Identification of novel biomarkers and prognostic value of the location (head, body, or tail) of pancreatic cancer

Shuwen Han
Huzhou Central Hospital
Xi Yang
Huzhou Central Hospital
Jiamin Xu
Huzhou University
Wei Wu
Huzhou Central Hospital
Jin Liu (liujin190701@163.com)
Huzhou Cent Hosp, Affiliated Cent Hops HuZhou University

Research article

Keywords: pancreatic cancer, competing endogenous RNAs, prognosis risk model, Nomogram

Posted Date: September 23rd, 2019

DOI: https://doi.org/10.21203/rs.2.14728/v1

Abstract

Objective : This study was designed to identify the differentially expressed mRNA, microRNA (miRNA), and long non-coding RNA (lncRNA) and their functions in pancreatic cancer (PC). Methods: The expression data of PC and normal samples were downloaded from the GEO database. The expression data of pancreatic head (H), body (B), and tail (T) were downloaded from the TCGA database. After data preprocessing, the differential analyses between PC vs. Normal, H vs. B, H vs. T, and T vs. B were performed. Overlapping genes between PC vs. Normal and the different locations (the union of genes among T vs. B, T vs. H, and B vs. H) were selected. The competing endogenous RNAs (ceRNA) network was constructed based on co-expression analysis and prediction of targets, followed by functional enrichment analysis. Construction of an mRNA prognosis risk model and screening of prognostic factors were performed using Cox univariate/multivariate regression analysis, followed by Nomogram model construction. Finally, the gene-drug interactions were predicted for the DE-mRNA. Results: A five-mRNA prognostic model (GRHL2+CACNA1A+GRM1+UPK1B+PKHD1) was constructed, and the risk score was relatively increased with the increased expression of the GRHL2, PKHD1, and UPK1B, and the decreased expression of CACNA1A and GRM1. Compared with pancreatic body/tail cancer, the expression of GRHL2 was increased, while the expression of CACNA1A and GRM1 was decreased in pancreatic head cancer. LncRNA AC006369.2-miR-146a-5p-CACNA1A/GRM1 was a regulatory axis in the ceRNA network. Verapamil was predicted to be an antagonist of CACNA1A. Conclusion: Our results provide a new direction for the accurate diagnosis and treatment of PC and for investigating the mechanism of PC.

Background

Pancreatic cancer (PC) is a lethal disease and is the fourth leading cause of cancer-related deaths globally [1]. The incidence and mortality of PC varies greatly in different countries and has been increasing annually, with the highest rates reported in developed countries [2]. It is an aggressive cancer that is difficult to diagnose at early stages due to the lack of early clinical symptoms [3]. Up to 80% of PC patients present locally advanced or late metastases at diagnosis; their median survival time is 4 months or less [1, 3].

Surgical resection is the main treatment for PC patients. The 5-year overall survival (OS) rate is approximately 5% (range, 2% to 9%) [2]. Survival of PC patients is correlated with various factors, including tumor stage, therapy method, and tumor location [4]. PC originates from exocrine/endocrine pancreatic cells and approximately 95% of PC cases display the histologic characteristics of pancreatic ductal adenocarcinoma, which can be divided into pancreatic head and pancreatic body/tail cancer based on anatomy [5, 6]. Accumulating evidence has highlighted the differences in pathological properties and course, incidence, therapy, as well as prognosis between pancreatic head cancer and pancreatic body/tail cancer [5, 7]. Compared with pancreatic head cancer, pancreatic body/tail cancer displays a more aggressive tumor biology, lower 3-year survival rates, and is less resectable. In addition, it is usually more advanced at diagnosis because of the lack of early symptoms of biliary obstruction [4, 8]. Considering that early diagnosis is less likely and the lack of effective therapies, prevention is considered

a meaningfully strategy for PC [9]. Hampering this, the etiology of PC is not fully understood even though several risk factors have been identified, such as smoking, genetics, obesity, and others. It is necessary to better understand the etiology and definitively identify the risk factors of PC.

Animal research has shown that the development of PC can be caused by the targeted activation of the KRAS2 oncogene coupled with inactivation of tumor protein P53 or cyclin dependent kinase inhibitor 2A [10]. Sequencing was used to demonstrate that partner and localizer of BRCA2 (PALB2) is a susceptibility gene of PC; its truncating mutation is found in patients with familial PC [11]. In addition, non-coding RNAs including microRNAs (miRNAs) and long non-coding RNAs (lncRNAs, which exceed 200 nucleotides in length), have been reported to have important roles in various biological processes in PC [12, 13]. MiRNAs are a type of short ncRNAs that are crucial in PC progression. For example, miR-96 serves as a tumor suppressor by directly targeting KRAS, which can drive PC [14]. Plasma miR-744 is highly expressed in PC, and this expression is related to lymph node metastasis and recurrences, suggesting its potential as a biomarker [15]. LncRNAs have been implicated in various cytological processes, tumor metastasis, and tumor progression [16, 17]. Moreover, they can serve as competing endogenous RNA (ceRNA) by competitively binding miRNAs to modulate the expression of miRNA targeted genes [17, 18]. One study reported the high expression of lncRNA regulator of reprogramming in PC, and its role as a tumor promoter in PC as a ceRNA to modulate the expression of the Nanog transcription factor by competitively binding miR-145 [19].

In this study, the differentially expressed genes (including miRNA, lncRNA, and messenger RNA [mRNA]) associated with PC and cancer locations (pancreatic head, body/tail) were identified, followed by the construction of a ceRNA network and functional enrichment analysis. In addition, screening for prognostic factors and construction of an mRNA prognosis risk model and Nomogram model construction were performed to identify potential prognostic factors and crucial mRNAs, and to predict their corresponding targeted drugs. The results should provide a theoretical basis and novel biomarkers in the study and treatment of PC.

Methods

Data sources

The human PC related microarray datasets GSE86436 (expression data of mRNA and lncRNA) and GSE85589 (expression data of miRNA) were downloaded from the Gene Expression Omnibus (GEO (http://www.ncbi.nlm.nih.gov/geo/) database. There were six primary PC tissue samples and six adjacent non-tumor tissue samples in the GSE86436 dataset, and all the samples were detected on the Arraystar Human LncRNA microarray V2.0 platform (Agilent-033010 Feature Number version). In addition, blood serum samples from 88 PC patients and 19 healthy individuals in the GSE85589 dataset were used, with the [miRNA-4] Affymetrix Multispecies miRNA-4 Array platform.

In addition, related data for pancreatic adenocarcinoma (PAAD) in The Cancer Genome Atlas (TCGA) was downloaded from the University of California Santa Cruz Genome Browser database

(http://xena.ucsc.edu/) [20], including clinical data, gene expression RNAseq log2(count+1) data, and miRNA expression log2(RPM+1) data. Pancreatic head, pancreatic body, and pancreatic tail samples were selected based on sample clinical phenotype information. A total of 137 pancreatic head samples, 14 pancreatic body samples, 14 pancreatic tail samples as well as their corresponding gene expression RNAseq log2(count+1) and miRNA expression log2(RPM+1) data were obtained.

Data preprocessing and lncRNA re-annotation

For the GSE86436 dataset, the standardized expression profile of mRNA and lncRNA, as well as probe sequences, were downloaded. The sequences of the probes were mapped to the GRCh38 human reference genome and “unique map” probes were selected. Then, the mapped gene for each probe was obtained based on the corresponding position and positive/negative strand information on the chromosome as well as the “Release 25” annotation file. Probes with the annotation of “protein_coding” were the corresponding mRNA corresponding, while probes with the annotations of “antisense”, “sense_intronic”, “lincRNA”, “sense_overlapping”, and “processed_transcript” were the corresponding lncRNA probes.

For the GSE85589 dataset, the standardized miRNA expression profile and annotation files were downloaded, followed by annotation of the probes. The probes with no mapped miRNA were removed, and when multiple probes mapped to one miRNA, the mean expression value was considered as the expression value of this miRNA.

For the data downloaded from the TCGA database, the gene expression RNAseq log2(count+1) data were converted to count values. Genes with a count value of 0 in more than half of the samples were filtered, followed by gene annotation based on the “Release 25” annotation file. Similarly, for the miRNA expression log2(RPM+1) data, miRNAs with a log2(RPM+1) of 0 in more than half of the samples were filtered, followed by their annotation based on miRbase database.

Differential analysis

The expression profiles of mRNA, lncRNA, and miRNA downloaded from the GEO database were analyzed to determine differential expression between the PC and normal groups. The corresponding P-value and log fold change (FC) were obtained using the classical Bayes method in the Limma package (Version 3.10.3, http://www.bioconductor.org/packages/2.9/bioc/html/limma.html). For the expression profiles of lncRNA and mRNA downloaded from the TCGA database, the raw count was standardized and converted into logCPM value using the TMM method in the edgeR package [21, 22] (Version: 3.4, http://www.bioconductor.org/packages/release/bioc/html/edgeR.html) to perform the differential expression analysis between pancreatic tail and vs. pancreatic body (T vs. B), pancreatic tail vs. pancreatic head (T vs. H), and pancreatic body vs. pancreatic head (B vs. H). The classical Bayes method in the Limma package was used to analyze the miRNA log2(RPM+1) between the three groups.

Differentially expressed mRNA (DE-mRNA) and DE-lncRNA were selected with the threshold of $P < 0.05$ and $|logFC| > 1$, while $P < 0.05$ and $|logFC| > 0.263$ were considered as the threshold for DE-miRNA. Overlapped mRNAs, lncRNAs, and miRNAs between PC vs. Normal and (union of genes among T vs. B, T vs. H, and B vs. H) were considered as the DE-mRNAs, DE-lncRNAs, and DE-miRNAs, respectively, related to PC and the locations of PC occurrence, and they were used in the subsequent analysis.

Co-expression analysis

The DE-mRNAs and DE-lncRNAs were used to calculate the Pearson correlation coefficient (r) of each mRNA and each lncRNA by one-to-one correspondence of the samples. The lncRNA-mRNA interaction pairs were selected with cut-offs of $r > 0$ and $P < 0.05$. Similarly, the r-value of each mRNA and each miRNA was also calculated by one-to-one correspondence of the samples, and the miRNA-mRNA interaction pairs were selected with cut-offs of $r < 0$ and $P < 0.05$.

ceRNA network construction

The miRNA-lncRNA interaction binding sites were predicted using miRanda software (version 3.3a), and the miRNA-lncRNA pairs were selected with a threshold score > 140 and threshold Energy < -20. In addition, the target genes of miRNAs were also predicted using miRWalk 2.0 (http://zmf.umm.uni-heidelberg.de/apps/zmf/mirwalk2/). The miRNA-mRNA pairs predicted to appear in at least three of nine databases (including miRWalk, Microt4, miRanda, miRMap, miRNAMap, PITA, RNA22, RNAhybrid, and Targetscan) were selected. The overlapped miRNA-mRNA pairs between co-expressed miRNA-mRNA pairs and predicted miRNA-mRNA pairs were considered the final miRNA-mRNA pairs. Finally, the lncRNA-miRNA-mRNA interactions were obtained based on the lncRNA-miRNA pairs, mRNA-miRNA pairs, and the co-expressed mRNA-lncRNA pairs, followed by the ceRNA network construction using Cytoscape software (version 3.4.0, http://chianti.ucsd.edu/cytoscape-3.4.0/).

Functional enrichment analysis

To explore the involved function of the DE-mRNAs, lncRNAs, and miRNAs, the clusterProfiler [23] (version 3.8.1, http://bioconductor.org/packages/release/bioc/html/clusterProfiler.html) in the R package was used to enrich the biological processes in Gene Ontology (GO) annotation and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. The significantly enriched function terms were selected with the threshold of P.adjust < 0.05 and enriched gene count > 2. Notably, the functions of lncRNAs and miRNA were obtained based on the functional enrichment of the mRNAs in co-expressed mRNA-lncRNA pairs and in the final obtained mRNA-miRNA pairs, respectively.

Construction of mRNA prognosis risk model

The mRNAs in the ceRNA network were considered as the candidate mRNAs. Cox univariate regression analysis was used to calculate the regression coefficient and P-value between each candidate mRNA and the survival time and state. The prognosis-related mRNAs were selected with the P-value threshold < 0.05, combined with hazard ratio (HR) risk (theoretically, up-regulated genes between PC vs. Normal should be

risk factors, corresponding to an HR > 1; otherwise HR < 1). The Risk Score was calculated as βgene1*exprgene1 + βgene2*exprgene2 + ... +βgenen*exprgenen, in which β is the prognostic correlation coefficient and exprgene is the expression value of the corresponding gene. The mRNAs were added to the model according to the P-value from small to large one-by-one. The high and low risk samples classified by the average value of the model constructed after adding an mRNA had the greatest significant correlation with survival (log-rank test). The area under the curve (AUC) of the high and low risk samples was maximized according to the expression value of the selected mRNA. In this case, the mRNA model was considered to be associated with prognosis.

Screening of prognostic factors and construction of Nomogram model

The clinical information of PC samples in the surveillance, epidemiology, and end results (SEER) and TCGA databases were downloaded. Clinical factors, including age, gender, location of cancer, clinical stage, and tumor histological grade, and the prognostic risk model score calculated as described above were used as independent variables. OS was used as the dependent variable to perform Cox univariate regression analysis for clinical factors having $P < 0.05$ to screen for prognostic factors. The Nomogram model was constructed based on the results of multivariate regression analysis, including conversion and assignment of regression coefficients, Nomogram plotting, and calibration curve plotting.

Construction of drug-gene interaction network

The targeted drugs for DE-mRNA were predicted using the DGIdb online database (http://www.dgidb.org/search_interactions) using the parameters settings of FDA Approved and Antineoplastic. The drug-gene interaction network was constructed based on the obtained drug-gene interactions using Cytoscape.

Results

DE analysis

Table 1 lists the identified miRNAs, mRNAs, and lncRNAs between PC vs. Normal, T vs. B, T vs. H, and B vs. H. Firstly, 182 miRNAs, 1077 mRNAs, and 127 lncRNAs were obtained after merging the miRNAs, mRNAs and lncRNAs between T vs. B, T vs. H, and B vs. H, respectively. Then, the overlapped miRNAs, mRNAs, and lncRNAs among PC vs. Normal and merged genes (182 miRNAs, 1077 mRNAs, and 127 lncRNAs) were screened (Figure 1A). A total of 10 DE-miRNAs, 204 DE-mRNAs, and 17 DE-lncRNAs were identified that were considered to be related to PC and locations of PC occurrence. Figure 1B presents heatmaps of these DE-miRNAs, DE-mRNAs, and DE-lncRNAs.

Co-expression analysis and ceRNA network construction

Co-expression analysis identified 1097 and 1472 lncRNA-mRNA interaction pairs from the GEO and TCGA databases, and a total of 679 overlapped lncRNA-mRNA interaction pairs were screened as the final co-expressed lncRNA-mRNA pairs. Similarly, there were 1020 and 407 miRNA-mRNA negative correlation

pairs identified from the GEO and TCGA databases, and a total of 218 overlapped miRNA-mRNA negative correlation pairs were screened. In addition, 131 predicted miRNA-lncRNA pairs and 55995 predicted miRNA-mRNA pairs were selected as described above. A total of 40 miRNA-mRNA pairs were screened among the 218 miRNA-mRNA negative correlation pairs and the 55995 predicted miRNA-mRNA pairs.

Finally, based on 679 lncRNA-mRNA pairs, 131 miRNA-lncRNA pairs, and 40 miRNA-mRNA pairs, a total of 107 lncRNA-miRNA-mRNA interactions pairs were obtained. The ceRNA network consisted of seven miRNAs, 14 lncRNAs, and 32 mRNAs (Figure 2), in which miR-183-3p regulated the most mRNA. Calcium Voltage-Gated Channel Subunit Alpha1 A (CACNA1A), and Glutamate Metabotropic Receptor 1 (GRM1) were regulated by miR-146a-5p, which interacted with lncRNA AC006369.2. In addition, Grainyhead Like Transcription Factor 2 (GRHL2) and PKHD1 Ciliary IPT Domain Containing Fibrocystin/Polyductin (PKHD1) were regulated by miR-346. MiR-1180-3p regulated Uroplakin 1B (UPK1B).

Functional enrichment analysis

The functional enrichment analysis for the 204 DE-mRNAs indicated that 44 GO_BP terms and eight KEGG pathways were significantly enriched. They included hsa04972~Pancreatic secretion, GO:0002526~acute inflammatory response, and GO:0055074~calcium ion homeostasis (e.g., GRM1). Table 2 lists the top 10 enriched GO_BP terms and KEGG pathways.

The functions of lncRNAs and miRNAs were obtained based on the functional enrichment of the mRNAs in co-expressed mRNA-lncRNA pairs and in the final obtained mRNA-miRNA pairs, respectively. A total of 85 GO_BP terms and nine KEGG pathways were significantly enriched for four miRNAs. The top five terms for each miRNAs are shown in Figure 3A. Similarly, 1938 GO_BP terms and 198 KEGG pathways were significantly enriched for 15 lncRNAs. The top five terms for each lncRNAs are shown in Figure 3B.

MiR-146a-5p was associated with various neural signal transduction processes, including long-term depression, synaptic transmission, glutamatergic activity, taste transduction, and others. Those functions were obtained based on the functional enrichment analysis of GRM1 and CACNA1A.

Construction of mRNA prognosis risk model

Cox univariate regression analysis was performed for the 32 mRNAs in the ceRNA network, and a total of 10 mRNAs were selected. These 10 mRNAs were added to the prognosis risk model according to their P-values from small to large one-by-one. As shown in Table 3, after adding five mRNAs (GRHL2+CACNA1A+GRM1+UPK1B+PKHD1), the high and low risk samples classified by the average value of the model had the most significant correlation with survival (minimum P = 0.006), and the AUC of the high and low risk samples was maximized according to the expression value of the selected mRNA (maximum AUC = 0.999).

Figure 4A displays the distribution of Risk score, survival time, and gene expression values of the high risk and low risk samples identified by the Risk model constructed using the GRHL2+CACNA1A+GRM1+UPK1B+PKHD1 mRNAs. The survival time of high_risk sample was lower than

the that of low_risk sample, and the Risk score was relatively increased with the increased expression of the GRHL2, PKHD1, and UPK1B, and with the decreased expression of CACNA1A and GRM1. A Kaplan-Meier curve was plotted to verify the correlation of the Risk model with prognosis (Figure 4B). As expected, the Risk model was highly correlated with prognosis, with a higher Risk score indicating lower survival.

Screening of prognostic factors

The clinical information of PC samples in the SEER and TCGA databases were downloaded to perform Cox univariate regression analysis. Age, neoplasm_histologic_grade, and pathologic_N were significantly correlated with prognosis in the two databases (Table 4). Hence, the three clinical factors, together with the Risk score were included in the Cox multivariate regression analysis. Risk score (P = 0.016) and pathologic_N (P = 0.019) were significantly correlated with prognosis (Table 5).

Nomogram model construction

The nomogram assigned different factors to points, followed by addition to obtain the total points corresponding to the survival rate. This clarified the results of the Coxph regression. The Nomogram model was constructed for Risk score and pathologic_N (Figure 5A). In addition, the consistency index (c-index) of each prognostic factor and composite factor (pathologic_N + Risk score) in the Nomogram model was calculated to fit the Coxph model. As shown in Table 6, composite factor (Nomogram_combined model) fit the Coxph model with a c-index of 0.641 and a maximum significance (P = 3.097E-05). Figure 5B displays the calibration curve of the Nomogram_combined model, which suggested a better prediction ability (close to 45°)

Construction of drug-gene interaction network

The drug-gene interaction network contained 44 drugs, 26 genes (13 up-regulated genes and 13 down-regulated genes), and 55 interactions (Figure 6). In this network, Tubulin Beta 2B Class IIb (TUBB2B), Interleukin 2 Receptor Subunit Alpha (IL2RA), and Interleukin 6 (IL6) were predicted to be interact with more drugs. Verapamil was predicted to be an antagonist of CACNA1A, a gene in the mRNA prognosis risk model.

Discussion

We identified 10 DE-miRNAs, 204 DE-mRNAs, and 17 DE-lncRNAs related to PC and locations of PC occurrence based on the related data in the GEO and TCGA databases. After co-expression analysis and prediction of targets, lncRNA-mRNA pairs, miRNA-lncRNA pairs, and miRNA-mRNA pairs were obtained, and the ceRNA network was constructed. The network contained seven miRNAs, 14 lncRNAs, and 32 mRNAs. The functional enrichment analysis showed that the DE-mRNAs were enriched in pancreatic secretion, calcium ion homeostasis (GRM1), and acute inflammatory response. The main function of miR-183-3p was inflammatory response related processes, while miR-146a-5p participated in

neurotransmission and calcium signaling pathways that involve CACNA1A and GRM1. The main functions implicated for lncRNAs were digestion and pancreatic secretion, including RP1−60O19.1, RP11−462G2.1, and RP11−528G1.2. The functional analysis suggested that the genes associated with pancreatic secretion, inflammatory response, and calcium signaling pathway were abnormally expressed in PC via a ceRNA mechanism.

The mRNA prognosis risk model analysis indicated that Risk score was relatively increased with the increased expressions of GRHL2, PKHD1, UPK1B, and decreased expressions of CACNA1A and GRM1, suggesting poor prognosis. GRHL2 encodes one of the grainyhead like transcription family members, which is crucial in epithelial morphogenesis and epithelial-specific functions [24]. GRHL2 has been reported to be involved in several cancers due to its important role in the regulation of epithelial-mesenchymal transition (EMT) [24, 25]. The overexpression of GRHL2 was observed in liver metastatic cells compared with primary invasive cells of pancreatic ductal adenocarcinoma (PDAC), and the proliferation of liver metastatic PDAC cells was inhibited by regulating EMT under the condition of GRHL2 knockdown [26]. These findings were consistent with our analysis. In addition, GRHL2 expression was significantly increased in pancreatic head compared with pancreatic tail, while the mRNA prognosis risk model indicated that the high expression of GRHL2 corresponded to a high Risk score. These findings might contribute to the accurate diagnosis and treatment of PC patients.

CACNA1A encodes the α1A subunit of voltage-dependent calcium channels, which regulate the transport of calcium ions and various calcium-related pathways [27]. CACNA1A expression in neuronal tissue is abundant to regulate the release of neurotransmitter [28]. GRM1 encodes one of the metabotropic glutamate receptors (mGluRs) that regulate the glutamatergic neurotransmission by G-protein-coupled receptors [29]. L-glutamate serves as the main excitatory neurotransmitter in the central nervous system and can activate mGluRs [30]. Regulation of voltage-dependent $Ca2^{+}$ channels by mGluRs is considered a crucial event in the release of neurotransmitter [31]. Nicotinic acetylcholine receptors are positioned in the cytoplasmic membrane and undergo a conformational change upon the binding of an agonist that leads to the opening of the ion channel followed by the entry of ions into cells [32]. This in turn causes an autocrine neurotransmitter loop and signaling cascades [33]. Data from Schuller et al. indicates the central regulatory role of neurotransmitters and their receptors in PC, and dysfunction of the neurotransmitter receptor may be a vital contributor in the development and progression of PC [6].

Nevertheless, PC is a heterogeneous disease involving individual differences in lifestyle, genetic, and environmental factors. For instance, smoking is a well-established causative factor for PC and is related with decreased survival [34, 35]. Nicotine promotes PC cell proliferation and migration by stimulating the production and release of stress neurotransmitters followed by activation of downstream signal cascades [36]. Nicotine also triggers the self-renewal of PC stem cells by increasing stress neurotransmitters coupled with decreased γ-aminobutyric acid [37]. Psychological stress can also mediate the release of stress neurotransmitters and γ-aminobutyric acid, as well as their downstream effectors [38, 39]. A close association of depression and PC has been reported[40]. Postsynaptic depolarization and calcium ion internal flow are reportedly essential for striatal long-term depression, and

the reduction of neurotransmitter (glutamate) release from presynaptic terminals can promote the expression of long-term depression at striatal synapses [41]. In our study, CACNA1A and GRM1 were enriched in long-term depression, glutamatergic synaptic transmission, and calcium signaling pathway. In addition, comparison of pancreatic body and pancreatic tail cancer revealed that CACNA1A and GRM1 were expressed in low levels in pancreatic head cancer, suggesting that CACNA1A and GRM1 are specifically down-regulated genes in pancreatic head cancer. The mRNA prognosis risk model analysis indicated that the Risk score was relatively increased with decreased expression of CACNA1A and GRM1, suggesting a poor prognosis. The collective data support the conclusion that the decreased expression of CACNA1A and GRM1 might contribute to the progression of pancreatic head cancer by mediating the production and release of excitatory neurotransmitter as well as their downstream effectors.

Notably, CACNA1A and GRM1 were target genes of miR-146a-5p. No reports have focused on the associations between miR-146a-5p and these two genes, but the effects of miR-146a in PC have been reported. The expression of miR-146a was shown to be decreased in PC cells, while its overexpression inhibited tumor cell invasion and metastasis [42]. Moreover, animal experiments revealed that decreased expression of miR-146a can promote cell growth by increasing the expression of epidermal growth factor receptor in PC [43]. Presently, lncRNA AC006369.2 interacted with miR-146a-5p, and was co-expressed with CACNA1A and GRM1. Despite the lack of knowledge of lncRNA AC006369.2, we speculate that AC006369.2 might function as a ceRNA in PC to mediate the expression of CACNA1A and GRM1 by competitively binding miR-146a-5p. The AC006369.2- miR-146a-5p- CACNA1A / GRM1 regulatory axis might be a potentially important mechanism. This must be verified in future studies.

The gene-drug network analysis we conducted supports the role of verapamil as an antagonist of CACNA1A. Verapamil is a calcium ion (Ca^{2+}) channel blocker that inhibits the growth of PC cells by blocking Ca^{2+} influx [44]. Similarly, Zhao et al. suggested that verapamil represses the proliferation and metastasis, and induces apoptosis of chemotherapy-resistant PC cells [45]. Our results, which indirectly verified the important effect of CACNA1A, are consistent with these reports. In both pancreatic body and pancreatic tail cancer, CACN1A was expressed at a significantly high level compared with pancreatic head cancer. As an antagonist of CACNA1A, we suggest that verapamil might be more useful in the treatment of pancreatic body/tail cancer.

Although several novel points were proposed in our study, there were some limitations. All the results were obtained by bioinformatics analysis, so further experimental verification is needed. Secondly, the predicted ceRNA mechanism and the corresponding functions need to be further explored. Thirdly, the gene-drug interactions should be further analyzed.

Conclusions

The five-mRNA prognostic model (GRHL2+CACNA1A+GRM1+UPK1B+PKHD1) was dependable in the prediction of PC survival. The high expression of GRHL2 and low expression of CACNA1A and GRM1 might be indicators of poor prognosis for patients with PC, especially pancreatic head cancer. Verapamil

might be more useful in the treatment of pancreatic body/tail cancer. Finally, the AC006369.2- miR-146a-5p-CACNA1A/GRM1 regulatory axis might be a potentially important mechanism in PC progression.

References

1. Zhang, Q., et al., Pancreatic Cancer Epidemiology, Detection, and Management. Gastroenterology research and practice, 2016. 2016: p. 8962321-8962321.
2. Ilic, M. and I. Ilic, Epidemiology of pancreatic cancer. World journal of gastroenterology, 2016. 22(44): p. 9694-9705.
3. Herreros-Villanueva, M. and L. Bujanda, Non-invasive biomarkers in pancreatic cancer diagnosis: what we need versus what we have. Annals of translational medicine, 2016. 4(7): p. 134-134.
4. van Erning, F.N., et al., Association of the location of pancreatic ductal adenocarcinoma (head, body, tail) with tumor stage, treatment, and survival: a population-based analysis. Acta Oncologica, 2018. 57(12): p. 1655-1662.
5. Lau, M.K., J.A. Davila, and Y.H. Shaib, Incidence and survival of pancreatic head and body and tail cancers: a population-based study in the United States. Pancreas, 2010. 39(4): p. 458-462.
6. Schuller, H.M. and H.A. Al-Wadei, Neurotransmitter receptors as central regulators of pancreatic cancer. Future oncology (London, England), 2010. 6(2): p. 221-228.
7. Ling, Q., et al., The diversity between pancreatic head and body/tail cancers: clinical parameters and in vitro models. Hepatobiliary & Pancreatic Diseases International, 2013. 12(5): p. 480-487.
8. Ruess, D.A., et al., The prognostic influence of intrapancreatic tumor location on survival after resection of pancreatic ductal adenocarcinoma. BMC surgery, 2015. 15(1): p. 123.
9. Korc, M., et al., Tobacco and alcohol as risk factors for pancreatic cancer. Best practice & research. Clinical gastroenterology, 2017. 31(5): p. 529-536.
10. Hidalgo, M., Pancreatic cancer. New England Journal of Medicine, 2010. 362(17): p. 1605-1617.
11. Jones, S., et al., Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene. Science (New York, N.Y.), 2009. 324(5924): p. 217-217.
12. Fu, Z., et al., LncRNA HOTTIP modulates cancer stem cell properties in human pancreatic cancer by regulating HOXA9. Cancer Letters, 2017. 410: p. S0304383517305694.
13. Zhou, Y., et al., LncRNA UCA1 impacts cell proliferation, invasion, and migration of pancreatic cancer through regulating miR-96/FOXO3. Iubmb Life, 2018. 70(4).
14. Yu, S., et al., miRNA-96 suppresses KRAS and functions as a tumor suppressor gene in pancreatic cancer. Cancer research, 2010. 70(14): p. 6015-6025.
15. Miyamae, M., et al., Plasma microRNA profiles: identification of miR-744 as a novel diagnostic and prognostic biomarker in pancreatic cancer. British journal of cancer, 2015. 113(10): p. 1467-1476.
16. Li, C., et al., A ROR1-HER3-lncRNA signalling axis modulates the Hippo-YAP pathway to regulate bone metastasis. Nature Cell Biology, 2017. 19(2): p. 106-119.

17. Cui, Y., et al., Upregulated lncRNA SNHG1 contributes to progression of non-small cell lung cancer through inhibition of miR-101-3p and activation of Wnt/Î²-catenin signaling pathway. Oncotarget, 2017. 8(11): p. 17785-17794.
18. Wang, H., et al., STAT3-mediated upregulation of lncRNA HOXD-AS1 as a ceRNA facilitates liver cancer metastasis by regulating SOX4. Molecular Cancer, 2017. 16(1): p. 136.
19. Gao, S., et al., ROR functions as a ceRNA to regulate Nanog expression by sponging miR-145 and predicts poor prognosis in pancreatic cancer. Oncotarget, 2016. 7(2): p. 1608-1618.
20. Tyner, C., et al., The UCSC genome browser database: 2017 update. Nucleic acids research, 2016. 45(D1): p. D626-D634.
21. Nikolayeva, O. and M.D. Robinson, edgeR for differential RNA-seq and ChIP-seq analysis: an application to stem cell biology, in Stem Cell Transcriptional Networks. 2014, Springer. p. 45-79.
22. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics, 2010. 26(1): p. 139-140.
23. Yu, G., et al., clusterProfiler: an R package for comparing biological themes among gene clusters. Omics: a journal of integrative biology, 2012. 16(5): p. 284-287.
24. Xiang, J., et al., Grhl2 reduces invasion and migration through inhibition of TGFβ-induced EMT in gastric cancer. Oncogenesis, 2017. 6(1): p. e284-e284.
25. Werner, S., et al., Dual roles of the transcription factor grainyhead-like 2 (GRHL2) in breast cancer. The Journal of biological chemistry, 2013. 288(32): p. 22993-23008.
26. Nishino, H., et al., Grainyhead-like 2 (GRHL2) regulates epithelial plasticity in pancreatic cancer progression. Cancer medicine, 2017. 6(11): p. 2686-2696.
27. Kors, E.E., et al., Delayed cerebral edema and fatal coma after minor head trauma: role of the CACNA1A calcium channel subunit gene and relationship with familial hemiplegic migraine. Annals of neurology, 2001. 49(6): p. 753-760.
28. Randall, A. and C.D. Benham, Recent advances in the molecular understanding of voltage-gated Ca2+ channels. Molecular and Cellular Neuroscience, 1999. 14(4-5): p. 255-272.
29. Esseltine, J.L., et al., Somatic Mutations in GRM1 in Cancer Alter Metabotropic Glutamate Receptor 1 Intracellular Localization and Signaling. Molecular Pharmacology, 2013. 83(4): p. 770-780.
30. Namkoong, J., et al., Metabotropic Glutamate Receptor 1 and Glutamate Signaling in Human Melanoma. Cancer Research, 2007. 67(5): p. 2298-2305.
31. Millán, C., et al., Co-expression of metabotropic glutamate receptor 7 and N-type Ca2+ channels in single cerebrocortical nerve terminals of adult rats. Journal of Biological Chemistry, 2003. 278(26): p. 23955-23962.
32. Lindstrom, J., et al., Structure and function of neuronal nicotinic acetylcholine receptors, in Progress in brain research. 1996, Elsevier. p. 125-137.
33. Al-Wadei, M.H., H.A. Al-Wadei, and H.M. Schuller, Pancreatic cancer cells and normal pancreatic duct epithelial cells express an autocrine catecholamine loop that is activated by nicotinic acetylcholine

receptors α3, α5, and α7. Molecular Cancer Research, 2012. 10(2): p. 239-249.
34. Zhu, B., et al., Genetic variants in the SWI/SNF complex and smoking collaborate to modify the risk of pancreatic cancer in a Chinese population. Molecular carcinogenesis, 2015. 54(9): p. 761-768.
35. Yuan, C., et al., Cigarette smoking and pancreatic cancer survival. Journal of Clinical Oncology, 2017. 35(16): p. 1822.
36. Al-Wadei, M.H., H.A. Al-Wadei, and H.M. Schuller, Effects of chronic nicotine on the autocrine regulation of pancreatic cancer cells and pancreatic duct epithelial cells by stimulatory and inhibitory neurotransmitters. Carcinogenesis, 2012. 33(9): p. 1745-1753.
37. Al-Wadei, M.H., et al., Nicotine induces self-renewal of pancreatic cancer stem cells via neurotransmitter-driven activation of sonic hedgehog signalling. European journal of cancer (Oxford, England : 1990), 2016. 52: p. 188-196.
38. Zabora, J., et al., The prevalence of psychological distress by cancer site. Psycho-oncology, 2001. 10(1): p. 19-28.
39. Bettison, T.M., et al., Understanding the pathophysiology of psychological distress and pancreatic cancer: a systematic review. Pancreas, 2018. 47(4): p. 376-381.
40. Boyd, A.D. and M. Riba, Depression and pancreatic cancer. Journal of the National Comprehensive Cancer Network, 2007. 5(1): p. 113-116.
41. Choi, S. and D.M. Lovinger, Decreased probability of neurotransmitter release underlies striatal long-term depression and postnatal development of corticostriatal synapses. Proceedings of the National Academy of Sciences of the United States of America, 1997. 94(6): p. 2665-2670.
42. Li, Y., et al., miR-146a suppresses invasion of pancreatic cancer cells. Cancer research, 2010. 70(4): p. 1486-1495.
43. Ali, S., et al., Deregulation of miR-146a expression in a mouse model of pancreatic cancer affecting EGFR signaling. Cancer letters, 2014. 351(1): p. 134-142.
44. Sato, K., et al., Inhibitory effect of calcium channel blockers on growth of pancreatic cancer cells. Pancreas, 1994. 9(2): p. 193-202.
45. Zhao, L., et al., Verapamil inhibits tumor progression of chemotherapy-resistant pancreatic cancer side population cells. International journal of oncology, 2016. 49(1): p. 99-110.

Declarations

Competing Interests

The authors declare that no conflicts of interest exist.

Tables

Table 1. Numbers of differentially expressed miRNAs, mRNAs, and lncRNAs in each comparable group.

Comparable group	miRNA	mRNA	lncRNA
PC vs. Normal	187	1661	661
Pancreatic body vs. Pancreatic head	141	760	83
Pancreatic tail vs. Pancreatic head	54	461	53
Pancreatic tail vs. Pancreatic body	36	136	24

Table 2. Top 10 enriched GO_BP terms and KEGG pathways

Terms ID	Description	P value	Count	Gene
GO:0002526	acute inflammatory response	3.35E-09	13	CCR7, SAA1, C4BPA, SERPINA3, REG3A, DEFB1, LBP, REG3G, IL6, CFB, SAA2, CFP, F3
GO:0007586	digestion	3.88E-09	14	CELA3A, CLPS, PPY, PRSS2, PLA2G1B, CTRL, SERPINA3, CTRB1, PNLIP, AMY2A, CTRB2, PRSS3, AMY2B, PRSS1
GO:0070098	chemokine-mediated signaling pathway	2.96E-07	9	CCL22, CCL19, XCR1, CCR7, CCR4, CCL23, CCL18, CXCR1, CXCL11
GO:0009235	cobalamin metabolic process	3.12E-06	5	CTRC, CTRB1, CTRB2, PRSS3, PRSS1
GO:0060326	cell chemotaxis	3.63E-06	13	CCL22, CCL19, CCR7, CCL23, CCL18, CXCR1, PLA2G1B, SAA1, LBP, CXCL11, IL6, HOXB9, SAA2
GO:0030593	neutrophil chemotaxis	4.10E-06	8	CCL22, CCL19, CCR7, CCL23, CCL18, PLA2G1B, SAA1, LBP
GO:1990266	neutrophil migration	6.97E-06	8	CCL22, CCL19, CCR7, CCL23, CCL18, PLA2G1B, SAA1, LBP
GO:0050727	regulation of inflammatory response	7.05E-06	14	CCR7, CCL23, CCL18, TLR10, CD28, SAA1, C4BPA, LBP, IL6, IL2RA, FABP4, CFB, CFP, PTGIS
GO:0055074	calcium ion homeostasis	7.82E-06	16	CCL19, PKHD1, XCR1, GRM1, CCR7, CCR4, CCL23, PLA2G1B, SAA1, KEL, TRPV6, CXCL11, SCGN, CACNA1A, CD52, CD19
GO:0006874	cellular calcium ion homeostasis	2.15E-05	15	CCL19, PKHD1, XCR1, GRM1, CCR7, CCR4, CCL23, PLA2G1B, SAA1, KEL, CXCL11, SCGN, CACNA1A, CD52, CD19
ID	Description	pvalue	Count	geneID
hsa04972	Pancreatic secretion	1.13E-18	20	CELA3A, CPB1, PLA2G10, PRSS2, PLA2G1B, CELA3B, CELA2B, CTRL, CTRB1, SLC4A4, CELA2A, PNLIP, CPA2, AMY2A, PRSS3, CPA1, CFTR, AMY2B, PNLIPRP1, PRSS1
hsa04974	Protein digestion and absorption	1.24E-11	14	CELA3A, SLC3A1, CPB1, PRSS2, CELA3B, CELA2B, CTRL, CTRB1, CELA2A, XPNPEP2, CPA2, PRSS3, CPA1, PRSS1
hsa04975	Fat digestion and absorption	0.000217645	5	CLPS, PLA2G10, PLA2G1B, PNLIP, PNLIPRP1
hsa04062	Chemokine signaling pathway	0.000250485	10	CCL22, CCL19, XCR1, CCR7, CCR4, CCL23, CCL18, CXCR1, CXCL11, WAS
hsa00982	Drug metabolism - cytochrome P450	0.000420645	6	UGT2B17, GSTA2, ADH4, GSTA1, UGT2B15, CYP2B6
hsa00980	Metabolism of xenobiotics by cytochrome P450	0.000563266	6	UGT2B17, GSTA2, ADH4, GSTA1, UGT2B15, CYP2B6
hsa04060	Cytokine-cytokine receptor interaction	0.000626393	12	CCL22, CCL19, XCR1, CCR7, CCR4, CCL23, CCL18, CXCR1, IL2RG, CXCL11, IL6, IL2RA
hsa00590	Arachidonic acid metabolism	0.001610217	5	PLA2G10, PLA2G1B, ALOX15, CYP2B6, PTGIS

Table 3. Construction of mRNA prognosis risk model

Model	P value	AUC
GRHL2	0.023	1.000
GRHL2+CACNA1A	0.019	1.000
GRHL2+CACNA1A+GRM1	0.015	0.997
GRHL2+CACNA1A+GRM1+UPK1B	0.024	0.997
GRHL2+CACNA1A+GRM1+UPK1B+PKHD1	**0.006**	**0.999**
GRHL2+CACNA1A+GRM1+UPK1B+PKHD1+CFTR	0.015	0.997
GRHL2+CACNA1A+GRM1+UPK1B+PKHD1+CFTR+CPLX2	0.015	0.995
GRHL2+CACNA1A+GRM1+UPK1B+PKHD1+CFTR+CPLX2+KBTBD12	0.008	0.990
GRHL2+CACNA1A+GRM1+UPK1B+PKHD1+CFTR+CPLX2+KBTBD12+TMPRSS13	0.007	0.988
GRHL2+CACNA1A+GRM1+UPK1B+PKHD1+CFTR+CPLX2+KBTBD12+TMPRSS13+OLFM4	0.009	0.992

Table 4. The results of Cox univariate regression analysis

	TCGA				SEER			
	HR	lower.95	upper.95	p.val	HR	lower.95	upper.95	p.val
age	1.023	1.001	1.045	**0.045**	1.024	1.021	1.027	**0.000**
gender	0.825	0.540	1.261	0.375	1.060	0.999	1.125	0.053
neoplasm_histologic_grade	1.394	1.031	1.884	**0.031**	1.462	1.407	1.519	**0.000**
Head	1.623	0.859	3.064	0.135	1.013	0.954	1.075	0.677
Body	0.466	0.170	1.274	0.137	1.053	0.980	1.133	0.161
Tail	0.820	0.376	1.787	0.618	0.937	0.872	1.005	0.070
pathologic_T	1.547	0.954	2.507	0.077	0.904	0.874	0.936	0.000
pathologic_N	2.483	1.437	4.289	**0.001**	0.910	0.858	0.966	**0.002**
tumor_stage	1.312	0.872	1.972	0.192	1.421	1.327	1.523	0.000
RiskScore	2.141	1.272	3.604	**0.004**				

Table 5. The results of Cox multivariate regression analysis

	HR	lower.95	upper.95	p.value
RiskScore	2.054	1.142	3.693	**0.016**
age	1.020	0.998	1.041	0.070
neoplasm_histologic_grade	1.237	0.916	1.670	0.166
pathologic_N	1.936	1.116	3.362	**0.019**

Table 6. The consistency index (c_index) of Nomogram factors fitting Coxph model

	c_index	lower.95	upper.95	P value
Nomogram_combind model	0.641	0.575	0.707	3.097E-05
Prognostic model	0.613	0.548	0.679	7.068E-04
pathologic_N	0.670	0.539	0.800	1.080E-02

Figures

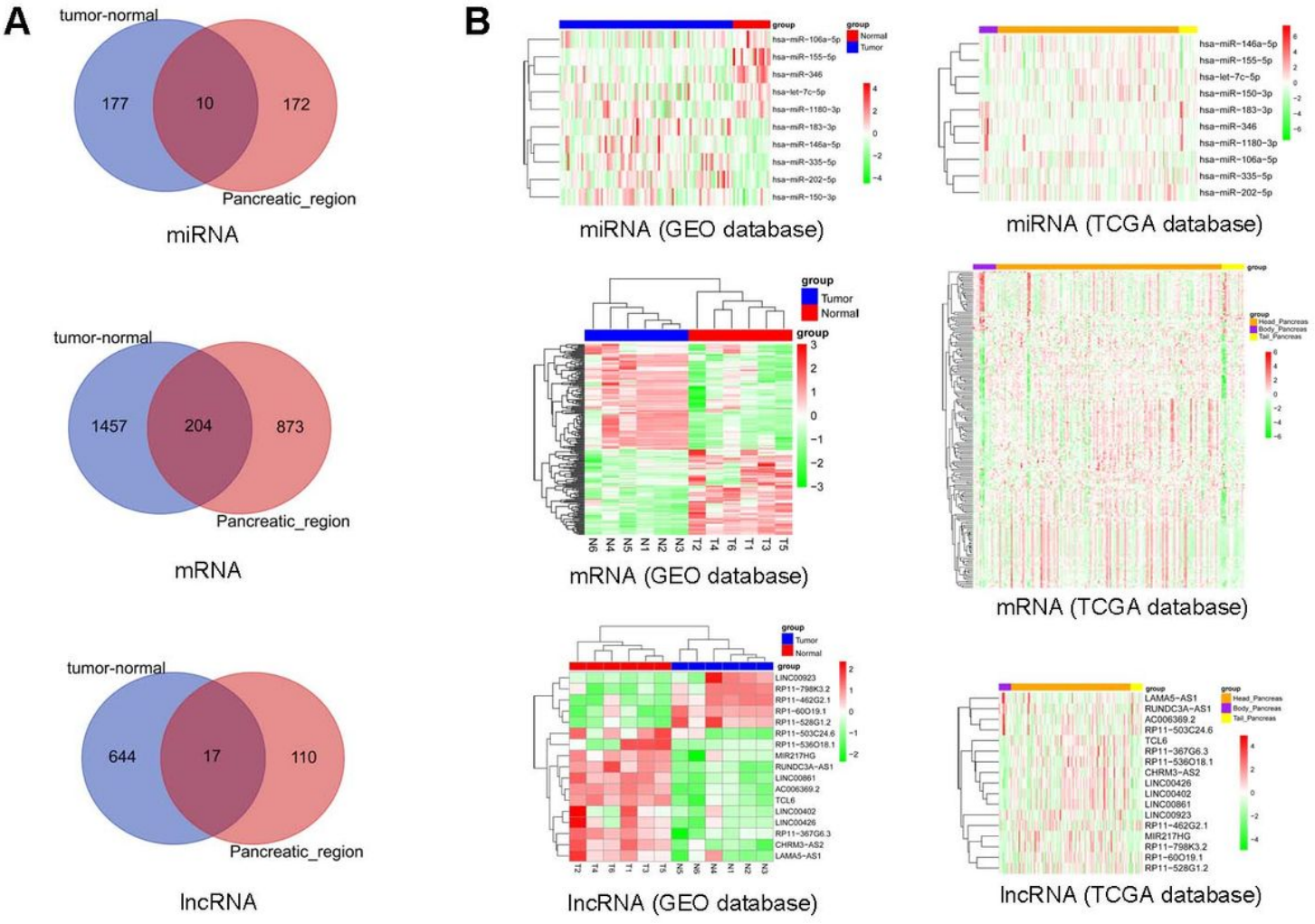

Figure 1

Results of differential expression analysis A, Venn diagram of differentially expressed miRNAs, mRNAs, and lncRNAs. B, Heatmaps of differentially expressed miRNAs, mRNAs, and lncRNAs in the two databases. Red represents normal samples, blue denotes tumor samples, yellow denotes pancreatic tail samples, orange denotes pancreatic head samples, and purple denotes pancreatic body samples.

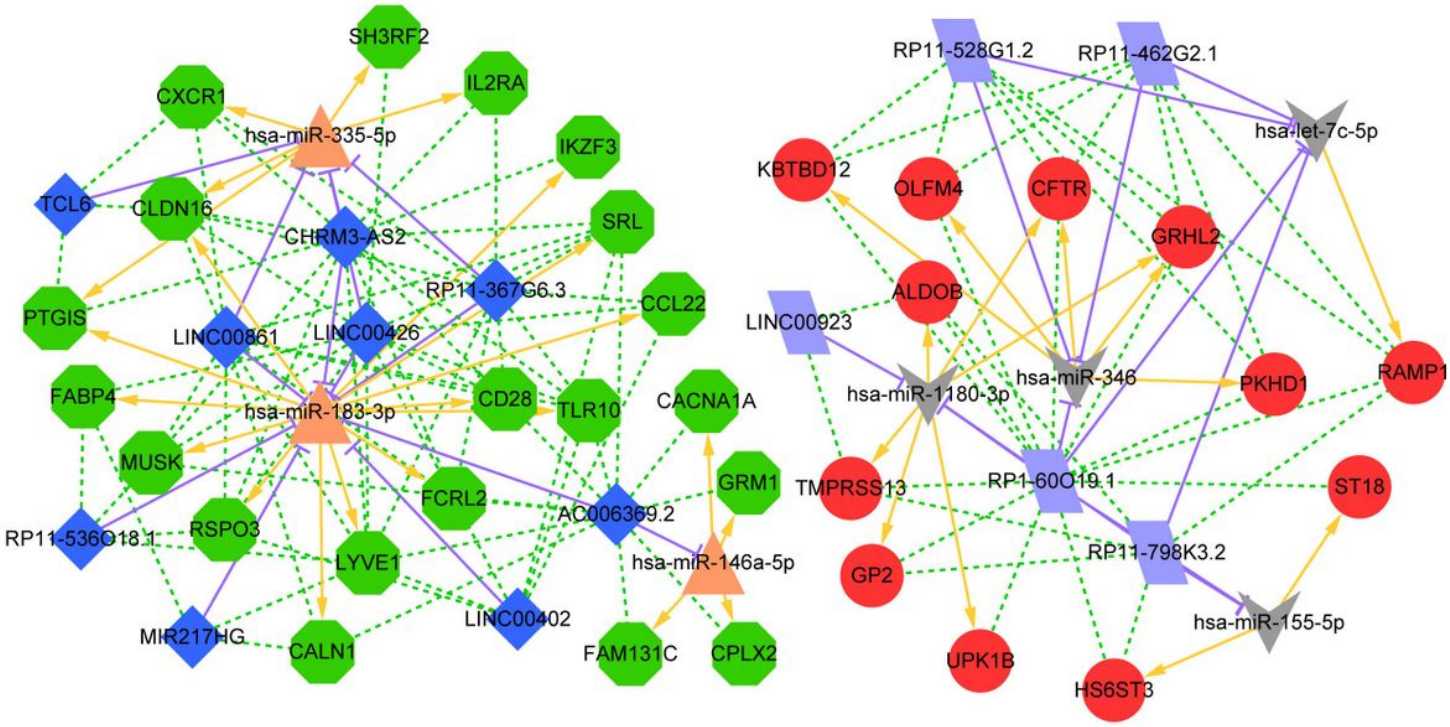

Figure 2

The ceRNA network Red circles represent up-regulated mRNAs and green octagonals represent down-regulated mRNAs. The light purple parallelogram represents up-regulated lncRNA, while the dark blue diamond represents down-regulated lncRNAs. Gray inverted triangles represent down-regulated miRNAs, while yellow positive triangles represent up-regulated miRNAs. Purple T-shaped lines represent the regulatory relationship between lncRNA-miRNA. Yellow arrows represent the regulatory relationship between miRNA-mRNA. Green dotted lines represent the co-expression relationship between mRNA and lncRNA.

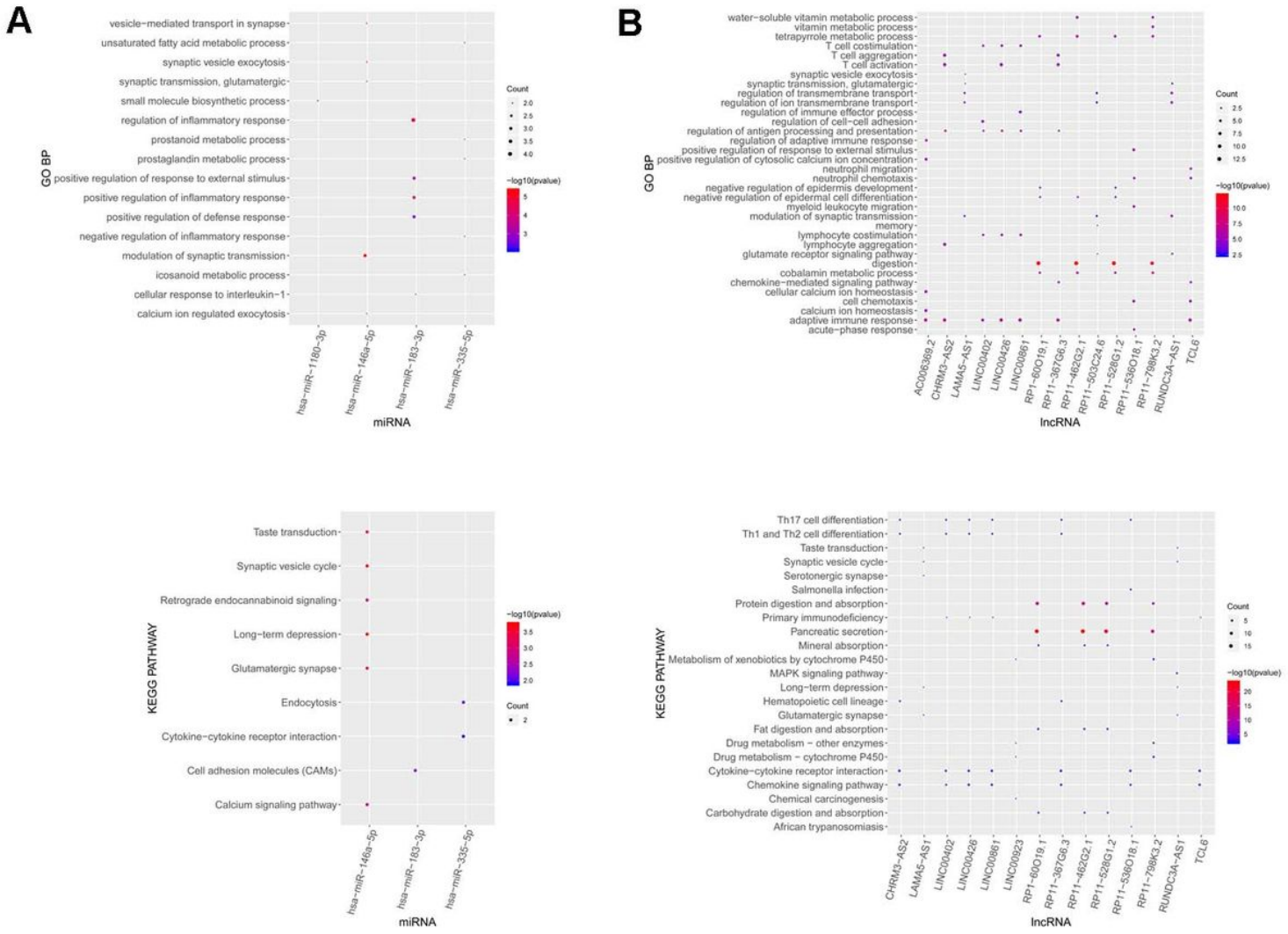

Figure 3

Results of functional enrichment analysis for miRNA and lncRNA A, Functional enrichment analysis including GO BP functional annotation(above) and KEGG pathway enrichment(below) for miRNA. B, Functional enrichment analysis including GO BP functional annotation(above) and KEGG pathway enrichment(below) for lncRNA. Colors ranging from blue to red indicate the decrease of p value, and bubble size indicates the number of genes.

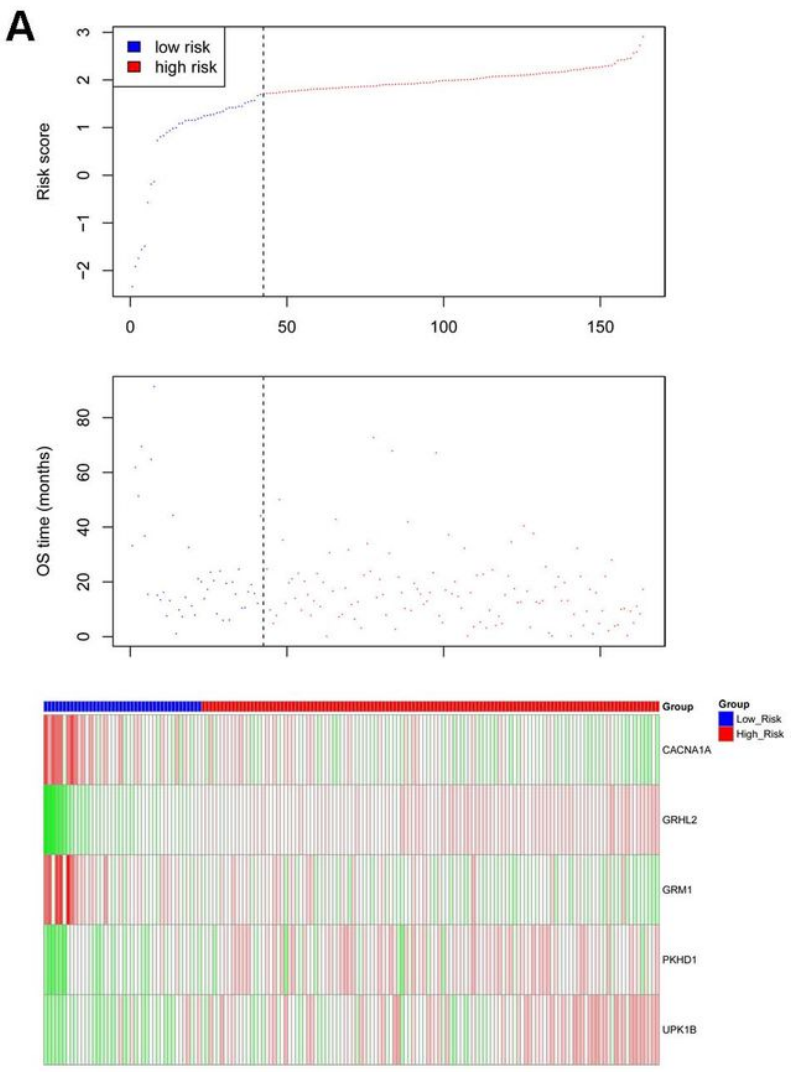

Figure 4

Results of mRNA prognosis risk model construction A, The distribution of Risk score, survival time, and gene expression values of the high risk and low risk samples identified by the Risk model constructed by GRHL2+CACNA1A+GRM1+UPK1B+PKHD1. B, Kaplan-Meier survival curve of the high risk and low risk groups.

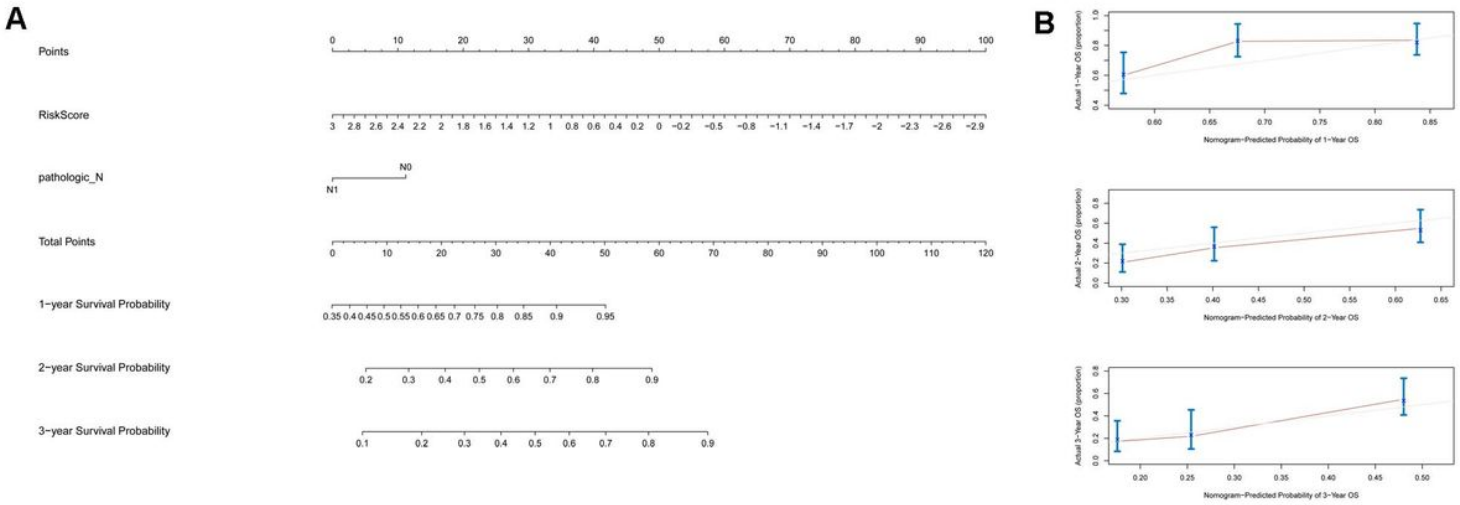

Figure 5

Nomogram predicting 1-, 2- and 3-year overall survival for patients with PC A, Nomogram and B, calibration curve of the Nomogram_combined model.

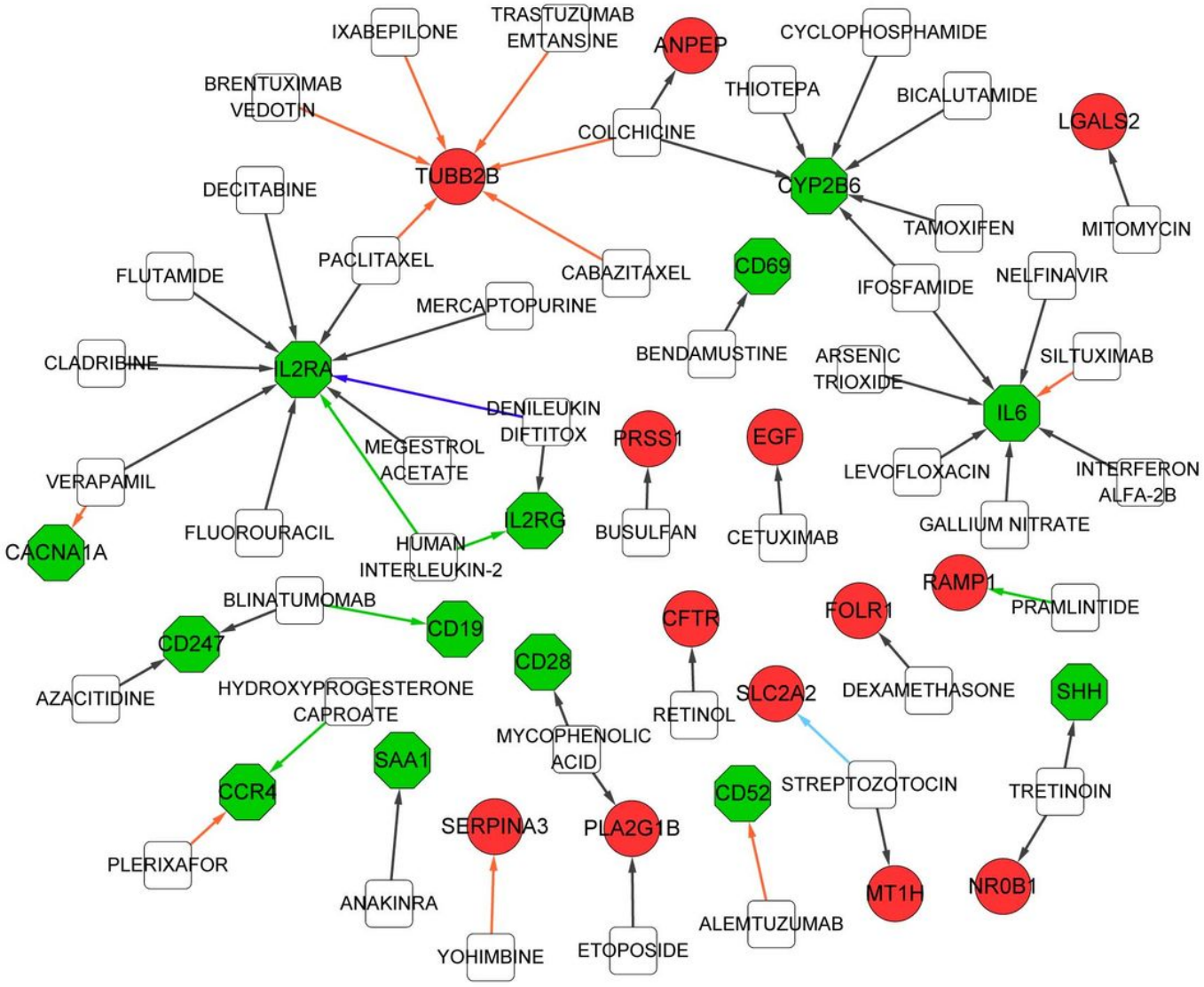

Figure 6

Gene-drug interaction network Red circles represent up-regulated mRNAs, green octagons represent down-regulated mRNAs, green arrows represent agonist effect, yellow arrows represent inhibitor effect, light blue arrows represent ligand effect, and dark blue represents effect of adhesives.