Original Article

Identification of novel hub genes and lncRNAs related to the prognosis and progression of pancreatic cancer by microarray and integrated bioinformatics analysis

Xing Liang[1#]^, Junfeng Peng[1#], Danlei Chen[1#], Liang Tang[1#], Anan Liu[1], Zhiping Fu[1], Ligang Shi[1], Keqi Wang[2], Chenghao Shao[1]

[1]Department of Pancreatic-biliary Surgery, Changzheng Hospital, Second Military Medical University, Shanghai, China; [2]Department of Gastroenterology, Changzheng Hospital, Second Military Medical University, Shanghai, China

Contributions: (I) Conception and design: X Liang; (II) Administrative support: C Shao; (III) Provision of study materials or patients: D Chen; (IV) Collection and assembly of data: J Peng; (V) Data analysis and interpretation: K Wang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Chenghao Shao. Department of Pancreatic-biliary Surgery, Changzheng Hospital, Second Military Medical University, Fengyang Road 415, Shanghai 200003, China. Email: shaochenghao_czyy@163.com; Keqi Wang. Department of Gastroenterology, Changzheng Hospital, Second Military Medical University, Shanghai 200003, China. Email: wangkeqi1001@163.com.

Background: Pancreatic cancer (PC) is one of the most invasive and metastatic neoplasms among the fatal malignancies of the digestive system. Abnormal expression of genes and long non-coding RNAs (lncRNAs) are reportedly linked to multiple cancers. However, the lncRNA-mRNA expression profiles and their molecular mechanisms in PC progression are poorly known. This study aimed to map the hub genes and lncRNAs which might play core roles in the development of PC.

Methods: This study used microarray expression analysis to screen for both differentially expressed genes (DEGs) and differentially expressed lncRNAs (DElncRNAs) between PC and matched adjacent non-tumor (AN) tissues. In order to clarify the functional classification of DEGs, we conducted GO and KEGG pathway enrichment analyses via the Enrichr database. LncRNA-mRNA co-expressed networks were also constructed to explore the probable core regulating DEGs and DElncRNAs. Subsequently, the hub genes and lncRNAs were validated via the ONCOMINE and GEPIA databases and the co-expressed networks.

Results: By analyzing an mRNA-lncRNA microarray, we identified 943 mRNAs and 1,138 lncRNAs differentially expressed in PC tumors compared with the matched AN tissues. GO analysis confirmed that both up-regulated and down-regulated DEGs were enriched in multiple terms. The KEGG pathways enrichment analyses revealed that DEGs were mostly enriched in the focal adhesion and glutathione metabolism pathways, amongst others. Co-expressed networks were established to reveal the differential interactions between DEGs and DElncRNAs, and to indicate the core regulatory factors located at the core nodes of the co-expressed networks. The expression levels of potential core-regulating DEGs were validated by the GEPIA and ONCOMINE databases, and the relationship between overall survival and tumor stage and the potential core-regulating DEGs was analyzed using the GEPIA database. As a result, five genes and sixteen lncRNAs were finally considered as the hub transcripts in PC.

Conclusions: This study identified DEGs and DElncRNAs between PC tumors and matched AN tissues, and these transcripts were connected with malignant phenotypes in PC through different BPs and signaling pathways. Furthermore, five hub genes and sixteen lncRNAs were identified, which are expected to represent candidate diagnostic biomarkers or potential therapeutic targets for PC.

^ ORCID: 0000-0003-4406-0753.

© Gland Surgery. All rights reserved.

Keywords: Pancreatic cancer (PC); long non-coding RNA (lncRNA); bioinformatics analysis; prognosis

Submitted Jan 27, 2021. Accepted for publication Mar 22, 2021.
doi: 10.21037/gs-21-151
View this article at: http://dx.doi.org/10.21037/gs-21-151

Introduction

Pancreatic cancer (PC) is one of the most aggressive human digestive cancers and has an extremely poor prognosis, demonstrating a 5-year survival rate of only 10% (1). Despite this, there has been little progress in the diagnosis and treatment of PC in the past two decades. Due to the lack of early symptoms and diagnosis, approximately 80% of PC patients lose the opportunity to have radical surgery because of the primary tumor's invasion of the adjacent organs or surrounding vital vessels, and many patients exhibit extensive metastasis even at their first diagnosis (2). In spite of continuing research, the mechanisms of PC development remain unclear. Therefore, there is an urgent need to explore the regulating mechanisms in PC proliferation, invasion, and metastasis, and to identify novel effective prognostic and therapeutic targets for PC.

Emerging evidence has confirmed that, in regard to hub genes, many long non-coding RNAs (lncRNAs) may also play core roles in PC (3-5) and other malignant phenotypes of multiple tumors and could represent novel targets for diagnosis and treatment (6-8).

In this study, we used an mRNA-lncRNA microarray expression analysis to determine differentially expressed genes (DEGs) and differentially expressed lncRNAs (DElncRNAs) in PC. To detect hub genes and hub lncRNAs in PC, we utilized a bioinformatics approach to analyze those DEGs and DElncRNAs. Using the Enrichr database, Gene Ontology (GO) analysis was conducted for the functional annotation assessment, and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was conducted for the enrichment assessment. LncRNA-mRNA co-expressed networks were constructed to explore the core regulating factors. The GEPIA and ONCOMINE databases were then used to identify hub genes which were associated with both overall survival and tumor stage in PC. Ultimately, we found five hub genes and sixteen hub lncRNAs that were strongly linked to the development and progression of PC. In this study, we used commercial microarrays and integrated bioinformatics analysis to screen out hub DEGs and DElncRNAs, which most of them never been reported before in PC. Different from co-expressed networks conducted by online databases, we established lncRNA-mRNA co-expressed networks in PC and paired adjacent non-tumor samples also considering their differential expressions, resulting a more convinced result. Furthermore, via these bioinformatics analysis, we map hub DElncRNAs as potential ceRNAs of hub genes, which can indicating our next experimental direction.

We present the following article in accordance with the MDAR reporting checklist (available at http://dx.doi.org/10.21037/gs-21-151).

Methods

Tissue samples

A total of five pairs of PC and paired adjacent non-tumor (AN) samples were obtained from PC patients who received surgery between 2012 and 2013 in the Department of Pancreatic Surgery at Changhai Hospital, Second Military Medical University, Shanghai, China. Among the PC patients, three were male and two were female, aging from 46 to 72. None of the patients in this study had received any neoadjuvant anti-tumor therapies. Two certified pathologists confirmed all diagnoses, which were based on pathological evidence. All samples were frozen immediately in liquid nitrogen with a TRIzol reagent (Invitrogen) after the tumor was removed surgically and were then transferred to a −80 ℃ refrigerator until required.

This study was approved by the Ethics Committee for Biomedical Research of the Second Military Medical University. Informed consent was obtained from all patients or their relatives. All procedures performed in this study involving human participants were in accordance with the Declaration of Helsinki (as revised in 2013).

LncRNA-mRNA microarrays

The frozen PC samples were delivered to KangChen Biotechnology Co. Ltd (Shanghai, China) for the mRNAs and lncRNAs extraction, quantity testing, labeling and

© Gland Surgery. All rights reserved.

analysis. In order to profile both human mRNAs and lncRNAs, the Arraystar Human mRNA-LncRNA Array v2.0 was designed. Array images were screened and analyzed using the Agilent Feature Extraction software (version 10.7.3.1), and acquired data was normalized and processed by the GeneSpring GX v11.5.1 software package (Agilent Technologies). LncRNAs and mRNAs, which were flagged as present or marginal in no less than half of the samples, were identified for further analysis. DElncRNAs and DEGs between the PC tissues and matched AN samples were confirmed through fold change filtering. Paired t-tests were performed to compare the mean differences. P-value/FDR and Volcano Plot filtering were used to identify the statistical significance of the DElncRNAs and DEGs between the PC tissues and matched AN samples. Hierarchical clustering analysis was conducted by the Cluster 3.0 and JAVA Treeview software packages. P<0.05 and $|\log_2 FC|$ >1 were considered statistically significant.

Enrichr database analysis

The Enrichr database (http://amp.pharm.mssm.edu/Enrichr) was used to perform the GO functional annotation and KEGG pathways enrichment analysis of the DEGs (9,10). Biological process (BP), cellular component (CC), and molecular function (MF) were included in the GO functional annotation. The top ten enriched GO functional annotations and KEGG pathways of DEGs were displayed as images downloaded from web pages. P<0.05 was considered statistically significant.

LncRNA-mRNA co-expressed network establishment

LncRNA-mRNA co-expressed networks in both PC and AN tissues were constructed to detect the correlations between DEGs and DElncRNAs (11) according to their normalized signal intensity of specific expression. The Pearson correlation was calculated for each pair of mRNA-lncRNA, mRNA-mRNA, and lncRNA-lncRNA (12). As the simplest and most important measure of an mRNA or lncRNA, indicating the centrality and relative importance in a network, degree centrality is defined as the number of associations one profile has with the others (13). Core regulatory DEGs and DElncRNAs are considered to possess the greatest degree differences between the PC and AN networks. Results were ranked by the absolute relative degree value |DiffK| (relative degree value, PC-AN, DiffK).

ONCOMINE database analysis

The expressions of potential core genes were assessed in the PC tissues compared to those in normal pancreatic tissues in the ONCOMINE (https://www.oncomine.org/) datasets. ONCOMINE is the world's largest microarray database and integrated data-mining platform of neoplasms at present. P<0.05 was considered statistically significant.

GEPIA database analysis

The Gene Expression Profiling Interactive Analysis (GEPIA, http://gepia.cancer-pku.cn/about.html) database is a newly developed interactive web server for analyzing RNA sequencing expression data of tumors and normal tissues obtained from The Cancer Genome Atlas (TCGA) and Genoytype-Tissue Expression (GTEx) databases. GEPIA now contains a total of 9,736 tumors and 8,587 normal samples (14). The GEPIA database was used to validate the expression levels of significant DEGs in PC. Cut-off criteria were determined as P<0.05 and $|\log_2 FC|$ >1. GEPIA was also used to identify the relationship between DEGs expressing levels and PC patients' pathological stages and overall survival. P<0.05 was considered statistically significant.

Statistical analysis

Statistical analysis was done based on the online databases respectively. In lncRNA-mRNA microarrays, paired t-tests were performed to compare the mean differences. P-value/FDR and Volcano Plot filtering were used to identify the statistical significance of the DElncRNAs and DEGs between the PC tissues and matched AN samples. P<0.05 and $|\log_2 FC|$ >1 were considered statistically significant.

Results

Identification of DEGs and lncRNAs

To explore the DEGs and DElncRNAs in PC, we had previously collected five pairs of PC samples from patients who had undergone surgery, and we used microarrays to analyze the DEGs and DElncRNAs between the PC tissues and paired AN tissues. (GEO, http://www.ncbi.nlm.nih.gov/geo/, ID: GSE101094). There were 943 mRNAs differentially expressed, of which 501 were up-regulated and 442 were down-regulated (fold change >2, P<0.05, https://cdn.amegroups.cn/static/public/10.21037gs-21-151-1.xlsx). We confirmed that 1,138 lncRNA transcripts

© Gland Surgery. All rights reserved.

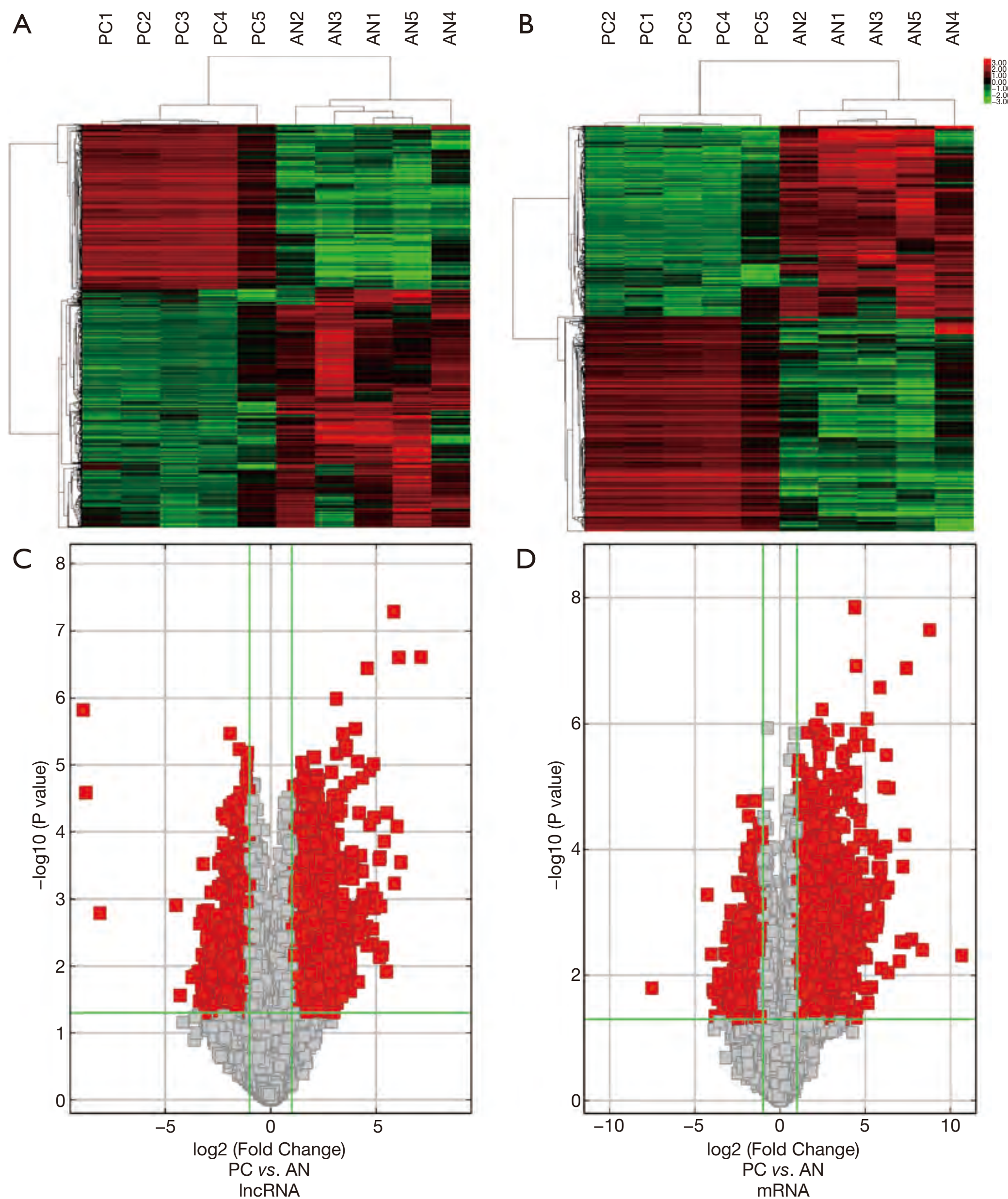

Figure 1 LncRNA and mRNA expression profile changes between pancreatic cancer tissues and matched adjacent non-tumor tissues. Hierarchical clustering of DElncRNAs (A) and DEGs (B) between PC tissues and matched AN tissues. Volcano plots of lncRNAs (C) and mRNAs (D) expression levels between PC tissues and matched AN tissues. DEG, differentially expressed gene; DElncRNA, differentially expressed long non-coding RNA; PC, pancreatic cancer; AN, adjacent non-tumor.

were differentially expressed, comprising 470 up-regulated and 668 down-regulated lncRNAs (fold change >2, P<0.05, https://cdn.amegroups.cn/static/public/10.21037gs-21-151-2.xlsx).

Hierarchical clustering analysis showed that the DEGs and DElncRNAs between the PC tissues and paired AN tissues were systematically variated (*Figure 1A,B*). In addition, volcano plots were constructed to identify the DEG and DElncRNA transcripts (*Figure 1C,D*). The results suggested that the DEGs and DElncRNAs were likely representative and most probably played a role in regulating the development of PC.

© Gland Surgery. All rights reserved.

GO annotation and KEGG pathway enrichment analyses

To determine the potential biological roles of the above 943 DEGs, we used the online Enrichr database to conduct the GO annotation and KEGG pathway enrichment analyses. The annotated results for the GO analyses contained three parts: BP, MF, and CC. GO BP analysis revealed that the top three significantly enriched terms of the up-regulated DEGs were adrenal gland development (GO:0030325), epithelial cell migration (GO:0010631), and peptide hormone secretion (GO:0030072); the top three enriched terms of the down-regulated DEGs were positive regulation of MAPK cascade (GO:0043410), organic anion transport (GO:0015711), and regulation of calcium ion transport (GO:0051924). In the GO MF analysis, the 501 up-regulated DEGs were mainly enriched in nuclear hormone receptor binding (GO:0035257), transmembrane receptor protein serine/threonine kinase binding (GO:0070696), and cadherin binding (GO:0045296); the 442 down-regulated DEGs were mainly enriched in sodium-independent organic anion transmembrane transporter activity (GO:0015347), organic anion transmembrane transporter activity (GO:0008514), and G-protein coupled photoreceptor activity (GO:0008020). In the GO CC analysis, the up-regulated DEGs were mostly enriched in focal adhesion (GO:0005925), the cytoplasmic vesicle membrane (GO:0030659), and the endoplasmic reticulum lumen (GO:0005788); the down-regulated DEGs were mostly enriched in the integral component of the plasma membrane (GO:0005887), the endocytic vesicle (GO:0030139), and the spanning component of the plasma membrane (GO:0044214) (*Figure 2A,B,C,D,E,F*).

The top three markedly enriched KEGG pathways for the 501 up-regulated DEGs were focal adhesion, human papillomavirus infection, and dilated cardiomyopathy (DCM); the 442 down-regulated DEGs were mostly enriched in glutathione metabolism, purine metabolism, and the PPAR signaling pathways (*Figure 2G,H*).

LncRNA-mRNA co-expressed network establishment and hub gene identification

In order to locate the core regulatory factors, we constructed lncRNA-mRNA co-expressed networks of PC tissues and AN tissues (*Figure 3A,B*, https://cdn.amegroups.cn/static/public/10.21037gs-21-151-3.xlsx, https://cdn.amegroups.cn/static/public/10.21037gs-21-151-4.xlsx). In a co-expressed network, core functional transcripts mostly have more relationships with other factors than nonfunctional transcripts, and always have the largest absolute relative degree values. Thus, mRNAs with the top 100 largest absolute values of relative degree differences between the PC and AN lncRNA-mRNA co-expressed networks (https://cdn.amegroups.cn/static/public/10.21037gs-21-151-5.xlsx) were selected as the potential hub genes for further study.

We compared these transcriptional levels of genes in PC tissues and normal pancreatic tissues by using the ONCOMINE and GEPIA databases, and identified that 22 mRNA levels, among the top 100 genes confirmed above, were significantly differentially expressed between PC and normal tissues, with the same expressed trend in our microarrays. They were *SFTA2*, *SPG21*, *COL17A1*, *TM4SF4*, *PPP2R5E*, *NR2F2*, *HOXA13*, *GSTA2*, *GGT5*, *C15orf57*, *PBX4*, *STOM*, *GJA1*, *RTN1*, *CHRAC1*, *TBRG4*, *NVL*, *GMEB1*, *CLPTM1*, *CXorf56*, *GIMAP2*, and *SERPINB9* (*Tables 1,2*, *Figure 4*). Among them, *HOXA13* and *NVL* were only identified as differentially expressed in the ONCOMINE database, and *PPP2R5E*, *GSTA2*, *C15orf57*, *PBX4*, *CHRAC1*, *GMEB1*, *CLPTM1*, and *CXorf56* were only confirmed to be differentially regulated in the GEPIA database.

The relationship of hub genes with overall survival and tumor stage

We next used the GEPIA database to study the association between overall survival and the 22 genes identified above by the median and quartile of their expression (Figure S1). We noticed that high levels of *SFTA2* (median: P=0.0089, HR =1.7; Quartile: P=0.00042, HR =3), *SPG21* (median: P=0.015, HR =1.7; quartile: P=0.00042, HR =1.7), *COL17A1* (median: P=0.00039, HR =2.1; quartile: P=0.00052, HR =2.8), *PPP2R5E* (quartile: P=0.027, HR =1.9), *HOXA13* (median: P=0.007, HR =1.8), and *CHRAC1* (quartile: P=0.027, HR =1.9) were associated with worse overall survival in PC. Conversely, high levels of *RTN1* (median: P=0.0039, HR =0.55; quartile: P=0.0024, HR =0.41) and *CLPTM1* (median: P=0.0081, HR =0.57; quartile: P=0.013, HR =0.47) were linked with improved overall survival in PC.

The GEPIA pathological stage plot analysis revealed that the levels of *SFTA2*, *COL17A1*, *HOXA13*, *PBX4*, *RTN1*, *CHRAC1*, *GMEB1*, and *CXorf56* varied significantly according to the tumor stage of the PC (Figure S1).

As a result of these findings, we considered *SFTA2*,

© Gland Surgery. All rights reserved.

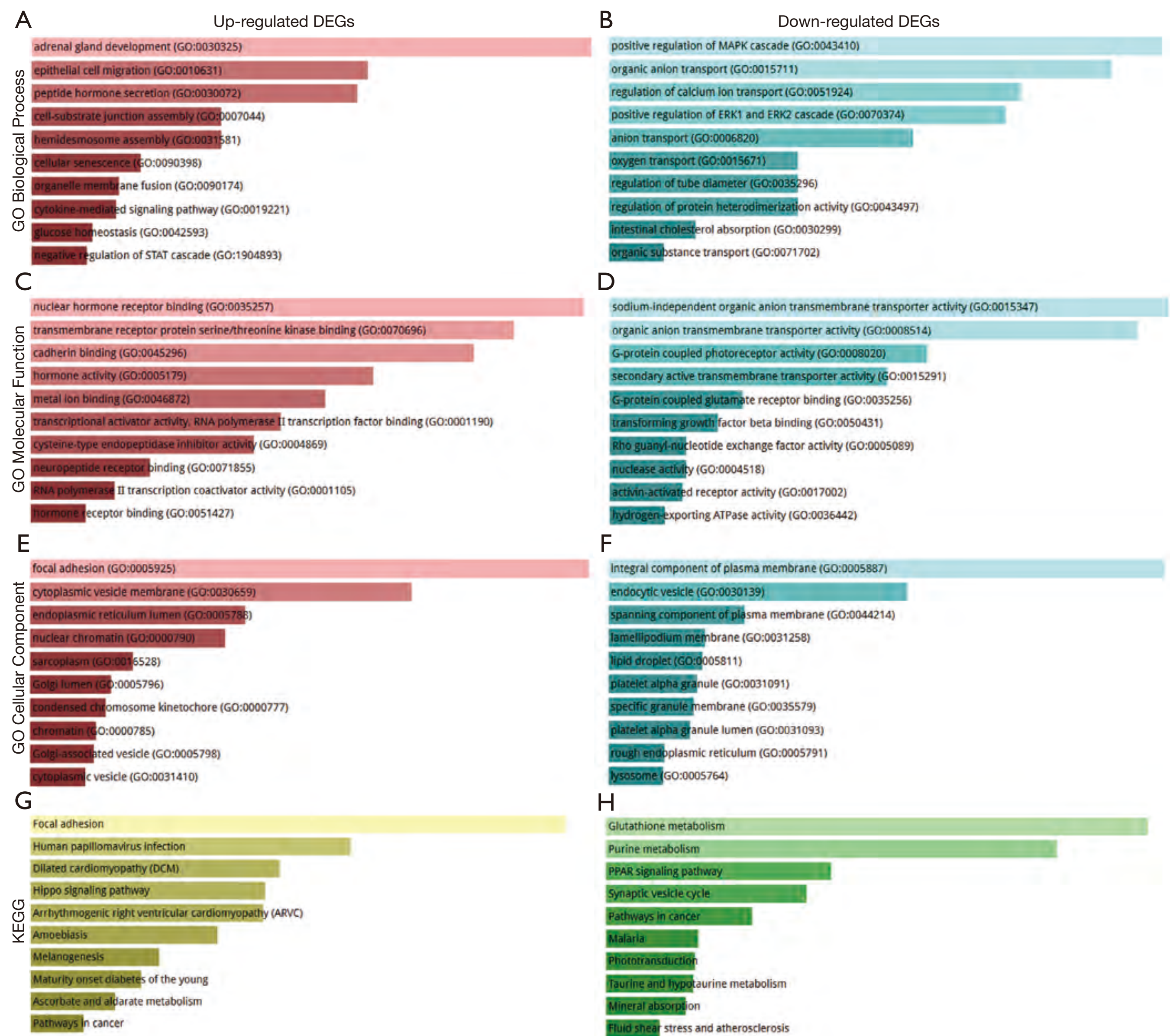

Figure 2 GO annotation and KEGG pathway enrichment analyses of DEGs between PC tissues and matched adjacent non-tumor tissues. The top 10 enriched GO biological process (A,B), molecular function (C,D), and cellular component (E,F) terms, as well as KEGG pathways (G,H). GO, gene ontology; KEGG, Kyoto encyclopedia of genes and genomes; DEG, differentially expressed gene.

COL17A1, *HOXA13*, *RTN1*, and *CHRAC1* to be the hub genes in PC, as their expression levels were associated with both overall survival and tumor stage.

Identification of hub lncRNAs

In order to identify the potential core lncRNAs in PC, we examined the co-expression correlations between DElncRNAs and DEGs in both the PC and AN tissues. We identified the lncRNAs with the most significant correlations to the five hub genes identified above (Pearson correlation coefficient >0.999, https://cdn.amegroups.cn/static/public/10.21037gs-21-151-6.xlsx). Of these potential core lncRNAs, we then selected 16 lncRNAs which were also within the top 100 largest absolute value of degree differences in the lncRNA-mRNA co-

© Gland Surgery. All rights reserved.

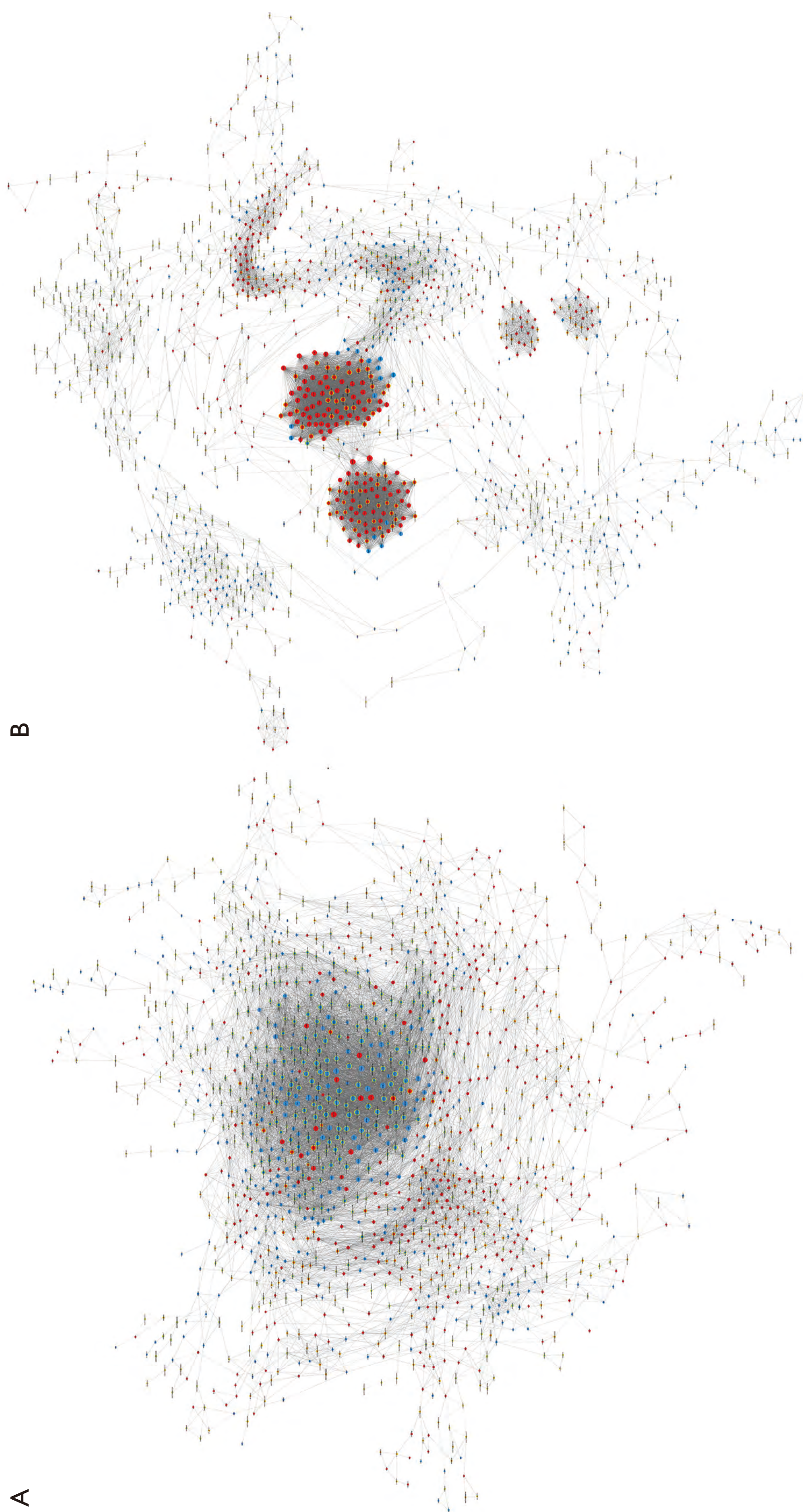

Figure 3 LncRNA-mRNA co-expressed network in pancreatic cancer tissues and matched adjacent non-tumor tissues. The view of the lncRNA-mRNA networks in PC tissues (A) and in matched AN tissues (B). The nodes without yellow rims represent mRNAs, and the nodes with yellow rims represent lncRNAs. The red nodes represent up-regulated mRNAs/lncRNAs. The blue nodes represent down-regulated mRNAs/lncRNAs. The solid line between two nodes represents a positive relationship, and the dashed line between two nodes represents a negative relationship. PC, pancreatic cancer; AN, adjacent non-tumor.

© Gland Surgery. All rights reserved.

Table 1 Significantly changed genes with the top 100 biggest absolute value of degree differences in the lncRNA-mRNA co-expressed network

Gene	Style	T_Degree	T_K	N_Degree	N_K	DiffK(AT-BN)	|DiffK|
SFTA2	Up	117	0.944	0	0.000	0.944	0.944
SPG21	Down	110	0.887	0	0.000	0.887	0.887
COL17A1	Up	115	0.927	5	0.041	0.886	0.886
TM4SF4	Up	13	0.105	113	0.934	–0.829	0.829
PPP2R5E	Up	107	0.863	6	0.050	0.813	0.813
NR2F2	Up	103	0.831	3	0.025	0.806	0.806
HOXA13	Up	0	0.000	96	0.793	–0.793	0.793
GSTA2	Down	0	0.000	96	0.793	–0.793	0.793
GGT5	Up	0	0.000	95	0.785	–0.785	0.785
C15orf57	Up	0	0.000	95	0.785	–0.785	0.785
PBX4	Up	1	0.008	95	0.785	–0.777	0.777
STOM	Up	0	0.000	94	0.777	–0.777	0.777
GJA1	Up	0	0.000	94	0.777	–0.777	0.777
RTN1	Up	3	0.024	96	0.793	–0.769	0.769
CHRAC1	Up	2	0.016	95	0.785	–0.769	0.769
TBRG4	Up	0	0.000	93	0.769	–0.769	0.769
NVL	Up	0	0.000	93	0.769	–0.769	0.769
GMEB1	Up	0	0.000	93	0.769	–0.769	0.769
CLPTM1	Up	0	0.000	93	0.769	–0.769	0.769
CXorf56	Up	0	0.000	93	0.769	–0.769	0.769
GIMAP2	Up	0	0.000	93	0.769	–0.769	0.769
SERPINB9	Up	3	0.024	95	0.785	–0.761	0.761

expressed network (https://cdn.amegroups.cn/static/public/10.21037gs-21-151-7.xlsx). As a result, we identified AL109748.6, AC125792.1, CTD-2292M14.1, AK027298, AK094441, BC084573, RP11-288L9.1, AC012005.5, NR_002827, NR_024058, NR_024427, AK056486, RP11-312H15.3, AC006427.3, BC035067, and CTD-2066L21.3 as the hub lncRNAs in PC (*Table 3*).

Discussion

PC is one of the most life-threatening diseases in the world. In recent decades significant advances have been made in understanding the molecular mechanisms of PC and in improving treatment, but to date no sensitive biomarker has been identified that could assist with early diagnosis or provide an effective target for drug therapy, and the mortality rate remains close to 100% (21). Therefore, the identification of new factors underlying PC development and novel individualized therapeutic targets are urgently needed for PC patients.

In the present study, we used commercial microarrays to screen out DEGs and DElncRNAs between human PC and paired AN tissues. Subsequent bioinformatics analyses were performed to identify the hub genes and hub lncRNAs correlated with PC. As a result, 943 DEGs and 1,138 DElncRNAs were identified. KEGG enrichment analysis revealed that focal adhesion was the most significantly enriched pathway for 501 up-regulated DEGs, and glutathione metabolism was the most significantly enriched pathway for 442 down-regulated DEGs. In cell biology, focal adhesions create combinatorial signaling complexes and mediate the actin-integrin function to regulate cellular

© Gland Surgery. All rights reserved.

Table 2 The expression of significantly changed potential hub genes in pancreatic cancer (ONCOMINE)

Gene ID	Fold Change	P value	t-test	References
SFTA2	8.818	1.36E-9	7.520	Pei Pancreas (15)
	3.531	5.46E-9	6.934	Badea Pancreas (16)
SPG21	2.512	0.002	4.087	Iacobuzio-Donahue Pancreas 2 (17)
COL17A1	12.996	1.17E-6	8.230	Iacobuzio-Donahue Pancreas 2 (17)
	5.813	1.17E-9	7.269	Pei Pancreas (15)
TM4SF4	2.944	0.007	2.798	Iacobuzio-Donahue Pancreas 2 (17)
NR2F2	2.174	1.46E-8	6.417	Badea Pancreas (16)
HOXA13	2.881	0.042	1.824	Grutzmann Pancreas (18)
GGT5	2.268	1.99E-13	9.039	Badea Pancreas (16)
STOM	2.936	6.65E-5	5.826	Segara Pancreas (19)
GJA1	4.055	0.001	3.676	Segara Pancreas (19)
	3.495	7.53E-9	6.796	Badea Pancreas (16)
RTN1	2.298	0.001	4.784	Logsdon Pancreas (20)
TBRG4	2.837	0.004	3.923	Iacobuzio-Donahue Pancreas 2 (17)
NVL	2.148	0.035	1.950	Grutzmann Pancreas (18)
GIMAP2	3.024	4.80E-4	4.873	Iacobuzio-Donahue Pancreas 2 (17)
SERPINB9	3.012	5.10E-6	6.499	Segara Pancreas (19)

behaviors such as cell motility, proliferation, differentiation, survival, and regulation of gene expression (22,23). Thus, increased activation of focal adhesions is closely associated with tumor metastasis and clinical prognosis, especially in PC. Shi *et al.* found that loss of linc01060 could induce PC invasion and metastasis through a vinculin-mediated focal adhesion turnover (24). Ennajdaoui *et al.* identified that *IGF2BP3* could alter the focal adhesion junction and promote PC cell invasiveness (25). Focal adhesion kinase (FAK) is known to be a scaffolding protein and an integral component of focal adhesions. Furuyama *et al.* suggested that FAK was upregulated in PC and was correlated with the size of PC (26). Moreover, glutathione is a vital intracellular antioxidant. Enzymes of glutathione metabolism have been found to be mediated by oncogenic Kras in PC resulting in a dysregulation of PC growth both *in vitro* and *in vivo* (27). Chio *et al.* found that AKT signaling inhibition and glutathione synthesis could prolong the survival of PC cells, suggesting a prospective target for therapy (28). As a result, it would appear that both focal adhesion and glutathione metabolism are essential pathways in PC progression, as we have identified with DEGs in PC via microarray and KEGG pathway enrichment analyses.

Based on the lncRNA-mRNA co-expressed network, we confirmed *SFTA2*, *COL17A1*, *HOXA13*, *RTN1*, and *CHRAC1* as the hub genes in PC, and their expression levels were associated with both overall survival and tumor stage, verified by the ONCOMINE and GEPIA databases. Previous studies have reported that *SFTA2* (surfactant-associated 2), which is a member of the surfactant protein family, assembles innate immune responses in the central nervous system and lung tissues (29,30). It was also identified as a core regulating gene related to carcinogenesis and prognosis in colorectal cancer (31,32). However, its detailed mechanism is still unclear. *COL17A1* (collagen type XVII alpha 1 chain), a collagen family member, maintains an integral part of the hemidesmosome structure. It was established as a novel gene upregulated in a dog bladder cancer organoid culture (33) and has been shown to be increased in multiple neoplasms, such as ovarian cancer, breast cancer, and lung cancer (34,35). *COL17A1* was also identified as a p53 downstreaming transcriptional target in breast tissues that inhibits cell migration and invasion and is associated with better prognosis (36). Moreover,

© Gland Surgery. All rights reserved.

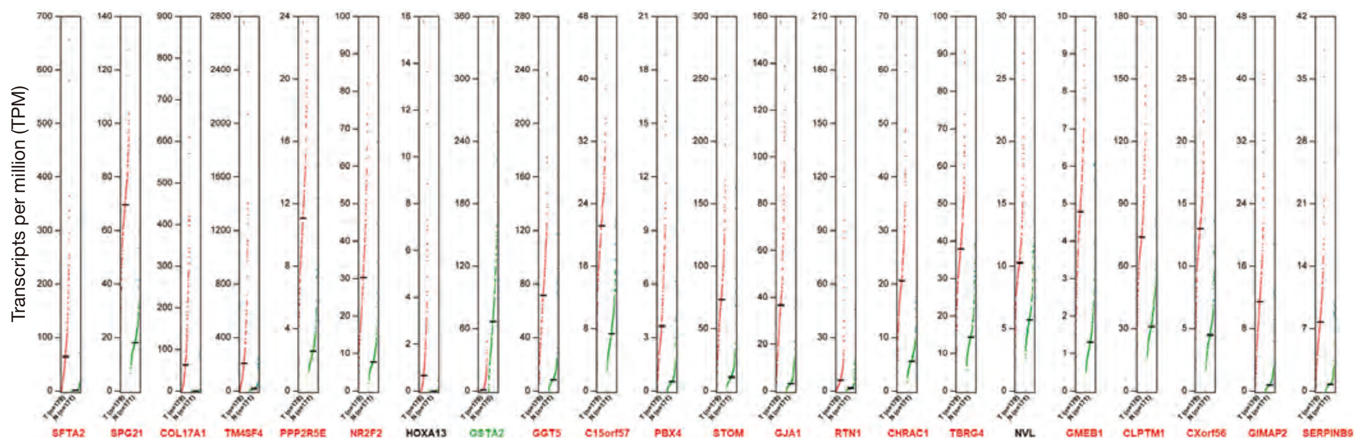

Figure 4 The expression of significantly changed potential hub genes in pancreatic cancer analyzed by the GEPIA databases.

Table 3 The interaction between hub genes and their relative lncRNAs in the lncRNA-mRNA co-expressed network

Hub genes	Relative lncRNAs	Interaction
HOXA13	AL109748.6	0.999981753
	AC125792.1	0.999957999
	CTD-2292M14.1	0.999789742
	AK027298	0.999586913
	AK094441	0.999319273
RTN1	AK027298	0.999999997
	AK094441	0.999967344
	BC084573	0.999820382
	RP11-288L9.1	0.999755758
	AL109748.6	0.999740607
	AC012005.5	0.999677021
	NR_002827	0.999568209
	NR_024058	0.999496232
	NR_024427	0.999347951
	AK056486	0.99929493
CHRAC1	AK056486	0.9999996
	RP11-312H15.3	0.999999069
	NR_024427	0.999997277
	AC006427.3	0.999992174
	BC035067	0.999981174
	NR_024058	0.999977522
	CTD-2066L21.3	0.999970842
	NR_002827	0.999958957
	AC012005.5	0.99991509

this gene was linked to overall survival and classic AJCC staging of PC patients in TCGA and the STRING online tool and external validation datasets (37,38). *HOXA13* is a member of the homeobox A cluster (*HOXA*) gene family with prognostic significance and is involved in a wide range of biological functions in human cancers such as esophageal squamous cell carcinoma, laryngeal squamous cell cancer, ovarian cancer, prostate cancer, and gastric cancer (39-43). It has also been shown to promote PC progression and gemcitabine resistance via a *HOTTIP-HOXA13* axis route (44). *RTN1* (reticulon 1) belongs to the reticulon-encoding gene family and is considered as a specific marker and potential therapeutic target for non-small-cell lung cancers, colorectal cancer, and prostate cancer (45-48). *CHRAC1* (Chromatin Accessibility Complex Subunit 1) is a histone-fold protein-encoding gene, functioning in DNA transcription, replication, and packaging as a sequence-specific DNA binding component, and associated with survival progression among breast cancer patients (49,50). However, little is known of the mechanisms of these five novel hub genes in PC, and, to the best of our knowledge, *SFTA2*, *RTN1*, and *CHRAC1* have not previously been reported in PC.

In order to comprehensively analyze the functions of DElncRNAs in the development and progression of PC and their status in the PC-regulating network, we mapped DElncRNAs in the lncRNA-mRNA co-expressed network of PC and AN tissues. A variety of interactions among DEGs and DElncRNAs were obtained. We identified hub lncRNAs with the largest absolute value of degree differences in a co-expressed network, and with the most significant correlation with the five novel hub genes identified above. As a result, we

© Gland Surgery. All rights reserved.

confirmed 16 hub lncRNAs in PC, which were AL109748.6, AC125792.1, CTD-2292M14.1, AK027298, AK094441, BC084573, RP11-288L9.1, AC012005.5, NR_002827, NR_024058, NR_024427, AK056486, RP11-312H15.3, AC006427.3, BC035067, and CTD-2066L21.3. Few functions and mechanisms of these hub lncRNAs have been known to us in human diseases. Among them, AK056486 has been reported down-regulated in lesional skin biopsies compared to non-lesional ones (51). RP11-288L9.4 can regulate HCV antibiral innate immunity via JAK-STAT signaling pathway (52). However, RP11-288L9.1 is still unclear currently.

As lncRNAs have been reported to work as ceRNAs in regulating the expression and function of genes, 16 hub lncRNAs were also indicated as ceRNAs with their correlated hub genes. As most of these hub lncRNAs have not been reported in PC before, additional experimental and clinical validation is required to support our findings.

Currently, the targeted treatment for pancreatic cancer and the study of its microenvironment is still at the early stage. As in in the newest NCCN clinical practice guideline of pancreatic adenocarcinoma, olaparib has been recommended for patients with germline *BRCA1/2* mutations, pembrolizumab has been recommended for MSI-H or dMMR tumors, larotrectinib and Entrectinib could been used if *NTRK* gene fusion positive in pancreatic cancer patients during metastatic stage as the first-line therapy (53). However, the positive rate of sequencing is quite low as far as our clinical experience in Chinses patients. On the other aspect, the effects of immunotherapies, such as CTLA-4 monoclonal antibody, CAR-T therapy and tumor vaccine, are still unsatisfactory in pancreatic cancer (54-56). As a result, further studies for the core factors in pancreatic cancer development and progression are urgently needed to develop newly therapies and diagnostic targets.

In summary, using an mRNA-lncRNA microarray assay of PC and paired AN tissues and integrated bioinformatics analysis, we identified DEGs and DElncRNAs in PC and constructed mRNA-lncRNA co-expressed regulatory networks. We further indicated five novel hub genes and sixteen novel related hub lncRNAs with potentially significant predictive and therapeutic values in PC. Further clinical studies with larger sample sizes, together with experimental studies focused on the detailed molecular mechanisms involved, are needed to validate the present findings.

Acknowledgments

Funding: This work was supported by the grant from Shanghai Changzheng Hospital Foundation (YQ676).

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at http://dx.doi.org/10.21037/gs-21-151

Data Sharing Statement: Available at http://dx.doi.org/10.21037/gs-21-151

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/gs-21-151). All authors report that this work was supported by the grant from Shanghai Changzheng Hospital Foundation (YQ676); the fee of microarray was supported by grants from the Science and Technology Commission of Shanghai Municipality (124119a0501); and payment for manuscript writing is provided by Shanghai Changzheng Hospital. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study was approved by the Ethics Committee for Biomedical Research of the Second Military Medical University. Informed consent was obtained from all patients or their relatives. All procedures performed in this study involving human participants were in accordance with the Declaration of Helsinki (as revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Grossberg AJ, Chu LC, Deig CR, et al. Multidisciplinary standards of care and recent progress in pancreatic ductal

© Gland Surgery. All rights reserved.

adenocarcinoma. CA Cancer J Clin 2020;70:375-403.
2. Poruk KE, Firpo MA, Adler DG, et al. Screening for pancreatic cancer: why, how, and who? Ann Surg 2013;257:17-26.
3. Liang X, Qi M, Wu R, et al. Long non-coding RNA CUDR promotes malignant phenotypes in pancreatic ductal adenocarcinoma via activating AKT and ERK signaling pathways. Int J Oncol 2018;53:2671-82.
4. Lei S, He Z, Chen T, et al. Long noncoding RNA 00976 promotes pancreatic cancer progression through OTUD7B by sponging miR-137 involving EGFR/MAPK pathway: Journal of experimental & clinical cancer research. CR 2019;38:470.
5. Peng WX, He RZ, Zhang Z, et al. LINC00346 promotes pancreatic cancer progression through the CTCF-mediated Myc transcription. Oncogene 2019;38:6770-80.
6. Xiong XD, Ren X, Cai MY, et al. Long non-coding RNAs: An emerging powerhouse in the battle between life and death of tumor cells. Drug Resist. Updat. 2016;26:28-42.
7. Qi P, Du X. The long non-coding RNAs, a new cancer diagnostic and therapeutic gold mine. Mod Pathol 2013;26:155-65.
8. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. Nature reviews. Genetics 2009;10:155-9.
9. Chen EY, Tan CM, Kou Y, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics 2013;14:128.
10. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 2016;44:W90-7.
11. Pujana MA, Han JD, Starita LM, et al. Network modeling links breast cancer susceptibility and centrosome dysfunction. Nat Genet 2007;39:1338-49.
12. Prieto C, Risueno A, Fontanillo C, et al. Human gene coexpression landscape: confident network derived from tissue transcriptomic profiles. PLoS One 2008;3:e3911.
13. Barabási AL, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet 2004;5:101-13.
14. Tang Z, Li C, Kang B, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 2017;45:W98-W102.
15. Li L, Zhang JW, Jenkins G, et al. Genetic variations associated with gemcitabine treatment outcome in pancreatic cancer. Pharmacogenet Genomics 2016;26:527-37.
16. Badea L, Herlea V, Dima SO, et al. Combined gene expression analysis of whole-tissue and microdissected pancreatic ductal adenocarcinoma identifies genes specifically overexpressed in tumor epithelia. Hepatogastroenterology 2008;55:2016-27.
17. Iacobuzio-Donahue CA, Maitra A, Olsen M, et al. Exploration of global gene expression patterns in pancreatic adenocarcinoma using cDNA microarrays. Am J Pathol 2003;162:1151-62.
18. Grützmann R, Pilarsky C, Ammerpohl O, et al. Gene expression profiling of microdissected pancreatic ductal carcinomas using high-density DNA microarrays. Neoplasia 2004;6:611-22.
19. Segara D, Biankin AV, Kench JG, et al. Expression of HOXB2, a retinoic acid signaling target in pancreatic cancer and pancreatic intraepithelial neoplasia. Clin Cancer Res 2005;11:3587-96.
20. Logsdon CD, Simeone DM, Binkley C, et al. Molecular profiling of pancreatic adenocarcinoma and chronic pancreatitis identifies multiple genes differentially regulated in pancreatic cancer. Cancer Res 2003;63:2649-57.
21. Ryan DP, Hong TS, Bardeesy N. Pancreatic adenocarcinoma. N Engl J Med 2014;371:1039-49.
22. Wozniak MA, Modzelewska K, Kwong L, et al. Focal adhesion regulation of cell behavior. Biochim Biophys Acta 2004;1692:103-19.
23. Geiger B, Spatz JP, Bershadsky AD. Environmental sensing through focal adhesions. Nature reviews. Nat Rev Mol Cell Biol 2009;10(1):21-33.
24. Shi X, Guo X, Li X, et al. Loss of Linc01060 induces pancreatic cancer progression through vinculin-mediated focal adhesion turnover. Cancer Lett2018;433:76-85.
25. Ennajdaoui H, Howard JM, Sterne-Weiler T, et al. IGF2BP3 Modulates the Interaction of Invasion-Associated Transcripts with RISC. Cell Rep 2016;15:1876-83.
26. Furuyama K, Doi R, Mori T, et al. Clinical significance of focal adhesion kinase in resectable pancreatic cancer. World J Surg 2006;30:219-26.
27. Son J, Lyssiotis CA, Ying H, et al. Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway. Nature 2013;496:101-5.
28. Chio IIC, Jafarnejad SM, Ponz-Sarvise M, et al. NRF2 Promotes Tumor Maintenance by Modulating mRNA Translation in Pancreatic Cancer. Cell 2016;166:963-76.
29. Krause M, Peukert N, Härtig W, et al. Localization, Occurrence, and CSF Changes of SP-G, a New Surface Active Protein with Assumable Immunoregulatory Functions in the CNS. Mol Neurobiol 2019;56:2433-9.
30. Xiao J, Lu X, Chen X, et al. Eight potential biomarkers

© Gland Surgery. All rights reserved.

for distinguishing between lung adenocarcinoma and squamous cell carcinoma. Oncotarget 2017;8:71759-71.
31. Li C, Shen Z, Zhou Y, et al. Independent prognostic genes and mechanism investigation for colon cancer. Biol Res 2018;51:10.
32. Gong B, Kao Y, Zhang C, et al. Identification of Hub Genes Related to Carcinogenesis and Prognosis in Colorectal Cancer Based on Integrated Bioinformatics. Mediators Inflamm 2020;2020:5934821.
33. Elbadawy M, Usui T, Mori T, et al. Establishment of a novel experimental model for muscle-invasive bladder cancer using a dog bladder cancer organoid culture. Cancer Sci 2019;110:2806-21.
34. López-Ayllón BD, de Castro-Carpeño J, Rodriguez C, et al. Biomarkers of erlotinib response in non-small cell lung cancer tumors that do not harbor the more common epidermal growth factor receptor mutations. Int J Clin Exp Pathol 2015;8:2888-98.
35. Januchowski R, wierczewska M, Sterzy ska K, et al. Increased Expression of Several Collagen Genes is Associated with Drug Resistance in Ovarian Cancer Cell Lines. J Cancer 2016;7:1295-310.
36. Yodsurang V, Tanikawa C, Miyamoto T, et al. Identification of a novel p53 target, COL17A1, that inhibits breast cancer cell migration and invasion. Oncotarget 2017;8:55790-803.
37. Wu M, Li X, Zhang T, et al. Identification of a Nine-Gene Signature and Establishment of a Prognostic Nomogram Predicting Overall Survival of Pancreatic Cancer. Front Oncol 2019;9:996.
38. Wu J, Li Z, Zeng K, et al. Key genes associated with pancreatic cancer and their association with outcomes: A bioinformatics analysis. Mol Med Rep 2019;20:1343-52.
39. Li J, Ye M, Zhou C. Expression Profile and Prognostic Values of Family Members in Laryngeal Squamous Cell Cancer. Front Oncol 2020;10:368.
40. Kelly Z, Moller-Levet C, McGrath S, et al. The prognostic significance of specific HOX gene expression patterns in ovarian cancer. Int J Cancer 2016;139:1608-17.
41. Luo Z, Rhie SK, Lay FD, et al. A Prostate Cancer Risk Element Functions as a Repressive Loop that Regulates HOXA13. Cell Rep 2017;21:1411-7.
42. Gu ZD, Shen LY, Wang H, et al. HOXA13 promotes cancer cell growth and predicts poor survival of patients with esophageal squamous cell carcinoma. Cancer Res 2009;69:4969-73.
43. Wu DC, Wang SSW, Liu C-J, et al. Reprogramming Antagonizes the Oncogenicity of HOXA13-Long Noncoding RNA HOTTIP Axis in Gastric Cancer Cells. Stem Cells 2017;35:2115-28.
44. Li Z, Zhao X, Zhou Y, et al. The long non-coding RNA HOTTIP promotes progression and gemcitabine resistance by regulating HOXA13 in pancreatic cancer. J Transl Med 2015;13:84.
45. Woenckhaus M, Klein-Hitpass L, Grepmeier U, et al. Smoking and cancer-related gene expression in bronchial epithelium and non-small-cell lung cancers. J Pathol 2006;210:192-204.
46. van de Velde HJ, Senden NH, Roskams TA, et al. NSP-encoded reticulons are neuroendocrine markers of a novel category in human lung cancer diagnosis. Cancer Res 1994;54:4769-76.
47. Lemire M, Qu C, Loo LWM, et al. A genome-wide association study for colorectal cancer identifies a risk locus in 14q23.1. Hum Genet 2015;134:1249-62.
48. Hamzeh O, Alkhateeb A, Zheng J, et al. Prediction of tumor location in prostate cancer tissue using a machine learning system on gene expression data. BMC Bioinformatics 2020;21:78.
49. Mahmood SF, Gruel N, Chapeaublanc E, et al. A siRNA screen identifies RAD21, EIF3H, CHRAC1 and TANC2 as driver genes within the 8q23, 8q24.3 and 17q23 amplicons in breast cancer with effects on cell growth, survival and transformation. Carcinogenesis 2014;35:670-82.
50. Lin W, Lin HD, Guo XY, et al. Allelic expression imbalance polymorphisms in susceptibility chromosome regions and the risk and survival of breast cancer. Mol Carcinog 2017;56:300-11.
51. Zouboulis CC, Nogueira da Costa A, Makrantonaki E, et al. Alterations in innate immunity and epithelial cell differentiation are the molecular pillars of hidradenitis suppurativa. J Eur Acad Dermatol Venereol 2020;34:846-61.
52. Liu X, Duan X, Holmes JA, et al. A Long Noncoding RNA Regulates Hepatitis C Virus Infection Through Interferon Alpha-Inducible Protein 6. Hepatology 2019;69:1004-19.
53. Tempero MA, Malafa MP, Al-Hawary M, et al. Pancreatic Adenocarcinoma, Version 2.2021, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw 2021.
54. Winograd R, Byrne KT, Evans RA, et al. Induction of T-cell Immunity Overcomes Complete Resistance to PD-1 and CTLA-4 Blockade and Improves Survival in Pancreatic Carcinoma. Cancer Immunol Res 2015;3:399-411.
55. Le DT, Picozzi VJ, Ko AH, et al. Results from a Phase IIb, Randomized, Multicenter Study of GVAX Pancreas and CRS-207 Compared with Chemotherapy in Adults with

© Gland Surgery. All rights reserved.

Previously Treated Metastatic Pancreatic Adenocarcinoma (ECLIPSE Study). Clin Cancer Res 2019;25:5493-502.
56. Stromnes IM, Schmitt TM, Hulbert A, et al. T Cells Engineered against a Native Antigen Can Surmount Immunologic and Physical Barriers to Treat Pancreatic Ductal Adenocarcinoma. Cancer Cell 2015;28:638-52.

(English Language Editor: D. Fitzgerald)

Cite this article as: Liang X, Peng J, Chen D, Tang L, Liu A, Fu Z, Shi L, Wang K, Shao C. Identification of novel hub genes and lncRNAs related to the prognosis and progression of pancreatic cancer by microarray and integrated bioinformatics analysis. Gland Surg 2021;10(3):1104-1117. doi: 10.21037/gs-21-151

© Gland Surgery. All rights reserved.

Figure S1 Overall survival and pathological stage plot analysis of the 22 significantly changed potential hub genes in pancreatic cancer patients. Data are presented as the hazard ratio (HR) with a 95% confidence interval. Log-rank P<0.05 was considered statistically significant in overall survival analysis. Pr(>F) <0.05 was considered statistically significant in pathological stage plot analysis.

© Gland Surgery. All rights reserved.