

Identification of differentially expressed genes in pancreatic ductal adenocarcinoma and normal pancreatic tissues based on microarray datasets

LIYING LIU^{1,2*}, SIQI WANG^{1,2*}, CHUNYUAN CEN^{1,2}, SHUYI PENG^{1,2}, YAN CHEN^{1,2},
XIN LI^{1,2}, NAN DIAO^{1,2}, QIAN LI^{1,2}, LING MA³ and PING HAN^{1,2}

¹Department of Radiology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology;

²Hubei Province Key Laboratory of Molecular Imaging, Wuhan, Hubei 430022;

³Advanced Application Team, GE Healthcare, Shanghai 201203, P.R. China

Received August 15, 2018; Accepted May 1, 2019

DOI: 10.3892/mmr.2019.10414

Abstract. Pancreatic ductal adenocarcinoma (PDAC) is a highly aggressive malignant tumor with rapid progression and poor prognosis. In the present study, 11 high-quality microarray datasets, comprising 334 tumor samples and 151 non-tumor samples from the Gene Expression Omnibus, were screened, and integrative meta-analysis of expression data was used to identify gene signatures that differentiate between PDAC and normal pancreatic tissues. Following the identification of differentially expressed genes (DEGs), two-way hierarchical clustering analysis was performed for all DEGs using the *gplots* package in R software. Hub genes were then determined through protein-protein interaction network analysis using NetworkAnalyst. In addition, functional annotation and pathway enrichment analyses of all DEGs were conducted in the Database for Annotation, Visualization, and Integrated Discovery. The expression levels and Kaplan-Meier analysis of the top 10 upregulated and downregulated genes were verified in The Cancer Genome Atlas. A total of 1,587 DEGs, including 1,004 upregulated and 583 downregulated genes, were obtained by comparing PDAC with normal tissues. Of these, hematological and neurological expressed 1, integrin subunit $\alpha 2$ (*ITGA2*) and S100 calcium-binding protein A6 (*S100A6*) were the top upregulated genes, and kinesin family member 1A, Dymeclin and β -secretase 1 were the top

downregulated genes. Reverse transcription-quantitative PCR was performed to examine the expression levels of *S100A6*, *KRT19* and *GNG7*, and the results suggested that *S100A6* was significantly upregulated in PDAC compared with normal pancreatic tissues. *ITGA2* overexpression was significantly associated with shorter overall survival times, whereas family with sequence similarity 46 member C overexpression was strongly associated with longer overall survival times. In addition, network-based meta-analysis confirmed growth factor receptor-bound protein 2 and histone deacetylase 5 as pivotal hub genes in PDAC compared with normal tissue. In conclusion, the results of the present meta-analysis identified PDAC-related gene signatures, providing new perspectives and potential targets for PDAC diagnosis and treatment.

Introduction

Pancreatic ductal adenocarcinoma (PDAC) is one of the most aggressive malignant tumors worldwide; it has a poor prognosis and a 5-year survival rate of ~8% in the USA (1). The median survival time for patients with advanced PDAC is <7 months, even with active treatment (2). Statistics from the National Cancer Institute Surveillance, Epidemiology and End Results database (<https://seer.cancer.gov>) indicate that PDAC is the fourth leading cause of cancer-related mortality, accounting for ~7% of all cancer deaths in the United States. The mortality rate of men with pancreatic cancer increased slightly from 2005 to 2014, at a rate of 0.3% per year (1). Based on this trend, Hezel *et al* predicted that pancreatic cancer would become the second most common cause of cancer-related death by 2030 (3). Resection of pancreatic tumor tissue during the early stage is currently the only curative treatment method. However, as symptoms and biomarkers for early diagnosis are lacking, >80% of patients with PDAC are not diagnosed until they are in the advanced stage, thereby missing the best timeframe for curative treatment (4). Additionally, the lack of effective and credible interventions is another key cause of the high mortality rate in patients with PDAC (3). Thus, new targets for PDAC diagnosis and better therapeutics for PDAC treatment are urgently needed.

Correspondence to: Professor Ping Han, Department of Radiology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, 1277 Jiefang Avenue, Wuhan, Hubei 430022, P.R. China
E-mail: cjr.hanping@vip.163.com

*Contributed equally

Key words: pancreatic ductal adenocarcinoma, differentially expressed genes, functional annotation, pathway enrichment analysis, protein-protein interaction

Pancreatic cancer has many risk factors, including individual, lifestyle-related, disease and drug-related factors, with PDAC being primarily caused by somatically acquired mutations (5). Jones *et al* investigated gene mutations involved in pancreatic cancer and found an average of 63 mutated genes (6). A number of previous studies have examined gene mutations in pancreatic cancer. For example, Nagata *et al* reported and validated that mucins, such as mucin (*MUC*) 1, *MUC5* and *MUC6*, were overexpressed in invasive ductal carcinoma (IDC) and in pancreatic intraepithelial neoplasia, which develops into IDC (7). Kirsten rat sarcoma viral oncogene homolog (*KRAS*) has been confirmed as the driver gene in PDAC (6-8). By establishing a genetically engineered mouse PDAC model, Rozenblum *et al* demonstrated that *SMAD4* is inactivated in most pancreatic cancers, which leads to inactivation of the tumor suppressor gene cyclin-dependent kinase inhibitor 2A and activation of *KRAS* (9,10). p53, another tumor suppressor gene, has been found to be mutated in >50% of patients with PDAC (3,9). Fong and Winter systematically summarized the presence of various biomarkers in cancer cells and body fluids in patients with PDAC based on previous studies; these biomarkers included the genes *KRAS*, Fanconi anemia complementation group I (*FANCI*) and *MUC1*, and the proteins CA-19-9, CA-125 and CEA, as well as microRNA (miR)-21 and miR-210 (11). Unexpectedly, although >2,000 studies have been conducted on biomarkers for pancreatic cancer, the CA-19-9 serum level is the only widely used Food and Drug Administration-approved tumor marker, albeit not for diagnosis. To date, no biomarkers are sufficiently accurate for widespread diagnostic use (12).

Since the emergence of high-throughput genomic technologies, including microarrays several decades ago, these techniques have been applied to study many diseases. For example, these technologies are used to identify disease subtypes (13), explore gene expression profiles of disease (14) and to identify potential novel pathogenic genes and diseases (15). Additionally, molecular signatures discovered by microarrays have become predictive biomarkers for certain diseases (16). An increasing number of studies have applied this technique to identify differentially expressed genes (DEGs) between tumor tissues and normal tissues (17). Although microarray technology has enabled the discovery of biomarkers, study results vary owing to inaccuracies or quality problems in microarray analyses and verification, and small sample sizes exacerbate this issue (18). A meta-analysis, as a large-sample study, has an advantage in addressing this limitation due to its enhanced statistical power (19). Prior meta-analyses have been applied to study tumors to confirm DEGs between tumor tissue and normal tissue in glioma (20,21), lung cancer (22), bladder cancer (23), breast cancer (24), osteosarcoma (25), liver cancer (26) and pancreatic cancer (27). In the present study, integrative meta-analysis of expression data (INMEX) was used to conduct a meta-analysis based on 11 qualified microarray datasets, with the aim to identify crucial DEGs between PDAC samples and normal pancreatic samples that may serve as biomarkers for PDAC treatment and prognosis.

Materials and methods

Patients and tissue collection. The mean age of patients was 56.3 years (range, 39-73). Paired normal pancreatic tissues and

Table I. Primer sequences.

| Gene name | Primer sequence (5'-3') |
|---------------|--|
| <i>GAPDH</i> | F:ACTTTGGTATCGTGGAAGGACTCAT R:GTTTTTCTAGACGGCAGGTCAGG |
| <i>S100A6</i> | F:CCATCTTCCACAAGTACTCCGG R:GCAGCTTCGAGCCAATGGT |
| <i>GNG7</i> | F:CGCATAGAAGCCGGGATTGA R:TTGTCCTTAAAGGGGTTCTCCG |
| <i>KRT19</i> | F:AGAATTGAACCGGGAGGTCCG R:CCTGATTCTGCCGCTCACTA |

S100A6, S100 calcium-binding protein; *GNG7*, G protein subunit 7; *KRT19*, keratin-19.

PDAC tissues were collected from 15 patients (female to male ratio, 4:11) with PDAC who underwent surgical procedures at the Union Hospital of Tongji Medical College, Huazhong University of Science and Technology (Wuhan, China) between July 2018 and September 2018. Human study protocols were approved by the ethics committee of Tongji Medical College, Huazhong University of Science and Technology (Wuhan, China), and all patients in the study provided written informed consent. The pathological diagnoses of the specimens were confirmed by two clinicopathological experts at the Department of Pathology, Union Hospital of Tongji Medical College, Huazhong University of Science and Technology prior to RNA extraction. Specimens were snap frozen in liquid nitrogen and stored at -80°C. According to the American Joint Commission on Cancer Staging System for patients with pancreatic adenocarcinoma (28), patients were divided into three groups: i) Stage I (n=7); ii) stage II (n=7); and stage III (n=1).

Identification of PDAC microarray datasets and validation in The Cancer Genome Atlas (TCGA). The keywords 'pancreatic cancer' were used to search for gene chips to study in the Gene Expression Omnibus (GEO) datasets of the National Center for Biotechnology Information database (<http://www.ncbi.nlm.nih.gov/geo>). The inclusion criteria for the qualified chips were as follows: i) The chip was a gene expression chip; ii) the genes were from normal pancreatic and/or PDAC tissues; iii) the GEO series (GSE) dataset included >3 samples. Gene chips that met any of the following criteria were excluded: i) Chips containing pancreatic cancer cell lines; ii) methylated gene chips; iii) genes originating from animals other than humans; iv) non-microarray gene chips. Following the determination of the microarray datasets, the GSE number, information regarding the expression platform and sample number, source literature and the number of normal and pancreatic cancer samples were collected from the GEO database. To validate the results in TCGA (<https://tcga-data.nci.nih.gov/tcga/>), DNA expression levels in pancreatic cancer and normal tissues and patients' survival time were obtained from the database.

Elimination of batch differences and individual data analysis. INMEX (<http://www.inmex.ca>) is an online tool that supports

Table II. Characteristics of datasets used in the meta-analysis of pancreatic ductal adenocarcinoma vs. Normal tissue.

| Author, year | Source accession | Country | Platform | Numbers | | (Refs.) |
|---|------------------|-------------|----------|---------|--------|---------|
| | | | | PDAC | Normal | |
| Badea <i>et al.</i> , 2009 | GSE15471 | Romania | GPL570 | 39 | 39 | (35) |
| Pei <i>et al.</i> , 2009 | GSE16515 | USA | GPL570 | 36 | 16 | (36) |
| Killary, 2011 | GSE22780 | USA | GPL570 | 8 | 8 | N/A |
| Hussain <i>et al.</i> , 2012 | GSE28735 | USA | GPL6244 | 45 | 45 | (37,38) |
| Donahue <i>et al.</i> , 2011 | GSE32676 | USA | GPL570 | 25 | 7 | (39) |
| Crnogorac-Jurcevic <i>et al.</i> , 2013 | GSE43288 | UK | GPL96 | 8 | 6 | (40) |
| Crnogorac-Jurcevic <i>et al.</i> , 2013 | GSE43288 | UK | GPL97 | 8 | 6 | (40) |
| Kim <i>et al.</i> , 2013 | GSE43795 | South Korea | GPL10558 | 6 | 5 | (41) |
| Raeder, 2017 | GSE46234 | Norway | GPL570 | 4 | 4 | N/A |
| Jamieson <i>et al.</i> , 2014 | GSE55643 | UK | GPL6480 | 45 | 8 | (42) |
| Janky <i>et al.</i> , 2016 | GSE62165 | Belgium | GPL13667 | 118 | 13 | (43) |

multiple microarray platforms and commonly used gene IDs; it has comprehensive processing capabilities for meta-analysis of multiple gene expression sets (29). The ComBat option using the empirical Bayes method in INMEX was applied to eliminate the differences between dataset batches on different platforms and on the same platform and to ensure that different microarray experiments were directly comparable to each other to avoid inaccuracies resulting from differences unrelated to the disease. Extremely high or low expression ratios for additional genes were robustly stabilized by contracting differences by the empirical Bayes method (30).

Meta-analysis of microarray datasets. All selected microarray datasets were log₂ transformed and each constructed relative expression value table was set to match INMEX's upload format: The rows contained gene information and the columns contained sample information; each table was then uploaded. The data were annotated, function IDs were matched, and outlier samples were determined from boxplots and principal component analysis plots to ensure normalized quantile data. The correctness and integrity of all datasets were checked prior to running the ComBat adjust batch effect function, and the meta-analysis was performed. Considering the characteristics of each statistical method, the Cochran Q test was performed on the uploaded INMEX data, which revealed notable heterogeneity among different studies. Therefore, a random-effects model combined with the moderated effect size (ES) and metaMA packages were selected for the DEG screening (31,32). The gplots package in R software (version 3.5.0; <https://www.r-project.org>) was used to perform hierarchical cluster analysis of the DEGs.

Functional enrichment analysis of the DEGs. To determine the potential functions of the screened DEGs, Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses were performed in the Database for Annotation, Visualization, and Integrated Discovery (DAVID; <https://david.ncifcrf.gov>) and the top 10 most enriched terms with P<0.05 were identified.

Network-based meta-analysis. To better understand DEG expression, protein-protein interaction (PPI) networks were constructed using NetworkAnalyst with an extensive and high-quality PPI database based in InnateDB (33). The Hub Explorer tool was used to retrieve information on node levels, betweenness centrality and expression levels. The types of data obtained were the degree of the node (number of connections to other nodes) and betweenness centrality (number of shortest paths through the node). The expression was defined as the log fold change value of the corresponding node. Nodes with the highest degree or betweenness values were considered crucial hub nodes.

Verification of gene expression and Kaplan-Meier analysis. To verify the expression of the top ten upregulated and downregulated genes in PDAC tissues compared with normal tissues, scatter plots of TCGA data were produced using GraphPad Prism 6 (GraphPad Software, Inc.). To explore the effects of DEGs on survival time, Kaplan-Meier curves were constructed using GraphPad Prism 6 based on DNA expression profiles in the PDAC tissue samples and the corresponding survival times extracted from TCGA. The median expression level of each DEG was set as the cutoff value to split TCGA cohorts into high- and low-expression groups. The log-rank tests were used to assess the differences between groups of patients exhibiting low or high expression of the selected DEGs. P<0.05 was considered to indicate a statistically significant difference.

RNA extraction and reverse transcription-quantitative PCR (RT-qPCR). To confirm *SI00A6*, *KRT19* and *GNG7* expression, PDAC tissues and paired normal pancreatic tissues were collected. Total RNA was extracted from tissues using TRIzol reagent (Invitrogen; Thermo Fisher Scientific, Inc.) following the manufacturer's protocol. First-strand cDNA was synthesized using the RevertAid First Strand cDNA Synthesis kit (Thermo Fisher Scientific, Inc.) following the manufacturer's instructions. RT-qPCR was performed using the FastStart Universal SYBR Green master mix (Roche Diagnostics). The thermocycling

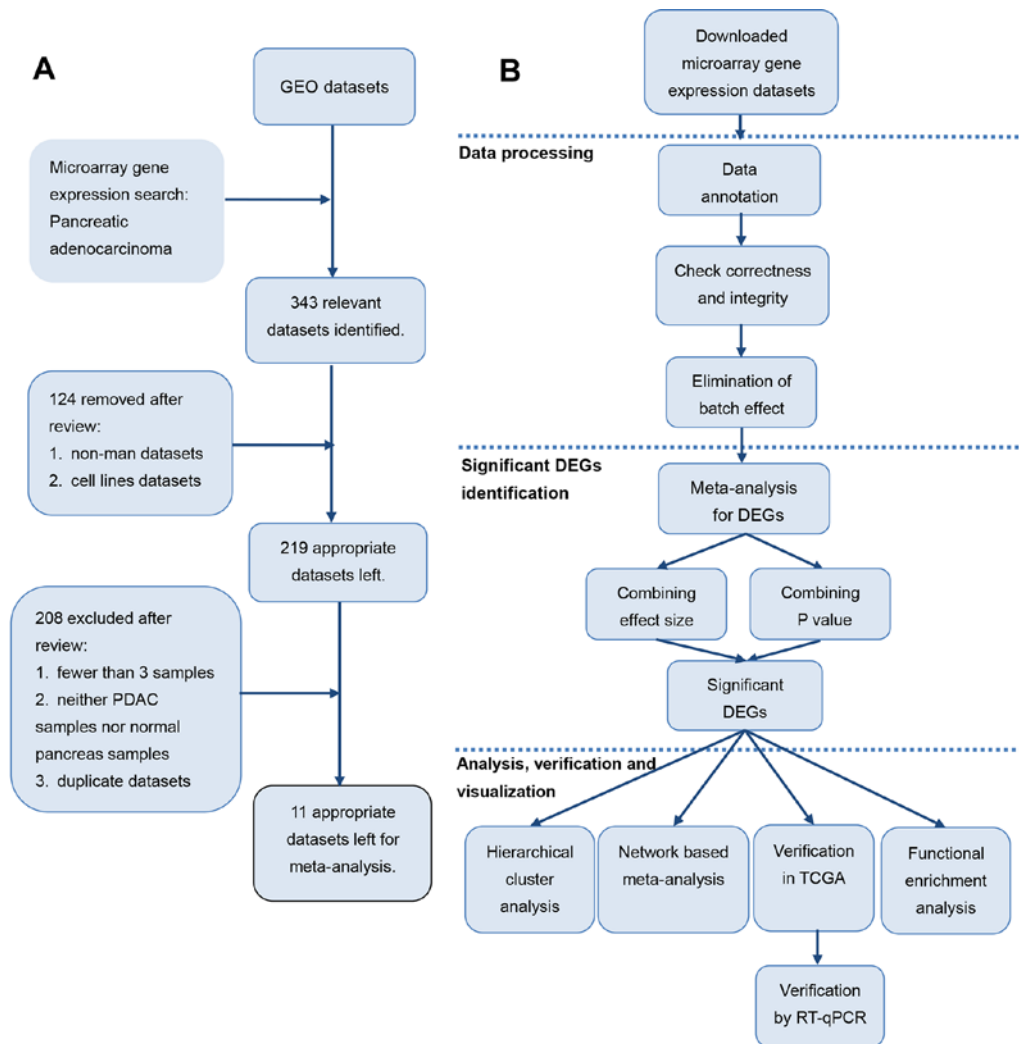


Figure 1. Study workflow. (A) Identification of eligible gene expression datasets for meta-analysis of PDAC. (B) Processing meta-analysis. DEGs, differentially expressed genes; GEO, Gene Expression Omnibus; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PDAC, pancreatic ductal adenocarcinoma; PPI, protein-protein interaction; TCGA, The Cancer Genome Atlas.

conditions were as follows: Initial denaturation at 95°C for 10 min, followed by 40 cycles of denaturation at 95°C for 15 sec and annealing/elongation at 60°C for 60 sec. The $2^{-\Delta\Delta C_q}$ quantification method was used for analyzing the qPCR data (34). The primers used are listed in Table I. The gene expression levels in each sample were normalized to *GAPDH*.

Statistical analysis. The ES combined random-effects model was used to perform the meta-analysis, and the threshold to screen for DEGs was set at $P < 0.05$. The Benjamini-Hochberg false discovery rate was used to correct the P-value to obtain more precise outcomes. The Mann-Whitney U test was used to conduct validation of gene expression in TCGA. *S100A6*, *KRT19* and *GNG7* expression results for the paired PDAC and normal samples were compared using paired Student's t-test or Wilcoxon test. $P < 0.05$ was considered to indicate a statistically significant difference.

Results

Meta-analysis. Raw data from the following 11 microarray gene expression profile datasets were downloaded from the

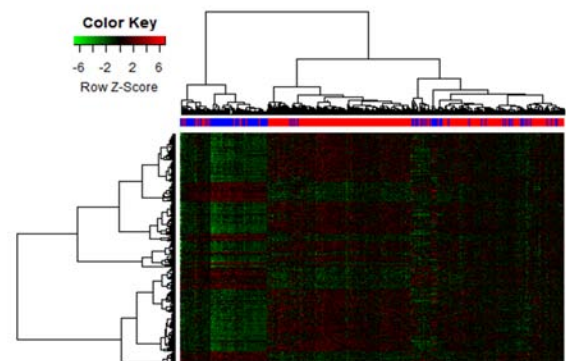


Figure 2. Two-way hierarchical clustering. Two-way hierarchical clustering based on 1,587 DEGs in PDAC vs. normal tissues across 11 datasets. PDAC samples, indicated in red label, and normal tissue, indicated in blue, were categorized into the two major clusters. PDAC, pancreatic ductal adenocarcinoma

GEO database: GSE15471 (35), GSE16515 (36), GSE22780, GSE28735 (37,38), GSE32676 (39), GSE43288-GPL96 (40), GSE43288-GPL97 (40), GSE43795 (41), GSE46234,

Table III. Top 20 differentially expressed genes identified in the meta-analysis of PDAC vs. normal tissues.

| A, Top 10 upregulated genes in PDAC vs. Normal tissue | | | | |
|---|---------------|--|-------------|--------------------------|
| Entrez ID | Gene | Gene name | Combined ES | Adjusted P-value |
| 51155 | <i>HNI</i> | Hematological and neurological expressed 1 | 2.4036 | 7.5536x10 ⁻¹⁰ |
| 3673 | <i>ITGA2</i> | Integrin subunit alpha 2 | 2.3434 | 5.0162x10 ⁻⁷ |
| 6277 | <i>S100A6</i> | S100 calcium binding protein A6 | 2.1598 | 1.1665x10 ⁻⁶ |
| 3624 | <i>INHBA</i> | Inhibin beta A subunit | 2.1409 | 2.4647x10 ⁻⁵ |
| 3880 | <i>KRT19</i> | Keratin 19 | 2.0556 | 5.215x10 ⁻⁵ |
| 129642 | <i>MBOAT2</i> | Membrane bound O-acyltransferase domain containing 2 | 2.046 | <0.001 |
| 1278 | <i>COL1A2</i> | Collagen type I alpha 2 chain | 2.0415 | 4.4806x10 ⁻⁵ |
| 1281 | <i>COL3A1</i> | Collagen type III alpha 1 chain | 1.993 | 1.0244x10 ⁻⁴ |
| 11343 | <i>MGLL</i> | Monoglyceride lipase | 1.8727 | 1.1349x10 ⁻¹² |
| 60 | <i>ACTB</i> | β actin | 1.8532 | 1.0806x10 ⁻⁴ |
| B, Top 10 downregulated genes in PDAC vs. Normal tissue | | | | |
| Entrez ID | Gene | Gene name | Combined ES | Adjusted P-value |
| 547 | <i>KIF1A</i> | Kinesin family member 1A | -1.9848 | 7.0263x10 ⁻⁷ |
| 54808 | <i>DYM</i> | Dymeclin | -1.8453 | 3.6218x10 ⁻⁷ |
| 23621 | <i>BACE1</i> | β-secretase 1 | -1.8387 | 1.302x10 ⁻⁴ |
| 4925 | <i>NUCB2</i> | Nucleobindin 2 | -1.8176 | 5.8813x10 ⁻⁷ |
| 29968 | <i>PSAT1</i> | Phosphoserine aminotransferase 1 | -1.809 | 6.8582x10 ⁻⁷ |
| 3977 | <i>LIFR</i> | LIF receptor alpha | -1.7681 | 1.2712x10 ⁻¹⁰ |
| 9024 | <i>BRSK2</i> | BR serine/threonine kinase 2 | -1.7642 | 6.1567x10 ⁻⁴ |
| 3777 | <i>KCNK3</i> | Potassium two pore domain channel subfamily K member 3 | -1.7365 | 1.1409x10 ⁻⁶ |
| 54855 | <i>FAM46C</i> | Family with sequence similarity 46 member C | -1.7315 | 4.01x10 ⁻¹¹ |
| 2788 | <i>GNG7</i> | G protein subunit gamma 7 | -1.6543 | 1.6178x10 ⁻⁶ |

Differentially expressed genes were ranked according to the combined ES. ES, effect size; PDAC, pancreatic ductal adenocarcinoma.

GSE55643 (42) and GSE62165 (43). In total, 485 samples were analyzed (334 tumor samples and 151 non-tumor samples); the two sample types were subjected to meta-analysis to identify DEGs. Specific information for each dataset included the first author, country, publication date, platform and numbers of PDAC and normal pancreatic tissue samples. The complete information and reference numbers of each dataset are presented in Table II. The detailed screening and processing workflow is depicted in Fig. 1.

Eliminating differences between batches. As differences between platforms lead to inaccuracies, biases in these datasets must be eliminated to increase the precision and reliability of the results. Therefore, prior to the meta-analysis, ComBat in INMEX was used to correct for differences between batches online. A principal component analysis plot demonstrated modest separation among the 11 datasets (data not shown).

Meta-analysis of gene expression in PDAC. DEGs with low but continuous expression levels in all profile datasets were

defined as gained genes, and DEGs that disappeared in the meta-analysis but were expressed in individual analyses or in the meta-analysis resulting from experimental errors in the platforms were defined as lost genes (20,44,45). In total, one gained gene, unkempt family zinc finger-like (*UNKL*), which was weakly expressed in all datasets (ES, -0.25395), and 1,278 lost genes were identified. Through the meta-analysis, 1,587 DEGs were identified, of which 1,004 were upregulated and 583 were downregulated (data not shown). Unsupervised hierarchical clustering of the DEGs was performed, and the results demonstrated that PDAC and normal samples were partitioned into two major groups (Fig. 2). Among all DEGs, the top 10 upregulated genes were hematological and neurological expressed 1 (*HNI*), integrin subunit α2 (*ITGA2*), S100 calcium-binding protein A6 (*S100A6*), inhibin βA subunit (*INHBA*), keratin-19 (*KRT19*), membrane-bound O-acyltransferase domain-containing 2 (*MBOAT2*), collagen type I α2 chain (*COL1A2*), collagen type III α1 chain (*COL3A1*), monoglyceride lipase (*MGLL*) and β-actin (*ACTB*) (Table III). The top 10 downregulated genes

Table IV. The top 10 hub genes in PDAC vs. Normal tissue.

| A, Upregulated genes | | | |
|----------------------|---------------|--------|-------------|
| ID | Gene | Degree | Betweenness |
| 2885 | <i>GRB2</i> | 90 | 53,725 |
| Q86VP6 | <i>CAND1</i> | 83 | 40,684 |
| Q13616 | <i>CUL1</i> | 82 | 43,269 |
| P08047 | <i>SP1</i> | 74 | 47,977 |
| P35222 | <i>CTNNB1</i> | 56 | 27,401 |
| P63279 | <i>UBE2I</i> | 45 | 24,406 |
| Q9UQL6 | <i>HDAC5</i> | 45 | 17,157 |
| Q9HCE7 | <i>SMURF1</i> | 41 | 16,525 |
| P22681 | <i>CBL</i> | 40 | 11,550 |
| Q92793 | <i>CREBBP</i> | 39 | 14,026 |

| B, Downregulated genes | | | |
|------------------------|---------------|--------|-------------|
| ID | Gene | Degree | Betweenness |
| Q9UQL6 | <i>HDAC5</i> | 45 | 17,157 |
| Q92793 | <i>CREBBP</i> | 39 | 14,026 |

CAND1, cullin-associated and neddylation-dissociated 1; *CBL*, Cbl proto-oncogene; *CREBBP*, CREB binding protein; *CTNNB1*, catenin β 1; *CUL1*, cullin 1; *GRB2*, growth factor receptor-bound protein 2; *HDAC5*, histone deacetylase 5; PDAC, pancreatic ductal adenocarcinoma; *SMURF1*, SMAD-specific E3 ubiquitin protein ligase 1; *SP1*, Sp1 transcription factor; *UBE2I*, ubiquitin-conjugating enzyme E2 I.

were kinesin family member 1A (*KIF1A*), Dymeclin (*DYM*), β -secretase 1 (*BACE1*), nucleobindin-2 (*NUCB2*), phosphoserine aminotransferase 1 (*PSAT1*), LIF receptor α (*LIFR*), BR serine/threonine kinase 2 (*BRSK2*), potassium two pore domain channel subfamily K member 3 (*KCNK3*), family with sequence similarity 46 member C (*FAM46C*) and G protein subunit γ 7 (*GNG7*) (Table III).

Identification of hub genes using network-based meta-analysis.

Hub genes serve key roles in interrelationships among DEGs. To identify hub genes among the PDAC DEGs, a PPI network analysis was conducted using NetworkAnalyst, which included comprehensive data curated from the literature by InnateDB. One large subnetwork comprising 8,527 nodes and 37,579 edges, and a smaller subnetwork containing 3 nodes and 2 edges were obtained, and the top 10 hub genes in PDAC vs. normal tissue were identified (Table IV). To display the results more clearly, a zero-edge network analysis was performed, which comprised 965 nodes and 2,806 edges (Fig. 3). Finally, hub genes in the network were ranked by degree. Growth factor receptor-bound protein 2 (*GRB2*), with a combined ES of 0.76104 and an adjusted P-value of 0.0053254, had the highest degree (90) and betweenness (53,725) among upregulated DEGs. Histone deacetylase 5 (*HDAC5*), with a combined ES of -0.37624 and an adjusted P-value of 0.0018425, had the highest degree (45) and betweenness (17,157) among downregulated DEGs.

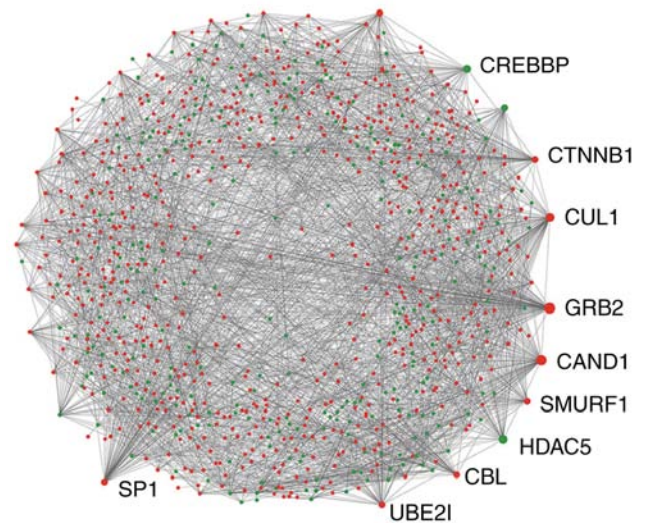


Figure 3. Identification of hub genes using network-based meta-analysis. 'Zero order' interaction network for DEGs by meta-analysis of PDAC in comparison with normal tissue in a forced atlas layout. *CAND1*, cullin-associated and neddylation-dissociated 1; *CBL*, Cbl proto-oncogene; *CREBBP*, CREB binding protein; *CTNNB1*, catenin β 1; *CUL1*, cullin 1; *GRB2*, growth factor receptor-bound protein 2; *HDAC5*, histone deacetylase 5; *SMURF1*, SMAD-specific E3 ubiquitin protein ligase 1; *SP1*, Sp1 transcription factor; *UBE2I*, ubiquitin-conjugating enzyme E2 I.

Functional annotation and pathway enrichment analyses. To examine the functions of the identified DEGs, the DEGs were mapped in DAVID to perform GO term and KEGG pathway analyses. The following three aspects were included in the GO analyses: Biological process (BP), cellular component (CC) and molecular function (MF). In the GO analyses, DEGs identified between normal and PDAC tissues were notably enriched in 'cell-cell adhesion' (BP), 'cytosol' (CC) and 'protein binding' (MF) (Fig. 4A-C). In total, 27 KEGG pathways were significantly enriched ($P < 0.05$) between normal and PDAC tissues. The top two enriched pathways with the lowest P value were 'ubiquitin-mediated proteolysis' ($P = 7.40 \times 10^{-9}$) and 'pathways in cancer' ($P = 6.23 \times 10^{-5}$) (Fig. 4D), which demonstrated that these DEGs may be strongly associated with pancreatic cancer occurrence.

TCGA validation. The top 10 upregulated and downregulated DEGs obtained by meta-analysis were validated in TCGA using 147 PDAC samples and four normal samples. The results revealed that compared with the expression levels in normal samples, *SI00A6* and *KRT19* were strongly upregulated in the PDAC samples (Fig. 5A and B). Compared with the expression levels in normal samples, *LIFR* and *GNG7* were significantly downregulated in the PDAC samples (Fig. 5C and D).

Kaplan-Meier analysis. The TCGA dataset comprised 157 pancreatic cancer cases with detailed clinical and prognostic information and gene expression data. To determine the effects of the identified DEGs on the survival time of patients with PDAC, Kaplan-Meier analysis was used for the top 10 upregulated DEGs and the top 10 downregulated DEGs among the PDAC samples in the TCGA dataset. The results revealed that patients with high *ITGA2* expression levels had shorter survival times compared with patients with lower *ITGA2* expression

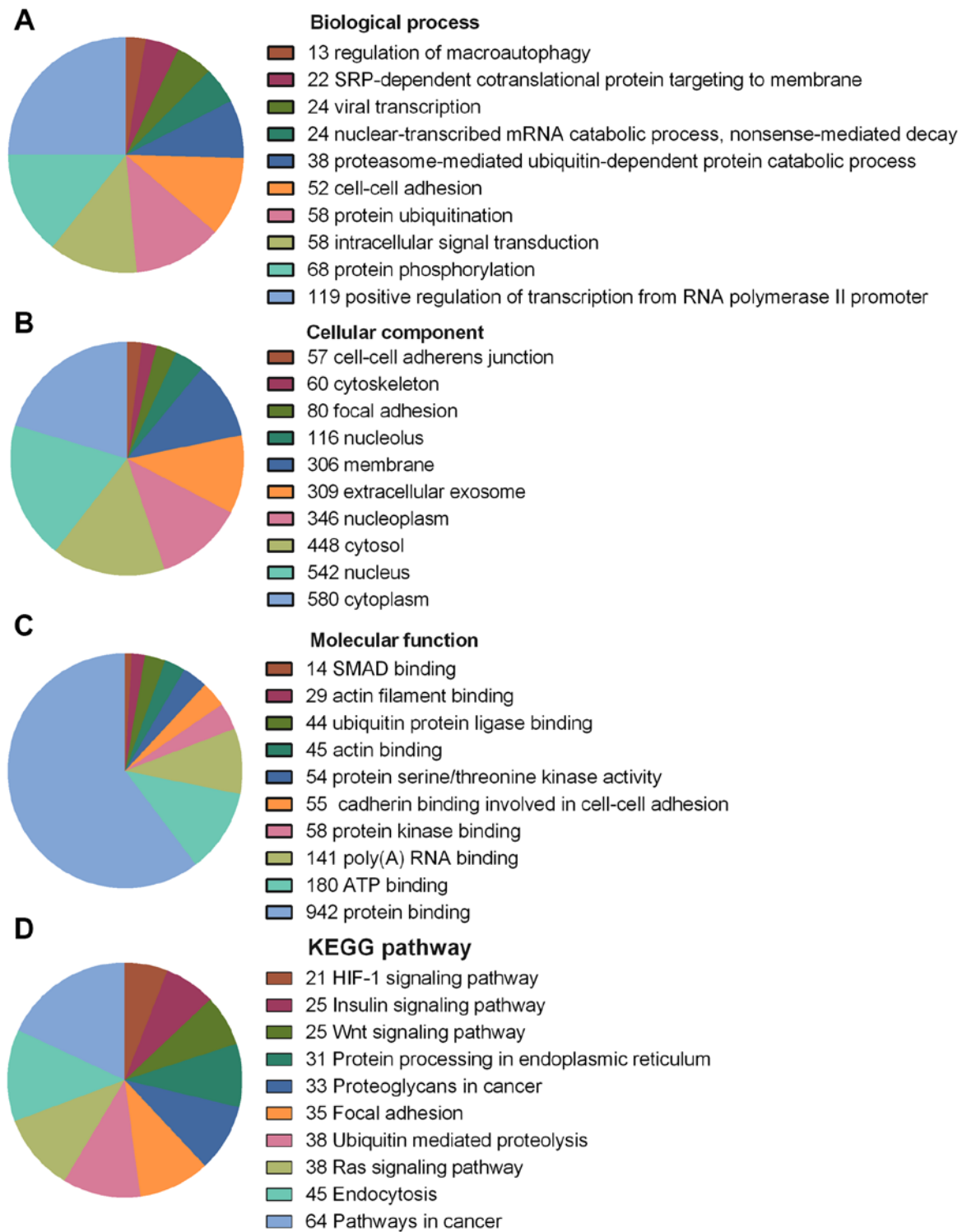


Figure 4. GO term and KEGG pathway analyses of differentially expressed genes. GO term analyses of (A) biological process, (B) cellular component and (C) molecular function. (D) KEGG pathway analysis. GO, Gene Ontology; HIF-1, hypoxia-inducible factor 1; KEGG, Kyoto Encyclopedia of Genes and Genomes; PDAC, pancreatic ductal adenocarcinoma; SMAD, Smad family of genes.

levels (Fig. 6A). *KRT19*, *MBOAT2* and *MGLL* exhibited similar results to *ITGA2* (Fig. 6C-E). By contrast, patients with high *FAM46C* expression had longer survival times compared with patients with low *FAM46C* expression (Fig. 6G). These findings further demonstrated that high expression levels of upregulated DEGs, such as *ITGA2*, *KRT19*, *MBOAT2* and *MGLL*, may represent a risk factor for poor survival in PDAC.

By contrast, high expression levels of downregulated DEGs, such as *FAM46C*, may positively influence survival times of patients with PDAC.

Validation of genes with significantly different expression in TCGA by RT-qPCR. To validate *S100A6*, *KRT19* and *GNG7* expression, RT-qPCR was performed using paired normal

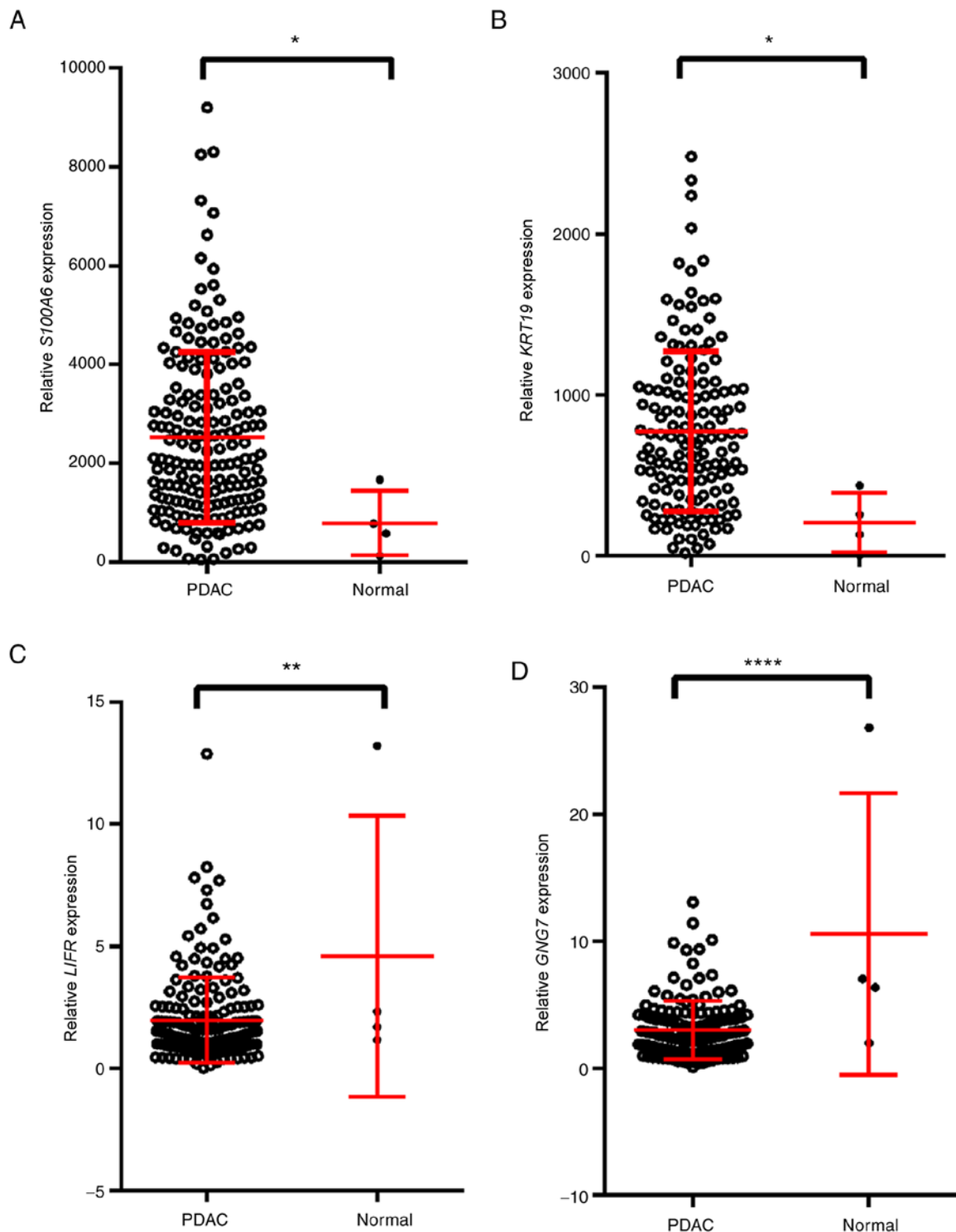


Figure 5. TCGA dataset validation of *S100A6*, *KRT19*, *LIFR* and *GNG7*. (A) *S100A6* and (B) *KRT19* were upregulated genes. (C) *LIFR* and (D) *GNG7* were downregulated genes. * $P < 0.05$; ** $P < 0.01$; **** $P < 0.0001$. *GNG7*, G protein subunit γ 7; *KRT19*, keratin-19; *LIFR*, LIF receptor α ; PDAC, pancreatic ductal adenocarcinoma; *S100A6*, S100 calcium-binding protein; TCGA, The Cancer Genome Atlas.

pancreatic tissues and PDAC tissues, which were collected from 15 patients. Among the three genes, *S100A6* was significantly upregulated in PDAC tissues compared with normal pancreatic tissues (Fig. 7A). By contrast, the expression of *KRT19* and *GNG7* (Fig. 7B and C) PDAC did not significantly differ from normal tissues, although enhanced/low expression levels of these genes were detected in some PDAC tissues.

Discussion

With the rapid development of high-throughput genomic technology, researchers have explored gene mutations in pancreatic cancer over the past two decades, including oncogenes such as *KRAS*, *SMAD4*, proto-oncogene *c-Myc* and *RAD51* recombinase (46), as well as tumor suppressor genes such as cyclin-dependent kinase inhibitor (*CDKN2A*)

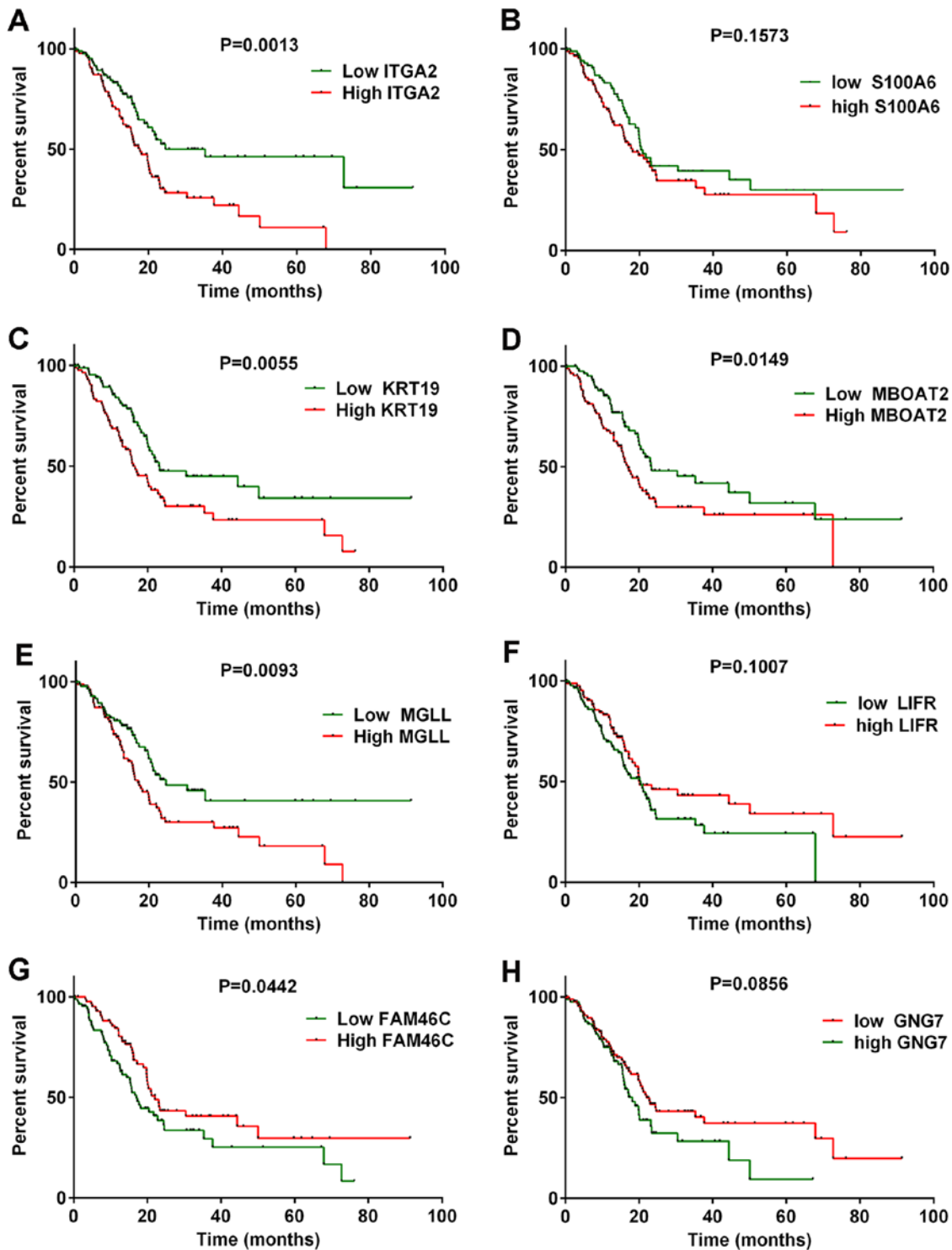


Figure 6. Kaplan-Meier survival curve analysis. Kaplan-Meier analysis of overall survival for patients with PDAC with high expression levels of upregulated and downregulated genes. (A) *ITGA2*, (B) *S100A6*, (C) *KRT19*, (D) *MBOAT2* and (E) *MGLL* were upregulated genes. (F) *LIFR*, (G) *FAM46C* and (H) *GNG7* were downregulated genes. Red lines represent high expression of DEGs and green lines represent low expression of DEGs. The censored subjects are marked with ticks on the Kaplan-Meier survival curves. DEG, differentially expressed gene; *FAM46C*, family with sequence similarity 46 member C; *GNG7*, G protein subunit 7; *ITGA2*, integrin subunit $\alpha 2$; *KRT19*, keratin-19; *LIFR*, LIF receptor; *MBOAT2*, membrane-bound O-acyltransferase domain-containing 2; *MGLL*, monoglyceride lipase; PDAC, pancreatic ductal adenocarcinoma; *S100A6*, S100 calcium-binding protein.

and *p53* (6). Although these studies initially identified several pancreatic cancer-related gene mutations, DEGs derived from individual studies may be unreliable due to the limited quality and quantity of microarray chips. Identifying specific genes that can be used as biomarkers to diagnose pancreatic cancer

or for targeted therapies is the key to reducing pancreatic cancer-related death and improving pancreatic cancer prognosis. Tang *et al* identified 205 pancreatic cancer-related genes (142 upregulated and 63 downregulated) by screening and merging DEGs from four transcriptome microarray datasets

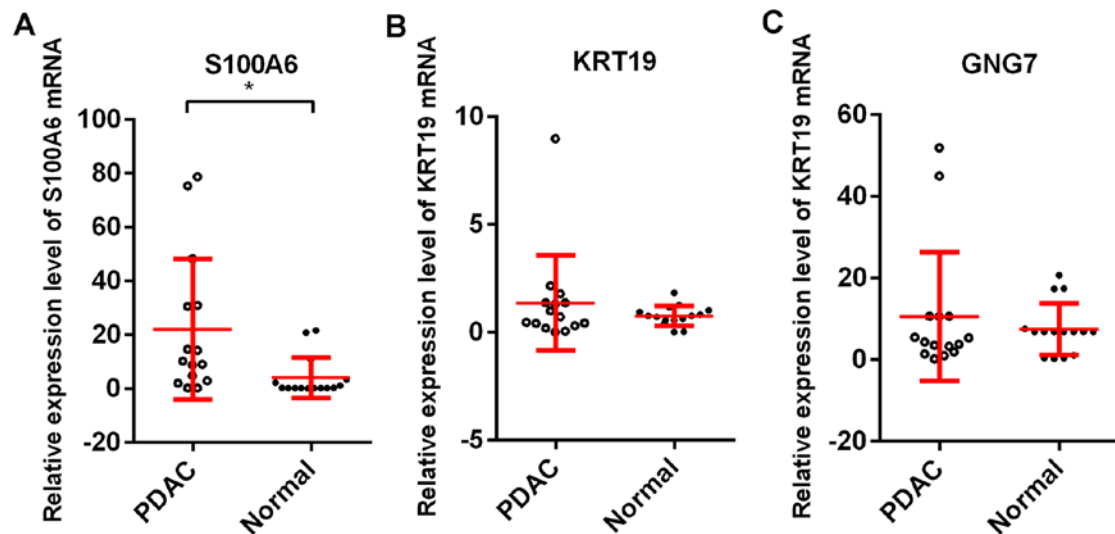


Figure 7. Enhanced expression of *S100A6*, *KRT19* and *GNG7* mRNA in PDAC. Scatter plots illustrate the relative expression of (A) *S100A6*, (B) *KRT19* and (C) *GNG7* mRNA normalized to GAPDH mRNA in each sample. * $P < 0.05$. *GNG7*, G protein subunit 7; *KRT19*, keratin-19; PDAC, pancreatic ductal adenocarcinoma; *S100A6*, S100 calcium-binding protein.

(226 PDAC samples and 65 normal pancreatic tissue samples) in the GEO database (27). Through functional analysis, including GO, KEGG pathway, Kaplan-Meier survival and PPI network analyses, the authors identified *DKK1* and *HMGA2* as candidate genes in PDAC. Notably, these two genes were not identified as DEGs in the present study. In the present study, a meta-analysis was performed using 11 high-quality microarray datasets (334 tumor samples and 151 normal samples) from the GEO database to identify reliable PDAC-related DEGs; 1,587 DEGs were identified, including 1,004 upregulated and 583 downregulated genes. One gained gene, *UNKL*, was identified with a combined ES value of -0.25395. This gained gene was previously unreported. Additionally, based on the PPI network, *GRB2* and *HDAC5* were identified as pivotal hub genes in PDAC. In addition to the aforementioned functional analysis by Tang *et al* (27), the expression of several top genes was verified by TCGA analysis and RT-qPCR.

Among the upregulated genes, *HNI* had the highest combined ES value of 2.4036. The present study was consistent with a previous study by Tang *et al* (27). In addition, many studies have confirmed that *HNI* mutations are related to cancer occurrence. For example, Lu *et al* reported that *HNI* is mutated in ovarian carcinoma epithelial tissue and can be used as a biomarker for ovarian cancer (47). Laughlin *et al* established a mouse glioma model and confirmed that murine *HNI* depletion led to a significant reduction in glioma volume (48). They also demonstrated that *HNI* was associated with the melanoma phenotype and regulated cell proliferation and melanogenesis (48). *HNI* is overexpressed in breast cancer and can promote invasion and metastasis in breast cancer (49) and prostate cancer cells (50). These studies illustrate that *HNI* gene expression is related to the proliferation, invasion and metastasis of tumor cells. The results of the bioinformatics analysis in the present study demonstrated that *HNI* was upregulated in PDAC samples, which indicated that PDAC occurrence may be related to *HNI* overexpression. This result is consistent with the previously studied mechanism of action of *HNI* (49). Conversely, RT-qPCR analysis of *HNI* did not demonstrate a significant

difference between PDAC samples and normal pancreatic tissue. Despite the lack of a statistical difference between PDAC patients with high levels and low *HNI* expression levels in the TCGA dataset, since the present study suggested that *HNI* expression is associated with PDAC, further exploration of the association between *HNI* gene expression, PDAC progression and patient survival is important.

ITGA2 is a collagen receptor expressed on cell membranes that mediates cell-cell and cell-matrix adhesions. A number of previous studies have indicated that *ITGA2* expression is upregulated in normal epithelial cells, and its expression changes when tumorigenesis occurs (51-53). Ramirez *et al* revealed a loss of ITG $\alpha 2\beta 1$ in breast and prostate cancer (51). Previous studies have also reported that loss of *ITGA2* serves a critical role in colon cancer metastasis (52) and that *ITGA2* is strongly expressed in PDAC (53). The results of the present study are consistent with previous studies on pancreatic cancer in that *ITGA2* expression was significantly higher in PDAC tissue compared with normal pancreatic tissue.

S100A6, or its product calcyclin, has been identified to be associated with various malignancies, such as osteosarcoma (54), gastric cancer (55) and malignant melanoma (56). Upregulation of *S100A6* expression in tumor tissue is associated with cell proliferation (57). In addition, Ohuchida *et al* quantitatively analyzed *S100A6* and demonstrated that the *S100A6* expression level was markedly higher in pancreatic cancer tissues compared with normal pancreatic tissue (58). *S100A6* is one of the 158 PDAC-related DEGs identified by Logsdon *et al* using microarray datasets (59). In the present study, *S100A6* expression was upregulated in PDAC samples compared with normal tissue, which was verified by TCGA and RT-qPCR; the findings were consistent with previous studies. Unfortunately, no association was observed between *S100A6* gene expression and patient survival time in the TCGA dataset.

In the present study, *KIF1A* was identified as the most strongly downregulated gene, with a combined ES of -1.9848. The protein encoded by *KIF1A* is microtubule-dependent molecular motor involved in important intracellular functions

such as organelle transport and cell division (60). De *et al* (61) found that *KIF1A* is highly expressed in breast cancer and its upregulation is associated with breast cancer resistance to docetaxel. By contrast, Hattori *et al* demonstrated that *KIF1A* expression was upregulated in human and murine adrenal tumors compared with normal adrenal tissue (62). However, few studies have shown that the PDAC occurrence is related to the downregulation of *KIF1A*, although mutation of *KIF1A* is strongly associated with cancer (63). In the present study, *KIF1A* was the most strongly downregulated DEG in PDAC, further indicating that *KIF1A* may be related to the disease.

DYM encodes a Golgi-related protein transported within cells, and previous studies have shown that *DYM* loss of function is closely related to microcephaly in Dyggve-Melchior-Clausen syndrome (64,65). In the present study, *DYM* was the second most downregulated gene, with a combined ES value of -1.8453, which suggested that it may have potential protective functions against PDAC. *KIF1A* and *DYM* expression levels were validated by TCGA, although no significant differences between PDAC and normal tissues were observed. In addition, RT-qPCR analysis of *KIF1A* and *DYM* was performed multiple times. The expression of these two genes could not be detected in some normal or PDAC tissues, potentially owing to sub-threshold expression levels of these two genes, or high sensitivity of the genes to pancreatin, leading to a certain extent of degradation. Additional specimens need to be collected for further study.

Overexpression of *BACE1*, which hydrolyzes transmembrane amyloid precursor protein into amyloid- β protein, has been demonstrated to serve an important role in Alzheimer's disease (66). Excessive amyloid- β protein aggregation damages neurons (67). Chen *et al* reported that the long non-coding RNA *BACE1* antisense RNA (*BACE1*-AS) may be used as a novel target of anisomycin to inhibit the proliferation and invasion of ovarian cancer stem cells; the authors also demonstrated that the increased *BACE1* levels caused by silencing of *BACE1*-AS resulted in poor cancer suppression by anisomycin (68). Therefore, *BACE1* may be a tumor suppressor. In the present study, *BACE1* was downregulated in PDAC compared with normal pancreatic tissues. Nonetheless, the relationship between *BACE1* and cancer, especially in PDAC, requires further study.

The top 10 upregulated and downregulated genes were validated by performing a meta-analysis on a dataset from TCGA. The results demonstrated that high expression levels of *ITGA2*, *KRT19*, *MBOAT2*, and *MGLL* were associated with shorter survival times, whereas high *FAM46C* expression was associated with longer survival times. Dong *et al* (69) reported that *ITGA2* is associated with the tumor, node and metastasis classification system and staging of gastric cancer, and that the expression level of *ITGA2* was increased in metastatic lymph nodes and distant metastases. They demonstrated that increased *ITGA2* levels were associated with reduced overall survival rates in patients with gastric cancer. Previous studies have shown that *KRT19* is a marker associated with metastasis and poor prognosis of pancreatic ductal adenocarcinoma (70,71). Badea *et al* (35) combined gene expression analysis of whole-tissue and microdissected PDAC, and found that upregulation of *MBOAT2* is inversely correlated with patient survival. Transcriptomic analyses identified that *MGLL* overexpression is an unfavorable prognostic marker in

primary gastrointestinal stromal tumors (72). The results of Caba *et al* (73) suggested that *FAM46C* was downregulated in patients with PDAC and could be used as prognostic indicator in these patients. The present results are consistent with previous studies and suggested that these genes may be novel targets for diagnosing and treating PDAC.

Using NetworkAnalyst for PPI network analysis, *GRB2*, with a degree of 90, and *HDAC5*, with a degree of 45, were identified as two hub genes among the upregulated and downregulated DEGs, respectively. The *GRB2* protein contains one Src homology 2 domain and is a linker protein in the tyrosine kinase receptor signaling pathway (74). *GRB2* expression directly affects cell proliferation and differentiation (75). Liang *et al* found that *GRB2* expression was higher in hepatocellular carcinoma compared with normal liver tissue and that high *GRB2* expression levels were associated with shorter survival times (76). This is consistent with the *GRB2* expression profile in PDAC described in the present study, although the expression level of *GRB2* did not appear to be significantly associated with overall survival rates in the present study. In addition, a previous study has confirmed that *Grb2* may be used as a therapeutic target for pancreatic cancer (77). Wang *et al* demonstrated that *GRB2* is a target of miR-329 through the *GRB2/pERK* pathway, which may lead to novel therapies for pancreatic cancer (78). The results of these studies verify the accuracy and credibility of the results of the present study and further demonstrate that *GRB2* is a potential target for treating pancreatic cancer.

Histone deacetylases are epigenetic regulators that suppress transcription by acting on gene promoters. *HDAC1*, *HDAC2*, *HDAC3* and *HDAC8* are targets of multiple epigenetic inhibitors, such as the REST corepressor 1, the nucleosome remodeling deacetylase and the SIN3 histone deacetylase complexes (79). A number of studies have reported high *HDAC5* expression in various tumors, including medulloblastomas (GSE15471, GSE16515, GSE22780, GSE28735, GSE32676, GSE43288-GPL96, GSE43288-GPL97, GSE43795, GSE46234, GSE55643 and GSE62165.) (80), breast cancer (81) and pancreatic neuroendocrine tumors (82). In addition, *HDAC5* overexpression is associated with tumor cell proliferation and invasion, as well as poor prognosis (80-82). He *et al* demonstrated that *HDAC5* is overexpressed in and promotes the proliferation of colorectal cancer cells by upregulating delta-like canonical Notch ligand 4 expression (83). However, Özdağ *et al* reported the opposite finding that, with the exception of high *HDAC5* expression levels in two samples of rectal cancer, *HDAC5* expression in all other rectal cancers was markedly downregulated and that *HDAC5* expression could be used to distinguish between rectal tumors and normal tissues (84). These results are similar to the results of the present study, which demonstrated that *HDAC5* was downregulated in PDAC tissues compared with normal tissues. Overall, findings to date demonstrate that *HDAC5* may serve a key role in tumorigenesis, but its role in PDAC requires further study.

In conclusion, 1,587 DEGs were identified in the present meta-analysis. *HNI*, *ITGA2* and *S100A6* were the top upregulated genes and may be promising potential targets for diagnosing and treating PDAC. *HNI*, which had the highest combined ES, has rarely been reported as being important in PDAC; thus, this gene should be further investigated. *KIF1A*, *DYM* and *BACE1* were the most markedly downregulated

genes in this study, providing new perspectives regarding PDAC mechanisms and targeted treatment. *GRB2* and *HDAC5* were identified as hub genes, serving the most pivotal roles in PDAC. However, a lack of adequate validation *in vitro* or *in vivo* is a limitation of this study. As the number of PDAC cases in the TCGA database is limited, the present results need to be confirmed by further experiments. Therefore, future research will include experimental verification of the meta-analysis results using immunohistochemistry and cell proliferation assays, and identification of genes associated with PDAC progression by screening differentially expressed genes at different stages.

Acknowledgements

Not applicable.

Funding

The present study was supported by The National Natural Science Foundation of China (grant nos. 81371661 and 81873895).

Availability of data and materials

The datasets used and analyzed during the present study are available from GEO GSE15471, GSE16515, GSE22780, GSE28735, GSE32676, GSE43288-GPL96, GSE43288-GPL97, GSE43795, GSE46234, GSE55643 and GSE62165 (<http://www.ncbi.nlm.nih.gov/geo>) and TCGA (<https://tcga-data.nci.nih.gov/tcga/>).

Authors' contributions

LL analyzed the results. LL and SW wrote the manuscript. PH and SW designed the study. CC performed the experiments. SP, YC and XL conceived the study and perform statistical analyses. ND, QL and LM collected and analyzed the data. All authors approved the manuscript for publication and agree to be accountable for all aspects of the research.

Ethics approval and consent to participate

Human study protocols were approved by the ethics committee of Tongji Medical College, Huazhong University of Science and Technology (Wuhan, China), in accordance with the Declaration of Helsinki (reference no. S298). All patients in the study provided written informed consent.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

- Siegel RL, Miller KD and Jemal A: Cancer Statistics, 2017. *CA Cancer J Clin* 67: 7-30, 2017.
- Crane CH, Varadhachary GR, Yordy JS, Staerckel GA, Javle MM, Safran H, Haque W, Hobbs BD, Krishnan S, Fleming JB, *et al*: Phase II trial of cetuximab, gemcitabine, and oxaliplatin followed by chemoradiation with cetuximab for locally advanced (T4) pancreatic adenocarcinoma: Correlation of Smad4(Dpc4) immunostaining with pattern of disease progression. *J Clin Oncol* 29: 3037-3043, 2011.
- Hezel AF, Kimmelman AC, Stanger BZ, Bardeesy N and Depinho RA: Genetics and biology of pancreatic ductal adenocarcinoma. *Genes Dev* 20: 1218-1249, 2006.
- Gillen S, Schuster T, Meyer Zum Büschenfelde C, Friess H and Kleeff J: Preoperative/neoadjuvant therapy in pancreatic cancer: A systematic review and meta-analysis of response and resection percentages. *PLoS Med* 7: e1000267, 2010.
- Stratton MR, Campbell PJ and Futreal PA: The cancer genome. *Nature* 458: 719-724, 2009.
- Jones S, Zhang X, Parsons DW, Lin JC, Leary RJ, Angenendt P, Mankoo P, Carter H, Kamiyama H, Jimeno A, *et al*: Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science* 321: 1801-1806, 2008.
- Nagata K, Horinouchi M, Saitou M, Higashi M, Nomoto M, Goto M and Yonezawa S: Mucin expression profile in pancreatic cancer and the precursor lesions. *J Hepatobiliary Pancreat Surg* 14: 243-254, 2007.
- Sausen M, Phallen J, Adloff V, Jones S, Leary RJ, Barrett MT, Anagnostou V, Parpart-Li S, Murphy D, Kay Li Q, *et al*: Clinical implications of genomic alterations in the tumour and circulation of pancreatic cancer patients. *Nat Commun* 6: 7686, 2015.
- Rozenblum E, Schutte M, Goggins M, Hahn SA, Panzer S, Zahurak M, Goodman SN, Sohn TA, Hruban RH, Yeo CJ and Kern SE: Tumor-suppressive pathways in pancreatic carcinoma. *Cancer Res* 57: 1731-1734, 1997.
- Bardeesy N, Cheng KH, Berger JH, Chu GC, Pahler J, Olson P, Hezel AF, Horner J, Lauwers GY, Hanahan D and DePinho RA: Smad4 is dispensable for normal pancreas development yet critical in progression and tumor biology of pancreas cancer. *Genes Dev* 20: 3130-3146, 2006.
- Fong ZV and Winter JM: Biomarkers in pancreatic cancer: Diagnostic, prognostic, and predictive. *Cancer J* 18: 530-538, 2012.
- Poruk KE, Gay DZ, Brown K, Mulvihill JD, Boucher KM, Scaife CL, Firpo MA and Mulvihill SJ: The clinical utility of CA 19-9 in pancreatic adenocarcinoma: Diagnostic and prognostic updates. *Curr Mol Med* 13: 340-351, 2013.
- Kim SM, Kwon IJ, Myoung H, Lee JH and Lee SK: Identification of human papillomavirus (HPV) subtype in oral cancer patients through microarray technology. *Eur Arch Otorhinolaryngol* 275: 535-543, 2018.
- Gunel T, Hosseini MK, Gumusoglu E, Kisakesen HI, Benian A and Aydinli K: Expression profiling of maternal plasma and placenta microRNAs in preeclamptic pregnancies by microarray technology. *Placenta* 52: 77-85, 2017.
- Qiao X, Wang H, Wang X, Zhao B and Liu J: Microarray technology reveals potentially novel genes and pathways involved in non-functioning pituitary adenomas. *Balkan J Med Genet* 19: 5-16, 2016.
- Li G, Li X, Yang M, Xu L, Deng S and Ran L: Prediction of biomarkers of oral squamous cell carcinoma using microarray technology. *Sci Rep* 7: 42105, 2017.
- Li J, Chen Z, Tian L, Zhou C, He MY, Gao Y, Wang S, Zhou F, Shi S, Feng X, *et al*: LncRNA profile study reveals a three-lncRNA signature associated with the survival of patients with oesophageal squamous cell carcinoma. *Gut* 63: 1700-1710, 2014.
- Ramasamy A, Mondry A, Holmes CC and Altman DG: Key issues in conducting a meta-analysis of gene expression microarray datasets. *PLoS Med* 5: e184, 2008.
- Rung J and Brazma A: Reuse of public genome-wide gene expression data. *Nat Rev Genet* 14: 89-99, 2013.
- Wang S, Jin F, Fan W, Liu F, Zou Y, Hu X, Xu H and Han P: Gene expression meta-analysis in diffuse low-grade glioma and the corresponding histological subtypes. *Sci Rep* 7: 11741, 2017.
- Zhang Y, Xia Q and Lin J: Identification of the potential oncogenes in glioblastoma based on bioinformatic analysis and elucidation of the underlying mechanisms. *Oncol Rep* 40: 715-725, 2018.
- Xiao W, Pacyna-Gengelbach M, Schluns K, An Q, Gao Y, Cheng S and Petersen I: Differentially expressed genes associated with human lung cancer. *Oncol Rep* 14: 229-234, 2005.
- Tang F, He Z, Lei H, Chen Y, Lu Z, Zeng G and Wang H: Identification of differentially expressed genes and biological pathways in bladder cancer. *Mol Med Rep* 17: 6425-6434, 2018.

24. Peng C, Ma W, Xia W and Zheng W: Integrated analysis of differentially expressed genes and pathways in triplenegative breast cancer. *Mol Med Rep* 15: 1087-1094, 2017.
25. Sun W, Ma X, Shen J, Yin F, Wang C and Cai Z: Bioinformatics analysis of differentially expressed pathways related to the metastatic characteristics of osteosarcoma. *Int J Mol Med* 38: 466-474, 2016.
26. Hoshida Y, Nijman SM, Kobayashi M, Chan JA, Brunet JP, Chiang DY, Villanueva A, Newell P, Ikeda K, Hashimoto M, *et al*: Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. *Cancer Res* 69: 7385-7392, 2009.
27. Tang Y, Zhang Z, Tang Y, Chen X and Zhou J: Identification of potential target genes in pancreatic ductal adenocarcinoma by bioinformatics analysis. *Oncol Lett* 16: 2453-2461, 2018.
28. Amin MB, Edge S, Greene F, Byrd DR, Brookland RK, Washington MK, Gershenwald JE, Compton CC, Hess KR, Sullivan DC, *et al* (eds): *AJCC Cancer Staging Manual*. 8th edition. Springer International Publishing, New York, NY, 2017.
29. Xia J, Fjell CD, Mayer ML, Pena OM, Wishart DS and Hancock RE: INMEX-a web-based tool for integrative meta-analysis of expression data. *Nucleic Acids Res* 41: W63-W70, 2013.
30. Johnson WE, Li C and Rabinovic A: Adjusting batch effects in microarray expression data using empirical bayes methods. *Biostatistics* 8: 118-127, 2007.
31. Cochran WG: The combination of estimates from different experiments. *Biometrics* 10: 101-129, 1954.
32. Marot G, Foulley JL, Mayer CD and Jaffrézic F: Moderated effect size and P-value combinations for microarray meta-analyses. *Bioinformatics* 25: 2692-2699, 2009.
33. Breuer K, Foroushani AK, Laird MR, Chen C, Sribnaia A, Lo R, Winsor GL, Hancock RE, Brinkman FS and Lynn DJ: InnateDB: Systems biology of innate immunity and beyond-recent updates and continuing curation. *Nucleic Acids Res* 41: D1228-D1233, 2013.
34. Livak KJ and Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25: 402-408, 2001.
35. Badea L, Herlea V, Dima SO, Dumitrascu T and Popescu I: Combined gene expression analysis of whole-tissue and microdissected pancreatic ductal adenocarcinoma identifies genes specifically overexpressed in tumor epithelia. *Hepatogastroenterology* 55: 2016-2027, 2008.
36. Pei H, Li L, Fridley BL, Jenkins GD, Kalari KR, Lingle W, Petersen G, Lou Z and Wang L: FKBP51 affects cancer cell response to chemotherapy by negatively regulating Akt. *Cancer Cell* 16: 259-266, 2009.
37. 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.
38. 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.
39. Donahue TR, Tran LM, Hill R, Li Y, Kovoichich A, Calvopina JH, Patel SG, Wu N, Hindoyan A, Farrell JJ, *et al*: Integrative survival-based molecular profiling of human pancreatic cancer. *Clin Cancer Res* 18: 1352-1363, 2012.
40. Crnogorac-Jurcevic T, Chelala C, Barry S, Harada T, Bhakta V, Lattimore S, Jurcevic S, Bronner M, Lemoine NR and Brentnall TA: Molecular analysis of precursor lesions in familial pancreatic cancer. *PLoS One* 8: e54830, 2013.
41. Park M, Kim M, Hwang D, Park M, Kim WK, Kim SK, Shin J, Park ES, Kang CM, Paik YK and Kim H: Characterization of gene expression and activated signaling pathways in solid-pseudopapillary neoplasm of pancreas. *Mod Pathol* 27: 580-593, 2014.
42. Lunardi S, Jamieson NB, Lim SY, Griffiths KL, Carvalho-Gaspar M, Al-Assar O, Yameen S, Carter RC, McKay CJ, Spoletini G, *et al*: IP-10/CXCL10 induction in human pancreatic cancer stroma influences lymphocytes recruitment and correlates with poor survival. *Oncotarget* 5: 11064-11080, 2014.
43. 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.
44. Jha PK, Vijay A, Sahu A and Ashraf MZ: Comprehensive gene expression meta-analysis and integrated bioinformatic approaches reveal shared signatures between thrombosis and myeloproliferative disorders. *Sci Rep* 6: 37099, 2016.
45. Wang X, Ning Y and Guo X: Integrative meta-analysis of differentially expressed genes in osteoarthritis using microarray technology. *Mol Med Rep* 12: 3439-3445, 2015.
46. Han H, Bearss DJ, Browne LW, Calaluze R, Nagle RB and Von Hoff DD: Identification of differentially expressed genes in pancreatic cancer cells using cDNA microarray. *Cancer Res* 62: 2890-2896, 2002.
47. Lu KH, Patterson AP, Wang L, Marquez RT, Atkinson EN, Baggerly KA, Ramoth LR, Rosen DG, Liu J, Hellstrom I, *et al*: Selection of potential markers for epithelial ovarian cancer with gene expression arrays and recursive descent partition analysis. *Clin Cancer Res* 10: 3291-3300, 2004.
48. Laughlin KM, Luo D, Liu C, Shaw G, Warrington KH Jr, Qiu J, Yachnis AT and Harrison JK: Hematopoietic- and neurologic-expressed sequence 1 expression in the murine GL261 and high-grade human gliomas. *Pathol Oncol Res* 15: 437-444, 2009.
49. Zhang C, Xu B, Lu S, Zhao Y and Liu P: HN1 contributes to migration, invasion, and tumorigenesis of breast cancer by enhancing MYC activity. *Mol Cancer* 16: 90, 2017.
50. Varisli L, Ozturk BE, Akuz GK and Korkmaz KS: HN1 negatively influences the β -catenin/E-cadherin interaction, and contributes to migration in prostate cells. *J Cell Biochem* 116: 170-178, 2015.
51. Ramirez NE, Zhang Z, Madamanchi A, Boyd KL, O'Rear LD, Nashabi A, Li Z, Dupont WD, Zijlstra A and Zutter MM: The α 2 β 1 integrin is a metastasis suppressor in mouse models and human cancer. *J Clin Invest* 121: 226-237, 2011.
52. Robertson JH, Yang SY, Winslet MC and Seifalian AM: Functional blocking of specific integrins inhibit colonic cancer migration. *Clin Exp Metastasis* 26: 769-780, 2009.
53. Shimoyama S, Gansauge F, Gansauge S, Oohara T and Beger HG: Altered expression of extracellular matrix molecules and their receptors in chronic pancreatitis and pancreatic adenocarcinoma in comparison with normal pancreas. *Int J Pancreatol* 18: 227-234, 1995.
54. Li Y, Wagner ER, Yan Z, Wang Z, Luther G, Jiang W, Ye J, Wei Q, Wang J, Zhao L, *et al*: The calcium-binding protein S100A6 accelerates human osteosarcoma growth by promoting cell proliferation and inhibiting osteogenic differentiation. *Cell Physiol Biochem* 37: 2375-2392, 2015.
55. Wang XH, Zhang LH, Zhong XY, Xing XF, Liu YQ, Niu ZJ, Peng Y, Du H, Zhang GG, Hu Y, *et al*: S100A6 overexpression is associated with poor prognosis and is epigenetically up-regulated in gastric cancer. *Am J Pathol* 177: 586-597, 2010.
56. Maelandsmo GM, Florenes VA, Mellingsaeter T, Hovig E, Kerbel RS and Fodstad O: Differential expression patterns of S100A2, S100A4 and S100A6 during progression of human malignant melanoma. *Int J Cancer* 74: 464-469, 1997.
57. He X, Xu X, Khan AQ and Ling W: High expression of S100A6 predicts unfavorable prognosis of lung squamous cell cancer. *Med Sci Monit* 23: 5011-5017, 2017.
58. Ohuchida K, Mizumoto K, Ishikawa N, Fujii K, Konomi H, Nagai E, Yamaguchi K, Tsuneyoshi M and Tanaka M: The role of S100A6 in pancreatic cancer development and its clinical implication as a diagnostic marker and therapeutic target. *Clin Cancer Res* 11: 7785-7793, 2005.
59. Logsdon CD, Simeone DM, Binkley C, Arumugam T, Greenon JK, Giordano TJ, Misek DE, Quic R and Hanah S: Molecular profiling of pancreatic adenocarcinoma and chronic pancreatitis identifies multiple genes differentially regulated in pancreatic cancer. *Cancer Res* 63: 2649-2657, 2003.
60. Okada Y, Yamazaki H, Sekine-Aizawa Y and Hirokawa N: The neuron-specific kinesin superfamily protein KIF1A is a unique monomeric motor for anterograde axonal transport of synaptic vesicle precursors. *Cell* 81: 769-780, 1995.
61. De S, Cipriano R, Jackson MW and Stark GR: Overexpression of kinesins mediates docetaxel resistance in breast cancer cells. *Cancer Res* 69: 8035-8042, 2009.
62. Hattori Y, Kanamoto N, Kawano K, Iwakura H, Sone M, Miura M, Yasoda A, Tamura N, Arai H, Akamizu T, *et al*: Molecular characterization of tumors from a transgenic mouse adrenal tumor model: Comparison with human pheochromocytoma. *Int J Oncol* 37: 695-705, 2010.

63. Brait M, Loyo M, Rosenbaum E, Ostrow KL, Markova A, Papagerakis S, Zahurak M, Goodman SM, Zeiger M, Sidransky D, *et al*: Correlation between BRAF mutation and promoter methylation of TIMP3, RARBeta2 and RASSF1A in thyroid cancer. *Epigenetics* 7: 710-719, 2012.
64. Osipovich AB, Jennings JL, Lin Q, Link AJ and Ruley HE: Dyggve-Melchior-Clausen syndrome: Chondrodysplasia resulting from defects in intracellular vesicle traffic. *Proc Natl Acad Sci U S A* 105: 16171-16176, 2008.
65. Dimitrov A, Paupe V, Guedry C, Sibarita JB, Raposo G, Vielemeyer O, Gilbert T, Csaba Z, Attie-Bitach T, Cormier-Daire V, *et al*: The gene responsible for Dyggve-Melchior-Clausen syndrome encodes a novel peripheral membrane protein dynamically associated with the Golgi apparatus. *Hum Mol Genet* 18: 440-453, 2009.
66. Shi Z, Hong Y, Zhang K, Wang J, Zheng L, Zhang Z, Hu Z, Han X, Han Y, Chen T, *et al*: BAG-1M co-activates BACE1 transcription through NF- κ B and accelerates A β production and memory deficit in Alzheimer's disease mouse model. *Biochim Biophys Acta Mol Basis Dis* 1863: 2398-2407, 2017.
67. Yi X, Hao Y, Xia N, Wang J, Quintero M, Li D and Zhou F: Sensitive and continuous screening of inhibitors of β -site amyloid precursor protein cleaving enzyme 1 (BACE1) at single SPR chips. *Anal Chem* 85: 3660-3666, 2013.
68. Chen Q, Liu X, Xu L, Wang Y, Wang S, Li Q, Huang Y and Liu T: Long non-coding RNA BACE1-AS is a novel target for anisomycin-mediated suppression of ovarian cancer stem cell proliferation and invasion. *Oncol Rep* 35: 1916-1924, 2016.
69. Dong J, Wang R, Ren G, Li X, Wang J, Sun Y, Liang J, Nie Y, Wu K, Feng B, *et al*: HMGA2-FOXL2 axis regulates metastases and epithelial-to-mesenchymal transition of chemoresistant gastric cancer. *Clin Cancer Res* 23: 3461-3473, 2017.
70. Yao H, Yang Z, Liu Z, Miao X, Yang L, Li D, Zou Q and Yuan Y: Glypican-3 and KRT19 are markers associating with metastasis and poor prognosis of pancreatic ductal adenocarcinoma. *Cancer Biomark* 17: 397-404, 2016.
71. Dugnani E, Sordi V, Pellegrini S, Chimienti R, Marzinotto I, Pasquale V, Liberati D, Balzano G, Doglioni C, Reni M, *et al*: Gene expression analysis of embryonic pancreas development master regulators and terminal cell fate markers in resected pancreatic cancer: A correlation with clinical outcome. *Pancreatology* 18: 945-953, 2018.
72. Li CF, Chuang IC, Liu TT, Chen KC, Chen YY, Fang FM, Li SH, Chen TJ, Yu SC, Lan J and Huang HY: Transcriptomic reappraisal identifies MGLL overexpression as an unfavorable prognosticator in primary gastrointestinal stromal tumors. *Oncotarget* 7: 49986-49997, 2016.
73. Caba O, Irigoyen A, Jimenez-Luna C, Benavides M, Ortuño FM, Gallego J, Rojas I, Guillen-Ponce C, Torres C, Aranda E and Prados J: Identification of gene expression profiling associated with erlotinib-related skin toxicity in pancreatic adenocarcinoma patients. *Toxicol Appl Pharmacol* 311: 113-116, 2016.
74. Wintgens JP, Wichert SP, Popovic L, Rossner MJ and Wehr MC: Monitoring activities of receptor tyrosine kinases using a universal adapter in genetically encoded split TEV assays. *Cell Mol Life Sci* 76: 1185-1199, 2019.
75. Dharmawardana PG, Peruzzi B, Giubellino A, Burke TR Jr and Bottaro DP: Molecular targeting of growth factor receptor-bound 2 (Grb2) as an anti-cancer strategy. *Anticancer Drugs* 17: 13-20, 2006.
76. Liang C, Xu Y, Ge H, Xing B, Li G, Li G and Wu J: miR-564 inhibits hepatocellular carcinoma cell proliferation and invasion by targeting the GRB2-ERK1/2-AKT axis. *Oncotarget* 8: 107543-107557, 2017.
77. Birchmeier C, Birchmeier W, Gherardi E and Vande Woude GF: Met, metastasis, motility and more. *Nat Rev Mol Cell Biol* 4: 915-925, 2003.
78. Wang X, Lu X, Zhang T, Wen C, Shi M, Tang X, Chen H, Peng C, Li H, Fang Y, *et al*: mir-329 restricts tumor growth by targeting grb2 in pancreatic cancer. *Oncotarget* 7: 21441-21453, 2016.
79. Witt O, Deubzer HE, Milde T and Oehme I: HDAC family: What are the cancer relevant targets? *Cancer Lett* 277: 8-21, 2009.
80. Milde T, Oehme I, Korshunov A, Kopp-Schneider A, Remke M, Northcott P, Deubzer HE, Lodrini M, Taylor MD, von Deimling A, *et al*: HDAC5 and HDAC9 in medulloblastoma: Novel markers for risk stratification and role in tumor cell growth. *Clin Cancer Res* 16: 3240-3252, 2010.
81. Li A, Liu Z, Li M, Zhou S, Xu Y, Xiao Y and Yang W: HDAC5, a potential therapeutic target and prognostic biomarker, promotes proliferation, invasion and migration in human breast cancer. *Oncotarget* 7: 37966-37978, 2016.
82. Klieser E, Urbas R, Stättner S, Primavesi F, Jäger T, Dinnewitzer A, Mayr C, Kiesslich T, Holzmann K, Di Fazio P, *et al*: Comprehensive immunohistochemical analysis of histone deacetylases in pancreatic neuroendocrine tumors: HDAC5 as a predictor of poor clinical outcome. *Hum Pathol* 65: 41-52, 2017.
83. He P, Liang J, Shao T, Guo Y, Hou Y and Li Y: HDAC5 promotes colorectal cancer cell proliferation by up-regulating DLL4 expression. *Int J Clin Exp Med* 8: 6510-6516, 2015.
84. Ozdağ H, Teschendorff AE, Ahmed AA, Hyland SJ, Blenkiron C, Bobrow L, Veerakumarasivam A, Burt G, Subkhankulova T, Arends MJ, *et al*: Differential expression of selected histone modifier genes in human solid cancers. *BMC Genomics* 7: 90, 2006.