Identification of Genomic Instability Related Lncrna Signature Associated with Immune Microenvironment in Pancreatic Cancer

Xiaole Zhu

Jiangsu Province Hospital and Nanjing Medical University First Affiliated Hospital

Rong Yu

Jiangsu Province Hospital and Nanjing Medical University First Affiliated Hospital

Yunpeng Peng

Jiangsu Province Hospital and Nanjing Medical University First Affiliated Hospital

Chaoqun Hou

Jiangsu Province Hospital and Nanjing Medical University First Affiliated Hospital

Chenyuan Shi

Jiangsu Province Hospital and Nanjing Medical University First Affiliated Hospital

Lingdi Yin

Jiangsu Province Hospital and Nanjing Medical University First Affiliated Hospital

Yi Miao

Jiangsu Province Hospital and Nanjing Medical University First Affiliated Hospital

Qiang Li (✉ liqiang020202@163.com)

Jiangsu Province Hospital and Nanjing Medical University First Affiliated Hospital <https://orcid.org/0000-0001-5129-2603>

Research article

Keywords: genome instability, somatic mutation, pancreatic cancer, long non-coding RNA, prognostic signature, tumor immune environment

Posted Date: July 29th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-741589/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Increasing evidence suggested that the critical roles for lncRNAs in the maintenance of genomic stability. However, the identification of genomic instability related lncRNA signature (GILncSig) and their clinical significance in tumor immune microenvironment of pancreatic cancer remain largely unexplored.

Methods: In the present study, a systematic analysis of lncRNA expression profiles and somatic mutation profiles was performed in pancreatic cancer patients from TCGA. We performed co-expression network and Gene Ontology (GO) enrichment analyses to determine the potential functions and pathways involved in lncRNAs are associated with genomic instability. We then development a risk score model to describe the characteristics of the model and verify its prediction accuracy. ESTIMATE algorithm, single-sample gene set enrichment analysis (ssGSEA), and CIBERSORT analysis were employed to reveal the characteristics of tumor immune microenvironment in pancreatic cancer. The correlation of risk signature with immune infiltration and immune checkpoint blockade (ICB) therapy was analyzed.

Results: We identified 206 GILncSig, of which five were screened to develop a prognostic GILncSig model. Multivariate Cox regression analysis and stratified analysis revealed that the prognostic value of the GILncSig was independent of other clinical variables. ROC analysis suggested that GILncSig is better than the existing lncRNA-related signatures in predicting survival. Additionally, the prognostic performance of the GILncSig was also found to be favorable in patients carrying wild-type KRAS, TP53 and SMAD4. Besides, a nomogram exhibited appreciable reliability for clinical application in predicting the prognosis of patients. Finally, the risk score significantly correlated with immune score, immune-related signature, infiltrating immune cells (i.e. B cells, etc.), and ICB key molecules (i.e. CTLA4, etc.).

Conclusion: In summary, the GILncSig identified by us may have crucial role in immune cell infiltration, immunotherapy and important indicator for clinical stratification management and therapy decisions for pancreatic cancer patients.

Introduction

Pancreatic cancer is one of the deadliest cancers, ranking as the fourteenth most common cancer and the seventh leading cause of cancer mortality worldwide. Due to the lack of obvious early symptoms, pancreatic cancer usually presents at an advanced stage, which results in a 5-year survival rate as low as 6% (ranges from 2–9%)[1]. Despite the great advances in surgery, chemotherapy, and radiotherapy for pancreatic cancer have been made in the past few years, long-term survival and prognosis remain terrible, with more than 80 percent of patients facing recurrence after resection[2]. More recently, a large number of previous studies have analyzed the relationship between the expression of molecular marker and clinicopathology and long-term survival in the molecular mechanism of pancreatic cancer. However, their impact on patient early diagnosis and treatment is still limited[3]. Therefore, searching for new prognostic markers that can predict the poor outcome of patients may become the target of intervention, and provide new treatment strategies for the treatment of pancreatic cancer.

Genomic instability refers to an increased tendency of the genome to acquire mutations, which is typically conferred by some mechanism dysfunction, such as DNA damage repair, DNA replication, transcription and so on. Genomic instability is a hallmark of cancer and is related to cancer initiation and progression[4]. In addition, genome stability status is also associated with survival and can be used as a prognostic marker for cancer patients[5]. Long non-coding RNAs (lncRNAs) are arbitrarily considered as non-protein coding transcripts over 200 nucleotides in length[6].

There is increasing evidence suggesting that lncRNAs are involved in a variety of biological processes and play a critical role in genome regulation[6–8]. Noticeably, the dysregulation of lncRNAs has been established to be associated with many complex diseases, including cancers[9–11]. A number of lncRNAs are abnormally expressed in tumor tissues, which have been considered as oncogenes, such as MALAT1[12], HOTAIR[13], H19[14], MEG3[15]. The main function of lncRNA is to regulate gene expression and indicate the tumor status better than the protein-coding RNAs, so it can be used as a novel biomarker with diagnostic and prognostic significance[16]. Currently, several lncRNA signatures have been developed in various cancers to predict patient prognosis with great predictive performance, including lung cancer[17], head and neck squamous cell carcinoma[18], ovarian cancer[19] and breast cancer[20, 21]. Recently, Lee et al. analyzed a non-coding RNA activated by DNA damage (or NORAD), maintained genomic stability by isolating PUMILIO protein[22]. Hu et al. reported that GUARDIN, as a p53-responsive lncRNA, kept genomic integrity under both stable and exposed status[23]. These results demonstrated the importance role of lncRNAs in maintaining genomic stability, but the lncRNAs associated with genomic instability need to be further explored.

In addition, studies have shown that immune cells act as tumor inhibitor or tumor promoter and may function as important players in the tumor immune microenvironment (TIME). Genomic instability has been termed as a promising indicator for predicting responsiveness to immune checkpoint blockade based on numerous researches.

Therefore, we constructed a GILncSig and to investigate whether the lncRNA signature could reflect the tumor immune microenvironment, and serve as an effective prognostic predictor for patients with pancreatic cancer.

Materials And Methods

1. Patient dataset:

The clinical information, RNA-seq expression data, lncRNA transcriptional profiles and somatic mutation information of patients with pancreatic cancer were obtained from The Cancer Genome Atlas (TCGA) project (<https://cancergenome.nih.gov/>)[24]. A total of 171 TCGA pancreatic cancer patients with lncRNA expression profiles somatic mutations, survival information and clinical features were utilized in our study. TCGA patients with pancreatic cancer were divided into an 84-sample training set and 87-sample testing set. The training set was used to identify the prognostic lncRNA signature and establish the prognostic risk model, while the testing set was used to independently validate its prognostic value.

2. Identification of genomic instability-associated lncRNAs

In order to identify with genomic instability-associated lncRNAs, a computational framework was constructed based on the lncRNAs expression profiles and somatic mutation profile of pancreatic cancer patients. As shown in Figure 1, the cumulative number of somatic mutations for per sample was calculated and arranged in descending order. The first 25% of patients were defined as genomic instability group (GU group), and the last 25% were defined as genomic stability group (GS group). Then compared the expression profiles of lncRNAs between GU group and GS group by the significance analysis of microarrays (SAM) method. The differentially expressed lncRNAs screened out by the filter of fold change and permutation correction were defined as genomic instability-related lncRNAs (fold change > 1.5 or <0.67 and false discovery rate (FDR) adjusted $P < 0.05$).

3. Statistical Analysis

We carried out a univariate regression analysis to determine the relationship between the expression level of lncRNAs and the overall survival of the training set. Those lncRNAs with p-value less than 0.05 were considered as the candidate prognostic lncRNAs of pancreatic cancer whose expression levels were significantly associated with overall survival of pancreatic cancer patients. In order to assess the contribution of those candidate lncRNA as an independent prognostic factor for survival, multivariate Cox regression analysis was further performed. A P value less than 0.05 was considered as significant. A prognostic risk score model of genomic instability-related lncRNAs signature (GILncSig) was constructed based on the expression level of lncRNAs and multivariate Cox regression coefficient to predict the prognosis of patients with pancreatic cancer as follows: $GILncSig(\text{patients}) = \sum_{i=1}^n \beta_i \times expression(lncRNA_i)$. In our formula, GILncSig (patients) is the prognostic risk score for pancreatic cancer patients. lncRNA_i is the each prognostic lncRNAs. Coefficient (lncRNA_i) represents the corresponding coefficient of multivariate Cox regression analysis, and expression(lncRNA_i) is the expression level of lncRNA_i.

According to above formula, the lncRNA expression-based risk scores for pancreatic cancer patients could be calculated and divided patients into high-risk and low-risk group with the cutoff of the median risk score from the training set. Kaplan-Meier survival curves was utilized to estimate the survival rate of the different patient groups, and the survival differences between high-risk group and low-risk group was assessed by the log-rank test. Time-dependent ROC analysis for overall survival was used to assess the performance of prognostic risk model for time dependent disease outcomes. Multivariate Cox regression and stratified analysis were performed to determine whether the GILncSig was independent of other clinical variables. Hazard ratio (HR) and 95% confidence intervals (CI) were estimated by Cox proportional hazards regression model. A nomogram was built in the training set to predict the 1-, 2-, and 3-year survival based on the results of multivariate cox regression analysis by R “rms” and “survival” package and applied to the testing set and the entire TCGA set for verification. The corrected plot was used to assess the prognostic accuracy of the nomogram. All statistical analyses were performed using R software and Bioconductor.

4. Functional Enrichment Analysis

We calculated Pearson correlation coefficient to evaluate their correlation by using paired lncRNA and mRNA expression profiles, and then established a lncRNA-mRNA co-expression network. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses of the co-expressed protein-coding genes with prognostic lncRNAs were performed to predict the biological function of the differentially expressed lncRNAs using clusterProfiler software in R- version 3.5.2[25].

5. Tumor Immune related analysis

For reflect the characteristics of tumor immune microenvironment, R package “ESTIMATE” was utilized to calculate Scores of immune and stromal cells. Immune infiltration information contains each tumor sample's immune cell fraction were obtained from Tumor Immune Estimation Resource (TIMER) (<https://cistrome.shinyapps.io/timer/>).

The correlation of tumor immune cell infiltrating with prognostic risk signature was further analyzed. We selected six key genes of immune checkpoint blockade-related genes in pancreatic cancer to investigate the potential role of lncRNA-based signature in ICB therapy of pancreatic cancer.

Results

1. Identification of genome instability-associated lncRNAs in patients with pancreatic cancer

To detect the potential Genome instability-related lncRNAs, the cumulative number of somatic mutations in each patient with pancreatic cancer was calculated from TCGA. The first 25% (n =43) and the last 25% (n =40) patients were classified into GU group and GS group by the descending order of cumulative number. Then the lncRNA expression profiles in GU group and GS group were analyzed by unsupervised clustering, the result show that total 206 lncRNAs were found to be significantly differentially expressed (Figure 2A). All patients with pancreatic cancer in TCGA were divided into GU-like group and GS-like group by unsupervised hierarchical clustering analysis based on the expression levels of the 206 differentially expressed lncRNAs. The cumulative number of somatic mutations was higher in GU-like group and lower in GS-like group (Figure 2B). As shown in Figure 2C, more mutated genes exist in GU-like group ($P < 0.001$ Mann-Whitney U test). As UBQLN4 gene is one of the driving factors of gene instability, the expression level of UBQLN4 gene in GU-like group and GS-like group was compared. The results showed that there was significant difference in the expression level of UBQLN4 between the two groups, and the expression level of UBQLN4 in GU-like group was significantly higher than that in GS like group. ($P < 0.001$, Mann–Whitney U test, Figure 2D).

To better understand the biological significance of the 206 differentially expressed lncRNAs, functional enrichment analysis was performed to predict potential functions. We selected the protein coding genes (PCGs) most related to the expression of each lncRNAs to construct an lncRNA-mRNA co-expression network (Figure 3A). According to the enriched results of the lncRNA-correlated PCGs, GO biological process (e.g., cellular component (CC), DNA binding in the molecular function (MF), and metabolism in the biological process (BP)) and KEGG pathway (e.g., MAPK signaling pathway, cAMP signaling pathway, Pancreatic secretion and Endocrine resistance) were annotated to be associated with genome instability (Figure 3B-C). Based on the above results, it is considered that the 206 lncRNAs were involved in the genomic instability-related biological process, and their altered expression may destruct the genomic stability of cells. Therefore, the 206 differentially expressed lncRNAs were recognized as candidate lncRNAs with genomic instability in pancreatic cancer.

2. Acquisition of a genomic instability-associated lncRNA prognostic signature from the training set

In order to screen out the prognostic lncRNAs with independent value, we performed univariate Cox proportional hazard regression analysis to analyze the relationship between expression levels of 206 lncRNA and OS in the training set, 17 candidate prognostic lncRNAs were found to be significantly associated with the prognosis of pancreatic cancer patients (Figure 4A). Furthermore, multivariate Cox proportional hazards regression was used to analysis on 17 candidate prognostic lncRNAs. Based on the multiCox model (Figure 4B), 5 of 17 candidate lncRNAs including AL121772.1, BX640514.2, LINC01133, AC087752.3 and LYPLAL1-AS1 were found to retain their prognostic significance and thus were identified as independent prognostic lncRNAs ($P < 0.05$). Among five prognostic lncRNAs, one lncRNAs (AC087752.3) having negative coefficients was shown to be a protective factor whose high expression level was closed associated with a longer survival, whereas the remaining four lncRNAs (AL121772.1, BX640514.2, LINC01133 and LYPLAL1-AS1) with positive coefficients tended to be prognostic risk factors and their high expression were associated with a shorter survival. A risk score model of genomic instability-associated lncRNA signature (GILncSig) based on the expression level and multivariate Cox analysis regression coefficients of 5 independent prognostic lncRNAs was generated to predict the outcome of pancreatic cancer patients as follows: $GILncSig = (-1.61 \times \text{expression value of AC087752.3}) + (0.63 \times \text{expression value of AL121772.1}) + (0.39 \times \text{expression value of BX640514.2}) + (0.02 \times \text{expression value of LINC01133}) + (0.38 \times \text{expression value of LYPLAL1-AS1})$. According to the GILncSig model, the prognostic risk score was computed for each patient in the training set. Using the median risk score as the cutoff point, all patients of training set were classified into a high-risk group (n = 38) and a low-risk group (n = 46). The Kaplan-Meier analysis indicated that the overall survival was

significantly different between the two risk groups and patients in low-risk subgroup had markedly longer overall survival than those in the high-risk group ($P=0.009$, log-rank test, Figure 5A). The time-dependent receiver operating characteristic (ROC) curves analysis for GILncRNA prognostic model achieved an area under the curve (AUC) of 0.653 at 1 year of overall survival (Figure 5C). These results demonstrated the GILncRNA had better prognosis prediction performance in patients with pancreatic cancer. Then we ranked the risk scores of patients in the training set. Figure 5B showed the expression pattern of the 5 Independent prognostic lncRNAs, the expression level of UBQLN4 and the count of somatic mutations. We found that for patients with high-risk scores, the expression levels of four risk lncRNAs (AL121772.1, BX640514.2, LINC01133, LYPLAL1-AS1) were up-regulated, while one protective lncRNA (AC087752.3) was expressed at a low level. In contrast, these prognostic lncRNAs expressed the opposite patterns in patients with low-risk scores. Similarly, there were significant differences in UBQLN4 expression level between high-risk group and low-risk group ($P=0.049$, Mann–Whitney U test; Figure 5D). Moreover, Figure 5D also revealed that the number of somatic mutations in high-risk group were slightly higher than those in low-risk group ($P=0.09$, Mann–Whitney U test; Figure 5D).

3.Validation of GILncSig in the testing set and entire TCGA set

To confirm our findings, the prognostic performance of the GILncSig was further evaluated in the testing set. Patients in the testing set were divided into the high-risk group ($n = 43$) and the low-risk group ($n = 44$) by using the same GILncSig and cutoff value deriving from the training set. Kaplan-Meier curves showed that there was a significant difference in overall survival between the high-risk group and the low-risk group, and the overall survival of the high-risk group was much lower than the low-risk group ($p<0.001$, log-rank test, Figure 5E), which were similar to those observed in the training set. Validation of the GILncSig in the testing set of 87 patients produced an ROC with an AUC of 0.806 at 1 year (Figure 5G). Figure 5F shows how the expression level of GILncSig, the count of somatic mutation, and the expression level of UBQLN4 in the testing set change with the increasing score. The analysis indicated that Somatic mutation counts and the expression level of UBQLN4 were significantly higher in the high-risk group as compared with those in the low-risk group ($P=0.0044$ – $P=0.00054$, Mann-Whitney U test; Figure 5H).

Similar results were observed when the prognostic performance of the GILncSig was further used to the entire TCGA set. Like the training and testing set, the GILncRNA was able to stratify 171 pancreatic cancer patients of the entire TCGA set into the high-risk group ($n=81$) and low-risk group ($n=90$) with obviously different overall survival ($P<0.001$, log-rank test, Figure 6A). The AUC of time-dependent ROC analysis for overall survival in the entire TCGA set was 0.724 (Figure 6B). The expression of GILncSig, somatic mutation counts and UBQLN4 expression level of pancreatic cancer patients in TCGA set were presented in Figure 6C, which were similar to those observed in the training set and testing set. The counts of somatic mutations in high-risk group were significantly higher than that in low-risk group ($P=0.0022$, Mann-Whitney U test, Figure 6D), as was the expression level of UBQLN4 ($P=0.0001$, Mann-Whitney U test, Figure 6D).

4.Comparison of the GILncSig and other lncRNA-related predictive signatures for survival prediction

Recently, two lncRNA-related signatures were reported to predict prognosis of pancreatic cancer patients. Therefore, we further compared the prognostic value of our GILncSig to that of different lncRNA-associated signatures for predicting outcomes: the five-lncRNA signature derived from Song's study (hereinafter referred to as SongSig)[26] and the three-lncRNA signature derived from Shi's study (hereinafter referred to as ShiSig)[27]. Utilizing the same TCGA patient set. Then we performed the time-dependent ROC analysis and calculated the area under the ROC curves to compare the prediction performance between the GILncSig and other two existing lncRNA-related

signatures in the entire TCGA set. The result demonstrated that the AUC at 1 year of overall survival for the GILncSig is 0.724, which was significantly higher than that of SongSig (AUC = 0.642) and ShiSig (AUC = 0.556) (Figure 7). For this reason, we believed that the GILncSig had better prognostic power than those two lncRNA-related signatures.

5. Independence of prognostic value of the GILncRNA from other clinical variables

To determine whether the prognostic value of the GILncRNA was independent of other clinical variables. Multivariate Cox regression analysis was performed in each patient set using prognostic risk score, age, gender, pathological grade and stage. Results from multivariate Cox analysis revealed that the GILncRNA was significantly associated with overall survival in each set when adjusted for age, gender, pathological grade and stage (Table 1). At the same time, we also observed that age, gender, pathological grade and stage were different in the multivariate analysis significantly. So we further performed data stratification analysis according to age and gender, pathological grade and stage. According to the age, pancreatic cancer patients could be stratified into an old patient group (age > 65, n=81) and a young patient group (age ≤65, n=90). The GILncRNA could subdivide each age group into high-risk group and low-risk group. There was significantly different overall survival between high-risk group and low-risk group in each age group. (log-rank test $p=0.016$ for the old patient group and log-rank test $p<0.001$ for the young patient group) (Figure 8A). Next, all patients were also stratified by gender. 78 female patients were classified into high-risk group (n =37) and low-risk group (n = 41) based on the GILncRNA. Similarly, 93 male patients were separated into two groups (high-risk group: n=44; low-risk group: n=49). The overall survival of patients in the low-risk group was significantly longer than that of patients in the high-risk group by analysis of the results. (log-rank test $p=0.002$ for the female group; log-rank test $p=0.001$ for the male group; Figure 8B). In addition, all patients in entire TCGA set were grouped according to tumor size, lymph node metastasis and distant metastasis. Each group was further separated into high-risk group and low-risk group by the GILncRNA, and the difference of overall survival between the two groups was compared. As shown in Figure 8, with the exception of the metastatic group (M1 group), there were statistically significant differences in overall survival between the high-risk and low-risk groups in each group, whereas marginally significant difference was existed in high-risk subgroup and low-risk subgroup of M1 group ($p=0.052$ for T1-2 group, $p<0.001$ for T3-4 group, Figure 8D; $p=0.027$ for N0 group, $p=0.03$ for N1 group, Figure 8G; $p=0.009$ for M0 group, $p=0.317$ for M1 group, Figure 8C; log-rank test). Finally, the same analysis method was applied to the pathological grade and stage of patients. All patients were divided into low-grade (G1/G2) and high-grade groups (G3/G4) according to pathological grade. The results of stratified analysis showed that the patients with high-grade were divided into either a high-risk group (n=24) with shorter survival or a low-risk group (n=25) with longer survival ($p=0.068$, log-rank test; Figure 8E). The patients in the low-grade group were similarly classified into two risk subgroups with significantly different survival time ($p<0.001$, log-rank test; Figure 8E). Furthermore, patients with pathologic stage I or II were combined into an early-stage group (n=161) and that with pathologic stage III or IV were combined into a late-stage group (n=7). The GILncRNA divided the early-stage group and the late-stage group into high-risk group and low-risk group respectively. The overall survival was significantly different between the two groups among the early-stage group ($p<0.001$, log-rank test; Figure 8F). Nevertheless, the difference in overall survival between the two groups was not significant probably due to the limited sample size in the late group ($p=0.549$, log-rank test; Figure 8F). Taken together, these results indicated that the GILncSig was an independent prognostic factor associated with overall survival in pancreatic cancer patients.

6. The prognostic significance of GILncSig better than KRAS, TP53, SMAD4 mutation status

KRAS, TP53 and SMAD4 were the most frequent mutant genes and associated with poor prognosis in pancreatic cancer. With this in mind, these three genes were included in the training set, testing set and TCGA set for analysis,

respectively. Then further stratified analysis was performed based on the mutation status of KRAS, TP53 and SMAD4 by GILncSig. The analysis showed that the proportion of patients with KRAS, TP53 and SMAD4 mutations in high-risk group was higher than that in low-risk group to varying degrees in each set. For KRAS, 66% of the high-risk group had KRAS mutations, significantly higher than 16% of the low-risk group in the training set (chi-square test $P < 0.001$). In the testing set, 72% of the high-risk group had KRAS mutation, which was significantly higher than 46% of the low-risk group (chi-square test $P = 0.040$). In the entire TCGA set, 69% of KRAS mutation in high-risk group significantly higher than 33% of low-risk group (chi square test $P < 0.001$). These results suggest that GILncSig is closely related to the mutation state of KRAS gene. Therefore, we applied GILncSig to patients with KRAS Wild type (KRAS Wild) and KRAS mutation type (KRAS mutation). Patients with KRAS Wild were divided into low-risk group (KRAS Wild/GS-like) and high-risk group (KRAS Wild/ GU-like), and patients with KRAS mutation were divided into low-risk group (KRAS Wild/GS-like) and high-risk group (KRAS mutation/GU-like). Through comparative analysis, we found that overall survival of KRAS Wild/GS-like group was significantly different from that of KRAS Wild/ GU-like group and KRAS Wild/ GU-like group and patients in KRAS Wild/GS-like group had better prognosis ($p = 0.01$, log-rank test; Figure 9A). For TP53, as shown in Figure 9B, 73% of TP53 mutations in the high-risk group were significantly higher than 26% in the low-risk group in the training set (chi square test $P < 0.001$). Similarly, in the TCGA set, the TP53 mutation in high-risk group was obviously higher than that in low-risk group (high-risk group 66% versus low-risk group 41%, chi square test $P = 0.004$). However, TP53 mutations were only slightly higher in the high-risk group than that in the low-risk group in the test set, and there was no significant difference between the two groups (high-risk group 58% versus low-risk group 54%, chi square test $P = 0.874$). In consequence, we believe that TP53 status can be predicted according to the GILncSig risk score. Then patients with TP53 mutation and TP53 wild type were further divided into TP53 mutation high-risk group (TP53 mutation/GU-like), TP53 mutation low-risk group (TP53 mutation/GS-like), TP53 wild high-risk group (TP53 wild/GU-like) and TP53 wild low-risk group (TP53 wild/GS-like). Survival analysis showed that patients in the TP53 wild/GS-like group had longer survival than those in the TP53 wild/GU-like group, and the higher risk scores were associated with lower survival rates in TP53 wild subgroups ($p = 0.002$, log-rank test; Figure 9B). For SMAD4, it has similar results with KRAS and TP53. The patients in training set, testing set and TCGA set were respectively divided into high-risk group and low-risk group by using GILncSig. In each set, the proportion of SMAD4 mutation in high-risk group was significantly higher than that in low-risk group ($p = 0.228$ for training set; $p = 0.028$ for testing set; $p = 0.009$ for TCGA set; chi-square test; Figure 9C). The patients with SMAD4 mutation type and SMAD4 wild type were further separated into SMAD4 mutation/GU-like group, SMAD4 mutation/GS-like, SMAD4 wild/GU-like group and SMAD4 wild/GS-like group. The results of survival analysis showed that the overall survival among the groups was slightly different ($p = 0.062$, log-rank test; Figure 9C). Therefore, the above findings suggested that the GILncSig is superior to KRAS, TP53 and SMAD4 mutation status in prognosis.

7. Development and validation of a nomogram for predicting survival in patients with pancreatic cancer

In order to improve the clinical application of the GILncSig, we established a prognostic nomogram model combined with risk score, age, gender, pathological grade and stage to predict the patients' survival at 1-, 2-, and 3- years in the training set by using "rms" and "survival" packages in software R (Figure 10A). In Figure 10B, the C-index of the nomogram of the training set was 0.650, and the AUC values predicted for 1-, 2- and 3-year survival is 0.806, 0.844 and 0.792, respectively. The C-index was 0.615 in the testing set and the AUCs of ROC for 1-, 2-, and 3-year survival predictions were 0.653, 0.776, and 0.856, respectively (Figure 10C). Likewise, the C-index was 0.618 in the whole TCGA set and the 1-, 2-, and 3-year AUCs were 0.724, 0.814, and 0.83, respectively (Figure 10D). The calibration plots in (Supplementary Figure 1) exhibited excellent accordance between the nomogram prediction and the actual values

in terms of the 1-, 2- and 3-year survival rates in the three datasets. The above results indicated that the prediction performance of the established nomogram is improved.

Correlation of Risk Score with TIME (Tumor Immune Environment Characterization)

Through the ESTIMATE evaluation method, TumorPurity, ImmuneScore and StromalScore were calculated. These results indicated that patients in the low-risk group have lower TumorPurity and higher ImmuneScore and StromalScore (Figure 11A,B,C). To further uncover the correlation between GILncSig and immune cell infiltration, the analysis showed that patients in low-risk group had more T cells CD8, B cells and T cells CD4 memory activated, while the Macrophages M0 was at low level (Figure 11 D). To further explore the influence of GILncSig upon TIME of pancreatic cancer we analyzed correlation of risk signature with immune cell infiltration type and level. The results indicated that the risk signature significantly correlated with infiltrating B cells ($r = -0.38$; $p = 1.2e - 05$), infiltrating CD4+T cells ($r = -0.34$; $p = 0.00011$), infiltrating plasma cells ($r = -0.19$; $p = 0.036$), CD8 T cells ($r = -0.35$; $p = 6.6e - 05$), macrophages M0 ($r = 0.32$; $p = 0.00023$), and macrophages M2 ($r = 0.21$; $p = 0.018$; Figure 11 E). Then, the ssGSEA algorithm was used to examine whether there was distinction of immune signatures between groups low/high risk. From the results, found that the infiltrating levels of B cells, CD8+T cells, DCs, Neutrophils, pDCs, Tfh, Th1 cells, Th2 cells were remarkably elevated and some immune signatures (i.e. CCR, checkpoint, inflammation promoting, IFN response type II) were significantly activated in the low-risk group (Figure 12 A,B).

Correlation of Risk Score with Immune Checkpoint Blockade Key Molecules

Six key immune checkpoint inhibitor genes (PDCD1, CD274, PDCD1LG2, CTLA-4, HAVCR2, and IDO1) were singled out for further research. We performed the correlation analysis of ICB key gene expression with risk signature to investigate the potential role of signature in the ICB therapy of pancreatic cancer (Figure 12 C). Correlation analysis results indicated that GILncSig had close relationship with CD274 ($r = -0.18$; $p = 0.018$), CTLA4 ($r = -0.32$; $p = 2e - 05$), HAVCR2 ($r = -0.29$; $p = 0.00012$), PDCD1 ($r = -0.4$; $p = 8.1e - 08$), and PDCD1LG2 ($r = -0.3$; $p = 6e - 05$; Figures 12D), indicating GILncSig might exert a nonnegligible player in ICB treatment outcome prediction in pancreatic cancer. Further correlation analysis presented that 32 of 47 (i.e. CD27, IDO2, etc.) immune check blockade-related genes expression levels were significantly different between two risk groups (Figures 13).

Discussion

Pancreatic cancer is the most common cause of cancer death worldwide[28]. It is characterized by high morbidity, high mortality, difficult early diagnosis and poor prognosis[29, 30]. Surgical resection is effective for patients with early pancreatic cancer, while palliative treatment is adopted for patients with locally advanced, metastatic and unresectable pancreatic cancer[31]. In recent years, the molecular research of pancreatic cancer has made great progress, and the survival rate of pancreatic cancer patients has been improved to some extent. However, the prognosis has not been improved[32]. As metastasis and recurrence are the main causes of poor prognosis, it is urgent to identify effective tumor biomarkers to evaluate the prognosis of patients with pancreatic cancer accurately.

Genomic instability is an important feature of human cancer, which is associated with poor prognosis, metastasis[4, 33]. It has been reported that genomic instability affects the prognosis of pancreatic cancer, and the pattern of genomic instability is quite heterogeneous in metastatic pancreatic cancer[34, 35]. It is known that the degree of genomic instability has diagnostic and prognostic implications, yet measuring genomic instability is a big challenge. Metturk, et al. constructed a 12-gene signature to assess genomic instability and predict clinical

outcomes in cancers[36]. Zhang S, et al. developed a biological rationale-driven genomic instability score to predict the prognosis of ovarian cancer[37].

LncRNAs have complex biological functions and have been proved to be closely related to the occurrence and development of cancers[11, 38]. Recently, increasingly more researchers pay attention to the clinical significance of LncRNAs in the prognosis of cancers. For instance, the high expression of lncRNA HOX transcript antisense RNA (HOTAIR) in lung tumor tissues is correlated with metastasis and poor prognosis in patients with lung cancer[39]. The lncRNA AOC4P induces a poor prognosis in gastric cancer patients through epithelial-mesenchymal transition[40]. It has been found that lncRNAs play an important role in maintaining genomic stability through continuous exploration of the function of lncRNAs[22, 23, 41]. Although some efforts have been made, few researches have been done on genomic instability-related lncRNAs in cancers. Therefore, there is an urgent need to investigate the prognostic value of genomic-instability associated lncRNAs in pancreatic cancer patients.

In our study, we identified 40 genomic instability-associated lncRNAs by analyzing the lncRNA expression profile and somatic mutation profile of 171 patients with pancreatic cancer. Then, the function of these lncRNAs was predicted by the lncRNA-mRNA co-expression network. The GO and KEGG enrichment results suggested that the genes co-expressed with these 206 lncRNAs were enriched at chromosomes and nucleoplasm in the cellular component, DNA binding in the molecular function and the transcription and compound synthesis and metabolism in the biological process can promote genomic instability, which lead to cancer eventually[42, 43]. We further divided all patients into training set and testing set. Cox proportional risk regression analysis was performed on the candidate genomic instability-associated lncRNAs in the training set, and a genomic instability-associated lncRNAs signature (GILncSig) consisting of 5 lncRNAs with independent prognostic value (AL121772.1, BX640514.2, LINC01133, AC087752.3 and LYPLAL1-AS1) was established to predict the prognosis of pancreatic cancer. The GILncSig can classify pancreatic cancer patients in the training set into the high-risk group and low-risk group with significantly different overall survival, which was verified in the testing set and the whole TCGA set. In addition, we also found that patients with pancreatic cancer in the high-risk group had significantly higher somatic mutation counts and UBQLN4 expression levels, both of which are characteristics of genomic instability. Comparison of our GILncSig and two recently reported lncRNA-related signatures with predictive value for pancreatic cancer in the same TCGA patient set suggested that the GILncSig has better prognostic ability in predicting survival than those two lncRNA-related signatures. Our study also found that the GILncSig was independent of other clinicopathological factors, including age, gender, pathological grade and stage. Furthermore, based on the GILncSig, the mutation states of KRAS, TP53 and SMAD4 in high-risk group were significantly higher than those in low-risk group. The survival time of KRAS, TP53 and SMAD4 wild-type patients in low-risk group was significantly longer than that of patients with mutant-type.

The above results indicated that the GILncSig may have greater prognostic significance than KRAS, TP53 and SMAD4 mutation states. Finally, a nomogram was constructed by combining GILncSig and the four independent prognostic factors of age, gender, pathological grade and stage in the training set, which further improved the predictive performance, and was verified on the testing set and the entire TCGA set.

What's more, numerous researches focusing on TIME have revealed the potential key role of lncRNAs on infiltrating immune cells. In this study, we find that GILncSig was significantly correlated with immune cell infiltration, ESTIMATE results showed that GILncSig was positively with tumor purity but negatively correlated with estimate score and immune score, suggesting GILncSig could serve as a novel immune indicator in pancreatic cancer. Besides, ssGSEA results indicated that in the low-risk group the infiltrating immune cells were significantly increased

and immune signatures were remarkably activated. The immune-activated condition in the low-risk group was associated with high ICB-relevant genes expression, suggesting samples in with high risk score might respond to immunotherapy. What'more, the correlation analysis between ICB-related genes and GILncSig indicated that our signature may possess the ability to predict clinical outcome of ICB therapy in pancreatic cancer.

Although the GILncSig identified here is reliable and promising as a prognostic signature in tumor immune microenvironment of pancreatic cancer, there are still several limitations. In addition to validation in the TCGA dataset, the GILncSig requires more independent datasets to verify. Meanwhile, it is necessary to further explore the regulatory mechanism of GILncSig in biological function to maintain genomic instability.

Conclusion

In summary, we have performed RNA-seq prognostic analysis in pancreatic cancer patients by bioinformatics methods to develop a genomic instability-derived lncRNA signature to predict the prognosis of pancreatic cancer patients and successfully validated it on the independent cohort. Moreover, we integrated GILncSig with age, gender, pathological grade and stage to construct a nomogram to improve its prediction performance. And further results unraveled that GILncSig was significantly correlated with immune cell infiltration and have important significance for genomic instability and ICB treatment of pancreatic cancer.

Declarations

Availability of data and material

Publicly available datasets were analyzed in this study. The data can be found:<https://portal.gdc.cancer.gov/>,<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi>.

Competing interests

The authors declare no conflicts of interest

Funding

This work was supported by grants from the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD, JX10231801) and The Innovation Capability Development Project of Jiangsu Province (No.BM2015004) and Jiangsu Biobank of Clinical Resources(BM2015004).

Author contributions

QL, YM, CQH, CYS, XLZ, RY, LDY and YPP all have made substantial contributions to conception, acquisition of data, analysis, and interpretation of data. All of them have been involved in drafting the manuscript and revising it critically for important intellectual content. All authors read and approved the final manuscript and take public responsibility for appropriate portions of the content and agreed to be accountable for all aspects of work.

Acknowledgements

Not applicable.

References

1. A. McGuigan, P. Kelly, R.C. Turkington, C. Jones, H.G. Coleman, R.S. McCain, Pancreatic cancer: A review of clinical diagnosis, epidemiology, treatment and outcomes, *World Journal of Gastroenterology*, 24 (2018) 4846-4861.
2. I. Garrido-Laguna, M. Hidalgo, Pancreatic cancer: from state-of-the-art treatments to promising novel therapies, *Nat Rev Clin Oncol*, 12 (2015) 319-334.
3. G. Garcea, C.P. Neal, C.J. Pattenden, W.P. Steward, D.P. Berry, Molecular prognostic markers in pancreatic cancer: a systematic review, *Eur J Cancer*, 41 (2005) 2213-2236.
4. P.H.G. Duijf, D. Nanayakkara, K. Nones, S. Srihari, M. Kalimutho, K.K. Khanna, Mechanisms of Genomic Instability in Breast Cancer, *Trends in Molecular Medicine*, 25 (2019) 595-611.
5. R. Gupta, S. Sinha, R.N. Paul, The impact of microsatellite stability status in colorectal cancer, *Curr Probl Cancer*, 42 (2018) 548-559.
6. L. Ma, V.B. Bajic, Z. Zhang, On the classification of long non-coding RNAs, *RNA Biology*, 10 (2014) 924-933.
7. J.L. Rinn, H.Y. Chang, Genome Regulation by Long Noncoding RNAs, *Annual Review of Biochemistry*, 81 (2012) 145-166.
8. T.R. Mercer, M.E. Dinger, J.S. Mattick, Long non-coding RNAs: insights into functions, *Nat Rev Genet*, 10 (2009) 155-159.
9. R. Spizzo, M.I. Almeida, A. Colombatti, G.A. Calin, Long non-coding RNAs and cancer: a new frontier of translational research?, *Oncogene*, 31 (2012) 4577-4587.
10. S.W. Cheetham, F. Gruhl, J.S. Mattick, M.E. Dinger, Long noncoding RNAs and the genetics of cancer, *Br J Cancer*, 108 (2013) 2419-2425.
11. E.A. Gibb, C.J. Brown, W.L. Lam, The functional role of long non-coding RNA in human carcinomas, *Mol Cancer*, 10 (2011) 38.
12. J.Z. Wang, J.J. Xiang, L.G. Wu, Y.S. Bai, Z.W. Chen, X.Q. Yin, Q. Wang, W.H. Guo, Y. Peng, H. Guo, P. Xu, A genetic variant in long non-coding RNA MALAT1 associated with survival outcome among patients with advanced lung adenocarcinoma: a survival cohort analysis, *BMC Cancer*, 17 (2017) 167.
13. G. Troiano, V.C.A. Caponio, L. Boldrup, X. Gu, L.L. Muzio, N. Sgaramella, L. Wang, K. Nylander, Expression of the long non-coding RNA HOTAIR as a prognostic factor in squamous cell carcinoma of the head and neck: a systematic review and meta-analysis, *Oncotarget*, 8 (2017) 73029-73036.
14. Y. Zhang, T. Shields, T. Crenshaw, Y. Hao, T. Moulton, B. Tycko, Imprinting of human H19: allele-specific CpG methylation, loss of the active allele in Wilms tumor, and potential for somatic allele switching, *American journal of human genetics*, 53 (1993) 113-124.
15. C. Braconi, T. Kogure, N. Valeri, N. Huang, G. Nuovo, S. Costinean, M. Negrini, E. Miotto, C.M. Croce, T. Patel, microRNA-29 can regulate expression of the long non-coding RNA gene MEG3 in hepatocellular cancer, *Oncogene*, 30 (2011) 4750-4756.
16. N. Hauptman, D. Glavac, Long non-coding RNA in cancer, *Int J Mol Sci*, 14 (2013) 4655-4669.
17. T. Lin, Y. Fu, X. Zhang, J. Gu, X. Ma, R. Miao, X. Xiang, W. Niu, K. Qu, C. Liu, Q. Wu, A seven-long noncoding RNA signature predicts overall survival for patients with early stage non-small cell lung cancer, *Aging*, 10 (2018) 2356-2366.
18. P. Diao, Y. Song, H. Ge, Y. Wu, J. Li, W. Zhang, Y. Wang, J. Cheng, Identification of 4-lncRNA prognostic signature in head and neck squamous cell carcinoma, *J Cell Biochem*, 120 (2019) 10010-10020.

19. M. Zhou, Y. Sun, Y. Sun, W. Xu, Z. Zhang, H. Zhao, Z. Zhong, J. Sun, Comprehensive analysis of lncRNA expression profiles reveals a novel lncRNA signature to discriminate nonequivalent outcomes in patients with ovarian cancer, *Oncotarget*, 7 (2016) 32433-32448.
20. J. Li, W. Wang, P. Xia, L. Wan, L. Zhang, L. Yu, L. Wang, X. Chen, Y. Xiao, C. Xu, Identification of a five-lncRNA signature for predicting the risk of tumor recurrence in patients with breast cancer, *Int J Cancer*, 143 (2018) 2150-2160.
21. J. Tang, J. Ren, Q. Cui, D. Zhang, D. Kong, X. Liao, M. Lu, Y. Gong, G. Wu, A prognostic 10-lncRNA expression signature for predicting the risk of tumour recurrence in breast cancer patients, *J Cell Mol Med*, 23 (2019) 6775-6784.
22. S. Lee, F. Kopp, T.C. Chang, A. Sataluri, B. Chen, S. Sivakumar, H. Yu, Y. Xie, J.T. Mendell, Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins, *Cell*, 164 (2016) 69-80.
23. W.L. Hu, L. Jin, A. Xu, Y.F. Wang, R.F. Thorne, X.D. Zhang, M. Wu, GUARDIN is a p53-responsive long non-coding RNA that is essential for genomic stability, *Nat Cell Biol*, 20 (2018) 492-502.
24. J. Li, L. Han, P. Roebuck, L. Diao, L. Liu, Y. Yuan, J.N. Weinstein, H. Liang, TANRIC: An Interactive Open Platform to Explore the Function of lncRNAs in Cancer, *Cancer Res*, 75 (2015) 3728-3737.
25. G. Yu, L.-G. Wang, Y. Han, Q.-Y. He, clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters, *OMICS: A Journal of Integrative Biology*, 16 (2012) 284-287.
26. J. Song, Q. Xu, H. Zhang, X. Yin, C. Zhu, K. Zhao, J. Zhu, Five key lncRNAs considered as prognostic targets for predicting pancreatic ductal adenocarcinoma, *J Cell Biochem*, 119 (2018) 4559-4569.
27. X. Shi, Y. Zhao, R. He, M. Zhou, S. Pan, S. Yu, Y. Xie, X. Li, M. Wang, X. Guo, R. Qin, Three-lncRNA signature is a potential prognostic biomarker for pancreatic adenocarcinoma, *Oncotarget*, 9 (2018) 24248-24259.
28. R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2020, *CA Cancer J Clin*, 70 (2020) 7-30.
29. J.-F. Peng, Noncoding RNAs and pancreatic cancer, *World Journal of Gastroenterology*, 22 (2016).
30. C. Sharma, Advances in diagnosis, treatment and palliation of pancreatic carcinoma: 1990-2010, *World Journal of Gastroenterology*, 17 (2011).
31. D. Li, K. Xie, R. Wolff, J.L. Abbruzzese, Pancreatic cancer, *Lancet (London, England)*, 363 (2004) 1049-1057.
32. G. Feldmann, A. Maitra, Molecular genetics of pancreatic ductal adenocarcinomas and recent implications for translational efforts, *J Mol Diagn*, 10 (2008) 111-122.
33. S.F. Bakhoun, L.C. Cantley, The Multifaceted Role of Chromosomal Instability in Cancer and Its Microenvironment, *Cell*, 174 (2018) 1347-1360.
34. I.H. Sahin, M.A. Lowery, Z.K. Stadler, E. Salo-Mullen, C.A. Iacobuzio-Donahue, D.P. Kelsen, E.M. O'Reilly, Genomic instability in pancreatic adenocarcinoma: a new step towards precision medicine and novel therapeutic approaches, *Expert Rev Gastroenterol Hepatol*, 10 (2016) 893-905.
35. P.J. Campbell, S. Yachida, L.J. Mudie, P.J. Stephens, E.D. Pleasance, L.A. Stebbings, L.A. Morsberger, C. Latimer, S. McLaren, M.-L. Lin, D.J. McBride, I. Varela, S.A. Nik-Zainal, C. Leroy, M. Jia, A. Menzies, A.P. Butler, J.W. Teague, C.A. Griffin, J. Burton, H. Swerdlow, M.A. Quail, M.R. Stratton, C. Iacobuzio-Donahue, P.A. Futreal, The patterns and dynamics of genomic instability in metastatic pancreatic cancer, *Nature*, 467 (2010) 1109-1113.
36. R.K. Mettu, Y.W. Wan, J.K. Habermann, T. Ried, N.L. Guo, A 12-gene genomic instability signature predicts clinical outcomes in multiple cancer types, *The International journal of biological markers*, 25 (2010) 219-228.
37. S. Zhang, Y. Yuan, D. Hao, A genomic instability score in discriminating nonequivalent outcomes of BRCA1/2 mutations and in predicting outcomes of ovarian cancer treated with platinum-based chemotherapy, *PloS one*, 9

(2014) e113169.

38. A. Fatica, I. Bozzoni, Long non-coding RNAs: new players in cell differentiation and development, *Nat Rev Genet*, 15 (2014) 7-21.
39. G. Loewen, J. Jayawickramarajah, Y. Zhuo, B. Shan, Functions of lncRNA HOTAIR in lung cancer, *Journal of hematology & oncology*, 7 (2014) 90.
40. K. Zhang, C. Lu, X. Huang, J. Cui, J. Li, Y. Gao, W. Liang, Y. Liu, Y. Sun, H. Liu, B. Wei, L. Chen, Long noncoding RNA AOC4P regulates tumor cell proliferation and invasion by epithelial-mesenchymal transition in gastric cancer, *Therap Adv Gastroenterol*, 12 (2019) 1756284819827697.
41. H. Liu, Linking lncRNA to genomic stability, *Science China Life Sciences*, 59 (2016) 328-329.
42. K.J. Barnum, M.J. O'Connell, Cell cycle regulation by checkpoints, *Methods Mol Biol*, 1170 (2014) 29-40.
43. E.C. Friedberg, How nucleotide excision repair protects against cancer, *Nature reviews. Cancer*, 1 (2001) 22-33.

Tables

Table 1. Univariate and Multivariate Cox regression analysis of the GILncSig and clinical features for the independent prognostic significance in different patient datasets

Variables	Univariable model				Multivariable model			
	HR	HR.95L	HR.95H	pvalue	HR	HR.95L	HR.95H	pvalue
TCGA set								
age	1.027207	1.005871	1.048994	0.012189	1.023761	1.001864	1.046137	0.033272
gender	0.873723	0.577194	1.322592	0.523359				
grade	1.391989	1.040839	1.861608	0.025759	1.250423	0.931461	1.678608	0.136906
stage	1.365182	0.936063	1.991023	0.105923				
riskScore	1.030134	1.014894	1.045604	9.46E-05	1.02821	1.013716	1.042912	0.000123
Testing set (								
id	HR	HR.95L	HR.95H	pvalue	HR	HR.95L	HR.95H	pvalue
age	1.02934	1.000879	1.058611	0.04324	1.02934	1.000879	1.058611	0.04324
gender	1.051846	0.60187	1.838237	0.859146				
grade	1.341221	0.926544	1.941486	0.119783				
stage	1.341823	0.799856	2.251018	0.265316				
riskScore	1.029268	0.964336	1.098572	0.38557				
Training set								
id	HR	HR.95L	HR.95H	pvalue	HR	HR.95L	HR.95H	pvalue
age	1.026642	0.994416	1.059912	0.106125				
gender	0.768614	0.409857	1.4414	0.412039				
grade	1.454163	0.904621	2.337542	0.122089				
stage	1.439642	0.82667	2.507129	0.19794				
riskScore	1.026271	1.011035	1.041736	0.000679	1.02934	1.000879	1.058611	0.04324

Figures

Image not available with this version

Figure 1

Computational framework of genomic instability-associated lncRNAs detection.

Figure 2

Identification of genome instability-related lncRNAs in patients with pancreatic cancer. (A) Unsupervised clustering analysis of the lncRNA expression profiles in GU group and GS group. (B) Unsupervised clustering analysis of 171 patients with pancreatic cancer according to the differential expression patterns of 206 genomic instability-related lncRNAs. (C) Boxplots of somatic mutations counts in the GU-like group and GS-like group. (D) Boxplot of the expression level of UBQLN4 in the GU-like group and GS-like group

Figure 4

Construction of the genomic instability-associated lncRNA prognostic signature from the training set. (A) Forest plot of 17 candidate prognostic lncRNAs associated with pancreatic cancer patients' overall survival based on univariate Cox regression analyses. (B) Forest plot of 5 candidate prognostic lncRNAs associated with pancreatic cancer patients' overall survival based on stepwise multivariate Cox proportional hazard regression.

Figure 5

Identification and validation of the genomic instability-related lncRNA signature (GILncSig) for outcome prediction in patients with pancreatic cancer in the training set and testing set. (A, E) Kaplan–Meier survival curves of patients in the high- and low-risk groups separated by the median GILncSig score in the training set (A) and testing set (E). (B, F) lncRNA expression patterns and the distribution of somatic mutation and UBQLN4 expression with increasing GILncSig score in the training set (B) and testing set (F). (C, G) Time-dependent ROC curves for 1-year survival prediction of the GILncSig in the training set (C) and testing set (G). (D, H) Boxplots of comparison of the somatic mutation counts and the UBQLN4 expression between the high- and low-risk groups in the training set (D) and testing set (H).

Figure 6

Predictive performance evaluation of the GILncSig in patients with pancreatic cancer in the TCGA set. (A) Kaplan-Meier survival curves of patients in the high- and low-risk groups separated by the median GILncSig score in the TCGA set. (B) LncRNA expression patterns and the distribution of somatic mutation and UBQLN4 expression with increasing GILncSig score in the TCGA set. (C) Time-dependent ROC curves for 1-year survival prediction of the GILncSig in the TCGA set. (D) Boxplots of comparison of the somatic mutation counts and the UBQLN4 expression between the high- and low-risk groups in the TCGA set.

Figure 7

The ROC analysis for 1-year survival prediction of the GlnLncSig and the other two existing signatures (SongLncSig, ShiLncSig), respectively.

Figure 8

Stratification analysis by age, gender and pathological grade and stage. (A) Kaplan-Meier curve analysis of overall survival in low-/high-risk groups for age. (B)Kaplan-Meier curve analysis of overall survival in low-/high-risk groups for gender. (C, D, G) Kaplan-Meier curve analysis of overall survival in low-/high-risk groups based on tumor size, lymph node metastasis and distant metastasis. (E) Kaplan-Meier curve analysis of overall survival in low-/high-risk groups for grade. (F) Kaplan-Meier curve analysis of overall survival in low-/high-risk groups for stage

Figure 9

Relationship between the GILncSig and KRAS, TP53, SMAD4 mutation. The proportion of KRAS (A), TP53 (B), SMAD4 (C) mutation in the high- and low-risk group in the training set, the testing set, and the TCGA set. Kaplan-Meier survival curves of overall survival for patients in groups divided based on KRAS (A), TP53 (B), SMAD4 (C) mutation status and the GILncSig.

Figure 10

Construction and assessment of a nomogram for survival prediction of patients with pancreatic cancer based on risk score, age, gender, pathological grade and stage. (A) The nomogram established in training set for predicting 1-, 2-, and 3-year survival of pancreatic cancer patients. (B–D) ROC curve analysis for 1-, 2-, and 3-year survival prediction of the nomogram in the training set (B), testing set (C), and TCGA set (D), respectively.

Figure 11

Correlation of prognostic risk score with TIME characterization (A-C) The correlation of estimate score, immune score, and tumor purity between these two subtypes. (D) Difference of infiltrating immune cell subpopulations and levels between low-/high-risk groups. (E) Correlation between tumor immune infiltration and GILncSig.

Figure 12

Correlation of prognostic risk score with TIME characterization (A) immune-related signature with corresponding immune-related scores in groups low/high risk. (B) Distinction of enrichment of immune-related signatures between low-/high-risk groups. (C) Association analyses between immune checkpoint inhibitors and GILncSig. (D) Association between risk model and CD274, CTLA4, HAVCR2, PDCD1, PDCD1LG2.

RISK ■ low ■ high

Figure 13

Comparison of 32 immune checkpoint blockade-related genes expression levels in low-/high-risk groups

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFigure1.png](#)