Original Article

Explore the ncRNA and TF-based target data and underlying molecular mechanisms of pancreatic cancer development and progression

Wei-Li Huang¹, Xiao-Hui Guan¹, Chun-Lei Liu¹, Xiao-Ying Li¹, Ying-Ze Ma¹, Yong-Chen Lv²

¹Department of Gastroenterology, Affiliated Hospital of Beihua University, Jilin 132001, Jilin, China; ²Department of General Surgery, Central Hospital of Jilin City, Jilin 132001, Jilin, China

Received February 25, 2020; Accepted May 24, 2020; Epub September 15, 2020; Published September 30, 2020

Abstract: Pancreatic cancer, like a digestive tract malignancy, has sharply raised morbidity and mortality in current years. The early diagnosis rate of pancreatic cancer is low, and the postoperative mortality is higher. Therefore, based on pancreatic cancer-related gene expression profile data, the purpose of this study was to explore the underlying molecular mechanisms of pancreatic cancer development and progression. Combining differential expression analysis, co-expression analysis, and enrichment analysis, we obtained modules related to pancreatic cancer cells. Subsequently, based on the hypergeometric test, we calculated the significance of the multi-regulatory regulation of potential modules and identified the regulatory mechanisms of crucial factors. We obtained a total of 11 modules of pancreatic cancer cell-associated genes, which involved module clustering of 4482 genes. In addition, pancreatic cancer-related module genes can participate in the regulation of pancreatic cancer progression via multiple biological processes and signal transduction pathways. Based on these modular genes, we predicted whether miR-16-5p and miR-195-5p have significant regulatory effects on pancreatic cancer development. At the same time, we know that transcription factors related to pancreatic cancer-related module genes have regulatory effects. This study will be conducive to deeply understand the potential etiopathogenesis of pancreatic cancer and also provide a new therapy direction of pancreatic cancer.

Keywords: Pancreatic cancer, dysfunction module, ADRB2, FENDRR, enrichment analysis

Introduction

Pancreatic cancer is one of the deadly high cancers. According to GLOBOCAN, there are more than 331,000 deaths per year due to pancreatic cancer [1]. Early symptoms are not visible, which leads to many patients with pancreatic cancer who have reached the end of the disease at the time of diagnosis, so the mortality rate is very high [2]. Studies have shown that metastatic progression and immune tolerance of pancreatic cancer are attributed to intercellular communication [3]. In recent years, considerable progress has been made in the detection of pancreatic cancer. However, after diagnosis, only about 4% of patients can survive for five years [4]. As the demand for treatment of cancer patients continues to grow, the need for low recurrence surgery and minimally invasive surgery is increasing urgently [5]. The over-all survival of pancreatic cancer patients is not high, which is mainly because of the deficiency of biomarkers for early diagnosis and disease prognosis. These markers can provide information for decision making and promote individual therapy and optimal clinical results [6]. Surgery is the therapy of first choice for pancreatic cancer, which can remove the lesions. Then, the second choice is additional chemotherapy. Due to the high incidence of pancreatic complications and primary unresectable pancreatic cancer, neoadjuvant chemotherapy is more and more used for pancreatic cancer [7]. For tumor microenvironment and peripheral pancreatic stem cells, the extra cellular therapy of pancreatic cancer is a new direction [8].

Pancreatic cancer is a highly aggressive malignancy for which currently available treatments are of only limited efficacy. For this reason, much research is directed at elucidating fundamental molecular mechanisms underlying the
The ncRNA and TF-based target data and pivotal regulator

biology of pancreatic cancer. These efforts are generating a rapidly growing body of information about oncogenes and tumor suppressor genes such as K-ras [9], p16 (CDKN2, p16-INK4a, MTS1) [10], p53 [11] and gene deleted in pancreatic carcinoma, locus 4 (DPC4/SM-AD4) [12], because of their prevalence and central roles in pathogenesis of PDAC, constitute the genetic signature of this cancer. Downstream signaling mediated by growth factor ligand-receptor interactions is likely to play important roles in a range of phenotypic features of PDAC, including growth, invasion, and angiogenesis, such as epidermal growth factor receptors (EGFR) [13], transforming growth factor-β(TGF-β) [14], and vascular endothelial growth factor (VEGF) [14]. Information on signaling cascades relevant to the behavior of pancreatic cancer cells is accumulating at a rapid pace, such as Raf/mitogen-activated protein kinase (MAPK) [15] cascade, phosphoinositide 3-kinase (PI3K)/AKT/mammalian target of rapamycin (mTOR) [16] signaling cascade, nuclear factor kappa B (NF-κB) signaling cascade [13], developmental cascades [17], telomerase [18]. Evidence suggests that pancreatic cancer develops in a step-wise progression, in which a parallel series of histologic and genetic alterations occur that ultimately lead to invasive pancreatic cancer.

According to the better comprehension of pancreatic cancer, studies have defined some disease precursors of pancreatic intraepithelial neoplasia, intraductal papillary myxoma, cystic tumors, and other cancer progressions [19]. Pancreatic cancer cells have a wide range of reprogramming metabolism, are driven by cancer gene-mediated cellular autonomic pathways, unique physiological environment of the tumor, and responses with noncancer cells [20]. Pancreatic cancer has recently been found to be characterized by a prominent proliferative/stromal response [21]. Somatically acquired mutations in cancer-associated genes may cause it, and modifications can also lead to cancer progression and metastasis [22]. Family history is a major hazard factor for the progression of pancreatic cancer, and it is the genetic risk that may be secondary to familial cancer susceptibility syndrome, genetic pancreatitis and familial pancreatic cancer [23]. Like many cancers, people are more and more interested in the effect of the immune system for pancreatic cancer progression. Especially that immunosuppression in the tumor microenvironment (TME) is believed to damage the host’s antitumor reaction [24]. Close attention is needed to be paid to the secondary signs of pancreatic cancer, such as dilated pancreatic duct, sudden changes of pancreatic duct diameter and parenchymal atrophy. It is of great importance for the diagnosis of pancreatic cancer [25]. Protein glycosylation and abnormalities in polysaccharides have functions in pancreatic tumorigenesis, affecting cancer progression, metastasis, immune response and chemo-resistance [26]. Radiation therapy has a specific effect on abnormal protein changes, and although this effect has been controversial, it is still developing [27]. The development and utilization of tumor microenvironment and surrounding stem cell therapy for pancreatic cancer will be the future therapeutic research direction [8]. Moriya C and Taniguchi H indicated that overexpression of targeting PRDM14 could inhibit the stem cell-like phenotype of cancer. Thus miRNA regulation and siRNA-based targeted therapy are promising new methods [28]. Also, Yamasaki and Nakayama et al. found that Liprin-α4 have functions in inducing malignant phenotypes such as pancreatic cancer proliferation and invasion. Liprin-α4 may be an underlying target for the treatment of pancreatic cancer [29]. CISD2 may be an independent factor for the prognosis of pancreatic cancer, indicating that the CISD2/Wnt/β-catenin pathway is involved in the hyperplasia of pancreatic cancer cells and the occurrence of EMT [30].

Here, a series of comprehensive analyses was made according to the data from pancreatic cancer cells to investigate the effect of key factors in the function and pathways related to the progression of pancreatic cancer. The integrated treatment of function obstacle modules not only is beneficial to investigate the pathogenesis and treatment progress of pancreatic cancer, but also provides a new way for the biologists in further design of the experiments.

Materials and methods

GEO database resource

In the GEO database, we download data on pancreatic cancer [31]. Also, we performed a differential analysis of pancreatic cancer and normal samples. The chip platform is the Affymetrix Human Genome U133 Plus 2.0 Array.
The differential expression of genes was judged to finally acquire differentially expressed genes related to pancreatic cancer

This study was carried out by the R language limma package [32]. For data, background correction function was first employed to carry out the background calibration and standardization of data. Then, according to the quantile normalization method, the control and the low expression probe were filtered out, while standardized high quality data were obtained. The default parameters were employed to analyze the lmFit and eBayes functions of the limma package, and screen the differential expression of genes in pancreatic cancer specimens. When P<0.05, the differential expression of genes was judged to finally acquire differentially expressed genes related to pancreatic cancer.

The correlation coefficients between genes are employed to structure the hierarchical clustering tree, with various branches of the clustering tree as various gene modules

To investigate the coordinate expression of the maladjusted genes, a gene expression profile matrix was structured for the weighted correlation network analysis (WGCNA) of these genes [33]. WGCNA is a method of system biology, which is employed to record the gene associated mode among different samples. It can be employed not only to judge coordinated changes in gene sets with high intrinsic correlation but also to link the relationship between gene expression behavior and sample phenotype. Thus, the weighted value of correlation coefficient, the gene correlation coefficient, is taken to the N-th power, calculating the correlation coefficient (Human Coefficient) between any two genes. Because the nodes in the network are constrained by scale free networks, their characteristics are consistent with the expression relationship between genes, so the algorithm is more biologically essential than other algorithms. After that, the correlation coefficients between genes are employed to structure the hierarchical clustering tree, with various branches of the clustering tree as various gene modules.

These signal pathways are affinitive with the development and progression of pancreatic cancer

The investigation on gene functions and signal pathway participation can be beneficial to research the disease molecular mechanisms. But, the functions and pathways participated in the module gene can characterize the function obstacle mechanism of the module. Therefore, for genes of these 11 functional modules, we use the R language and the cluster profile package to perform Go functions (p-value cutoff = 0.01, q value Cutoff = 0.01) and KEGG pathways (p-value cutoff = 0.05, q value Cutoff = 0.2) for enrichment analysis [34]. These signal pathways are affinitive with the development and progression of pancreatic cancer.

The ncRNA and TF-based target data were predicted as background sets, and the pivotal regulator of the regulatory dysfunction module was obtained

Data on target regulatory relationships during transcription and post-transcription were contained in the TRRUST v2.0 database and the RAID v2.0 database [35, 36]. Among them, we downloaded and used all human transcription factors’ target data in the TRRUST v2 database, involving 2492 transcription factors and 9396 TF-Module interaction pairs. In the RAID 2.0 database, we screened 431937 ncRNA-mRNA interactions of 5431 ncRNAs. These regulators participate in the mediation of the disease progression. With an aim of investigating the driving forces of co-expression modules in pancreatic cancer-related genes, according to these response data, we performed a pivotal analysis. Pivot analysis means that in the target pair, we seek at least two interacting drivers, and we calculate the response significance between the driver and the module based on hypergeometric test. Based on the context of TF_Pivot data, we performed Pivot analysis and screened p_value <0.01, resulting in 18 transcription factors including 18 TF-Module interaction pairs, and 408 ncRNAs including 447 ncRNA-Module interaction pairs. Finally, a statistical analysis of the pivot was performed, and the pivoting pivot was identified as the core pivot. The ncRNA and TF-based target data were predicted as background sets, and the pivotal regulator of the regulatory dysfunction module was obtained.

Result

Some genes are dysregulated in patients with pancreatic cancer. Based on the expression profiles of the selected pancreatic cancer, we obtained 4482 differentially expressed genes (Schedule 1). Among them, the change ratio of
The ncRNA and TF-based target data and pivotal regulator

REG4 and MUC17 is higher. These differentially expressed genes are all associated with pancreatic cancer so that they may have functions in the development of a disease. In the patient sample, to research the functional mechanism of pancreatic cancer-relevant genes, widespread analysis was performed. First, expression profiles were structured in 4,472 pancreatic cancer-associated genes in patient specimens. Then, according to co-expression network analysis, we obtained the expression of pancreatic cancer modules for 11 groups (Figure 1A-C), and then found that the 11 groups of modules have the synergistic expression in the group.

Disorder factors are involved in cancer-related functions and signaling pathways

These function modules may be related to various functions and pathways, which represent different adjustment mechanisms in the mediation of the development and progression of pancreatic cancer. To study the possible function obstacle of modular gene imbalance, separate enrichment analysis of functions and pathways was performed for each module. The GO function of the 11 module genes and the enrichment analysis of KEGG pathway showed that the genes participated in the function were 3252, the molecular function term was 4613, and the biological process was 24261 (Figure 2A). According to the statistical analysis, it was found that the relevant functional module genes are mostly participated in cell proliferation, regulation of intracellular protein transport and other biological processes. In another way, KEGG analysis indicated that these module genes affected 1265 KEGG pathway enrichment. They are mostly participated in signal pathways such as serotonergic synapse and ether lipid metabolism, which are primarily related to the development and progression of pancreatic cancer (Figure 2B). We understand that the function and pathway results of module gene enrichment are closely related to pancreatic cancer. Therefore, these 11 groups of modules were identified as dysfunction modules.

Drive TF and ncRNA of the pancreas

The regulation before and after the transcription of genes has been always believed as a
The ncRNA and TF-based target data and pivotal regulator

A crucial factor that regulates the occurrence and development of diseases, and ncRNA is recognized to be an important regulation factor. For the regulation of dysfunction, the prediction of the module gene will help us to further investigate the transcriptional regulation mechanism of pancreatic cancer. For this purpose, we conducted a critical analysis according to the targeted relationship between ncRNA and genes to research ncRNA regulatory factors causing module function obstacle. The predicted results (Figure 3A) reveal that 408 ncRNAs play an important role in the regulation of the module, which involve 447 ncRNA-Module interaction pairs. In addition, the statistical analysis of the results showed that miR-16-5p and miR-195-5p have substantial organizational relationships with the three function obstacle modules. FENDRR, let-7c-5p, miR-125a-3p and miR-1279 have been identified and have regulatory functions in two dysfunction modules, which may become potential pathogenic factors for pancreatic cancer. Other ncRNAs also exhibit essential modulation of dysfunction modules and contribute primarily to the pathogenesis of pancreatic cancer. Transcription factors are equally important for the transcriptional regulation of genes.

Figure 2. Modules involved in gene function and pathway identification of pancreatic cancer dysfunction modules. A. GO function of module gene participation. The deeper the color, the stronger the enrichment. The larger the circle, the more significant the proportion of the gene in the module that accounts for the GO function. B. The KEGG signaling pathway involved in the modular gene. The deeper the color, the stronger the enrichment. The larger the circle, the more significant the proportion of the gene in the KEGG pathway entry.
Figure 3. Regulators that regulate the module. A. ncRNA pivotal regulation network map of pancreatic cancer-associated modules. B. TF pivot regulation network diagram for pancreatic cancer-related modules.
Many research results have indicated that the regulation obstacle of transcriptors may cause a variety of diseases. The occurrence of pancreatic cancer is also not separated from the unbalancing of transcriptors, also reflecting in the regulation of transcriptors for function obstacle modules. Hence, we conducted a pivot analysis of the module according to the regulatory relationship between the transcriptors and the gene. The results (Figure 3B) indicate that for the pancreatic cancer dysfunction module, 18 transcriptors in total have transcriptional regulation, involving 18 TF-Module regulatory pairs. With statistical analysis, we found that tumor suppressor genes including CREBBP, CRX, DEAF1, E2F4, and FOXP1 essentially regulated two dysfunction modules. It has potential functions in controlling pancreatic cancer. These transcription factors, which have essential regulatory effects on multiple dysfunction modules, have been identified as transcription factors for the core of pancreatic cancer. We also found that FENDRR, an essential gene affecting pancreatic cancer, may promote pancreatic cancer cell proliferation by targeting ADRB2 (Figure 4).

**Discussion**

After five years of diagnosis of pancreatic cancer, the patient survival rate was only 8%. Even after surgical resection, most patients will have recurrence [37]. In recent years, scientists have focused on the discovery of particular genes or proteins, as well as related signaling pathways, and have achieved specific results. With an aim of deep understanding the molecular mechanisms in the progression of pancreatic cancer, we have integrated various analytical methods to investigate the mystery of disease development. First, we structured all expression profiles of pancreatic cancer specimens for various analyses and screened 4482 underlying pathogenic genes containing REG4 and MUC17.

Recent studies have shown that overexpression of genes is associated with the development of pancreatic cancer [38]. In the digestive tract, MUC17 is a type 1 membrane-bound glycoprotein whose abnormal overexpression is associated with the malignant potential of pancreatic ductal adenocarcinoma [39]. Co-expression modules can affect the development of pancreatic cancer after participating in multiple functions and signaling pathways. Based on functional analysis results, we found that the module genes are mainly involved in the regulation of intracellular protein transport and cell proliferation. According to related research, insulin can induce PLK1 to promote the spread of pancreatic ductal epithelial cells through PI3K/AKT and NF-κB pathways, indicating that insulin may be one of the mechanisms involved in developing pancreatic cancer [40]. Besides, studies have shown that IL-9 essentially promotes proliferation, invasion, and migration of pancreatic cancer cells. It confirms that the biological process of module involvement in cell proliferation can encourage the development of pancreatic cancer [41].

On the other hand, due to the abundant pathway results, it was found that the module gene also participated in the pathogenesis of pancreatic cancer through 5-hydroxytryptamine synapse, ether lipid metabolism, and other signal pathways. The microenvironment of pancreatic tumor has a large content of growth factors and inflammatory cytokines to support tumor growth. In addition, it has a highly immunosuppressive effect, and up regulated cytokine pathway may adjust the progression of PDAC [42]. Based on Carreras Torres R and other scholars’ researches, the increase of fasting
insulin levels of genes is positively correlated with the increase of the pancreatic cancer risk. Therefore, insulin resistance may be believed as a hazard factor of pancreatic cancer [43]. All in all, these functions and pathways that participated in the modular gene produce a comprehensive network effect, and the extensive regulation of dysfunction modules mediates the underlying pathogenesis of pancreatic cancer.

Subsequently, we explored a series of driving factors for these dysfunction modules involving ncRNA (miR-16-5p and miR-195-5p) and transcription factors (CREBBP, CRX, DEAF1, E2F4, and FOXP1). MiR-195, one of the miR-16/15/195/424/497 family members, has been shown to play an important role in tumorigenesis, as a tumor suppressor. Here, we assessed miR-195 expression in colorectal cancer, and found that downregulation of miR-195 correlates with lymph node metastasis and poor prognosis in colorectal cancer [44]. In terms of the driving force of ncRNA, both miR-16-5p and miR-195-5p have regulatory effects on three functional barrier modules of pancreatic cancer. Kuśnierz-Cabala B passed the control group of patients with acute pancreatitis, and it was concluded that miR-16-5p could be used to diagnose acute pancreatitis in different degrees [45]. Also, other ncRNAs such as FENDRR, miR-125a-3p and miR-1279 have also been identified to have regulatory functions in two dysfunction modules. It may be a potential factor affecting pancreatic cancer. FENDRR has been suggested to serve important functions in vertebral development and is downregulated in numerous types of cancer [46-48]. A low level of FENDRR expression promotes local invasion and lymphatic metastasis in gastric cancer, which are key steps for cancer progression [49]. Liu et al. found that miR-125a-3p is responsible for the chemosensitivity of PDAC. By direct targeting, it can inhibit epithelial-mesenchymal transition to Fyn [50]. As previously reported, patients with type 2 diabetes (T2DM) have a higher risk of developing pancreatic cancer, so it is speculated that diabetes may be closely related to the pathogenesis of pancreatic cancer [51]. Moreover, in patients with type 2 diabetes, miR-16-5p was found to have functions in regulating the transcription of different tissues. It may be a potential key regulator of pancreatic cancer [52].

In terms of transcription factor regulation, transcription factors such as CREBBP, CRX, DEAF1, E2F4, and FOXP1 essentially regulate two dysfunction modules and exert potential regulatory effects on pancreatic cancer. Recent studies have shown that by directly targeting RBL2 and miR-17-5p, one can disrupt the RBL2/E2F4-related gene suppressor complex. It may promote the proliferation of adenocarcinoma cells and alter the cell cycle spectrum [53]. FOXP1 is considered to be a tumor suppressor gene, which supports the inhibition of pancreatic cancer development [54]. There are some factors in this study that regulate dysfunction modules. Although there is no specific literature for reference, they all have regulatory functions and participate in the progression of pancreatic cancer. Based on the above series of analyses, we believe that FENDRR is a crucial regulatory gene that promotes the proliferation of pancreatic cancer cells.

Acknowledgements

This work was supported by Jilin Provincial Department of Education “The 13th Five-year” Science and Technology Projects (JJKH2019-0656KJ).

Disclosure of conflict of interest

None.

Address correspondence to: Yong-Chen Lv, Department of General Surgery, Central Hospital of Jilin City, No. 4, Nanjing Street, Chuanying District, Jilin 132001, Jilin, China. Tel: +86-0432-62167219; E-mail: bqu87v@163.com

References

The ncRNA and TF-based target data and pivotal regulator

World J Gastroenterol 2014; 20: 12118-12131.


The ncRNA and TF-based target data and pivotal regulator


[38] He XJ, Jiang XT, Ma YY, Xia YJ, Wang HJ, Guan TP, Shao QS and Tao HQ. REG4 contributes to the invasiveness of pancreatic cancer by up-regulating MMP-7 and MMP-9. Cancer Sci 2012; 103: 2082-2091.


