolomics and transcriptomics revealed a fatty acid network exerting growth inhibitory effects in human pancreatic cancer. Clin Cancer Res 19: 4983-4993, 2013.

16. Xing Z, Chu C, Chen L and Kong X: The use of Gene Ontology terms and KEGG pathways for analysis and prediction of oncogenes. Biochim Biophys Acta 1860: 2725-2734, 2016.
17. Kanehisa M, Sato Y, Furumichi M, Morishima K and Tanabe M: New approach for understanding genome variations in KEGG. Nucleic Acids Res 47: D590-D595, 2019.
18. Kanehisa M, Furumichi M, Tanabe M, Sato Y and Morishima K: KEGG: New perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res 45: D353-D361, 2017.
19. Kanehisa M and Goto S: KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28: 27-30, 2000.
20. Yu G, Wang LG, Han Y and He QY: clusterProfiler: An R package for comparing biological themes among gene clusters. OMICS 16: 284-287, 2012.
21. 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.
22. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT and Lin CY: cytoHubba: Identifying hub objects and sub-networks from complex interactome. BMC Syst Biol 8 (Suppl 4): S11, 2014.
23. 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.
24. Li B and Dewey CN: RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12: 323, 2011.
25. Wang X, Wang L, Mo Q, Dong Y, Wang G and Ji A: Changes of Th17/Treg cell and related cytokines in pancreatic cancer patients. Int J Clin Exp Pathol 8: 5702-5708, 2015.
26. Shen S, Gui T and Ma C: Identification of molecular biomarkers for pancreatic cancer with mRMR shortest path method. Oncotarget 8: 41432-41439, 2017.
27. Thomas PD: The gene ontology and the meaning of biological function. Methods Mol Biol 1446: 15-24, 2017.
28. Duffy MJ, Duggan C, Mulcahy HE, McDermott EW and O'Higgins NJ: Urokinase plasminogen activator: A prognostic marker in breast cancer including patients with axillary node-negative disease. Clin Chem 44: 1177-1183, 1998.
29. Nielsen TO, Andrews HN, Cheang M, Kucab JE, Hsu FD, Ragaz J, Gilks CB, Makretsov N, Bajdik CD, Brookes C, *et al*: Expression of the insulin-like growth factor I receptor and urokinase plasminogen activator in breast cancer is associated with poor survival: Potential for intervention with 17-allylamino geldanamycin. Cancer Res 64: 286-291, 2004.
30. Pavón MA, Arroyo-Solera I, Céspedes MV, Casanova I, León X and Mangues R: uPA/uPAR and SERPINE1 in head and neck cancer: Role in tumor resistance, metastasis, prognosis and therapy. Oncotarget 7: 57351-57366, 2016.
31. Amos S, Redpath GT, Dipierro CG, Carpenter JE and Hussaini IM: Epidermal growth factor receptor-mediated regulation of urokinase plasminogen activator expression and glioblastoma invasion via C-SRC/MAPK/AP-1 signaling pathways. J Neuropathol Exp Neurol 69: 582-592, 2010.
32. Chaudhary A, Hilton MB, Seaman S, Haines DC, Stevenson S, Lemotte PK, Tschantz WR, Zhang XM, Saha S, Fleming T and St Croix B: TEM8/ANTXR1 blockade inhibits pathological angiogenesis and potentiates tumoricidal responses against multiple cancer types. Cancer Cell 21: 212-226, 2012.
33. Xu ZY, Chen JS and Shu YQ: Gene expression profile towards the prediction of patient survival of gastric cancer. Biomed Pharmacother 64: 133-139, 2010.
34. Liu P, Weng Y, Sui Z, Wu Y, Meng X, Wu M, Jin H, Tan X, Zhang L and Zhang Y: Quantitative secretomic analysis of pancreatic cancer cells in serum-containing conditioned medium. Sci Rep 6: 37606, 2016.
35. Borradori L and Sonnenberg A: Structure and function of hemidesmosomes: More than simple adhesion complexes. J Invest Dermatol 112: 411-418, 1999.
36. Järveläinen H, Sainio A, Koulu M, Wight TN and Penttinen R: Extracellular matrix molecules: Potential targets in pharmacotherapy. Pharmacol Rev 61: 198-223, 2009.
37. Yodsurang V, Tanikawa C, Miyamoto T, Lo PHY, Hirata M and Matsuda K: Identification of a novel p53 target, COL17A1, that inhibits breast cancer cell migration and invasion. Oncotarget 8: 55790-55803, 2017.
38. Chowdhury N and Sapru S: Association of protein translation and extracellular matrix gene sets with breast cancer metastasis: Findings uncovered on analysis of multiple publicly available datasets using individual patient data approach. PLoS One 10: e0129610, 2015.
39. Rizwan A, Bulte C, Kalaichelvan A, Cheng M, Krishnamachary B, Bhujwalla ZM, Jiang L and Glunde K: Metastatic breast cancer cells in lymph nodes increase nodal collagen density. Sci Rep 5: 10002, 2015.
40. Zou X, Feng B, Dong T, Yan G, Tan B, Shen H, Huang A, Zhang X, Zhang M, Yang P, *et al*: Up-regulation of type I collagen during tumorigenesis of colorectal cancer revealed by quantitative proteomic analysis. J Proteomics 94: 473-485, 2013.