Comprehensive analysis of prognostic immune-related genes associated with the tumor microenvironment of pancreatic ductal adenocarcinoma

SHIBAI YAN¹, JUNTAO FANG², YUANQIANG ZHU³, YONG XIE⁴ and FENG FANG⁴

¹Department of Medical Oncology, The First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi Zhuang Autonomous Region 530021, P.R. China; ²Laboratory of Experimental Cardiology, Department of Cardiology, University Medical Center Utrecht, 3584 CX Utrecht, The Netherlands; ³Department of Infertility and Sexual Medicine, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong 510630; ⁴Department of Obstetrics and Gynecology, The First People's Hospital of Foshan, Foshan, Guangdong 528000, P.R. China

Received February 13, 2020; Accepted September 4, 2020
DOI: 10.3892/ol.2020.12228

Correspondence to: Mr. Feng Fang, Department of Obstetrics and Gynecology, The First People's Hospital of Foshan, 81 Lingnan North Avenue, Foshan, Guangdong 528000, P.R. China
E-mail: fangfeng_go@163.com

Key words: pancreatic ductal adenocarcinoma, immune-related genes, tumor microenvironment, prognosis

Abstract. Pancreatic ductal adenocarcinoma (PDAC) is a malignant tumor with a specific tumor immune microenvironment (TIME). Therefore, investigating prognostic immune-related genes (IRGs) that are closely associated with TIME to predict PDAC clinical outcomes is necessary. In the present study, 459 samples of PDAC from the Genotype-Tissue Expression database, The Cancer Genome Atlas (TCGA), International Cancer Genome Consortium (ICGC) and Gene Expression Omnibus (GEO) were included and a survival-associated module was identified using weighted gene co-expression network analysis. Based on the Cox regression analysis and least absolute shrinkage and selection operator analysis, four IRGs (2′-5′-oligoadenylate synthetase 1, MET proto-oncogene, receptor tyrosine kinase, interleukin 1 receptor type 2 and interleukin 20 receptor subunit β) were included in the prognostic model to calculate the risk score (RS), and patients with PDAC were divided into high- and low-RS groups. Kaplan-Meier survival and receiver operating characteristic curve analyses demonstrated that the low-RS group had significantly improved survival conditions compared with the high-RS group in TCGA training set. The prognostic function of the model was also validated using ICGC and GEO cohorts. To investigate the mechanism of different overall survival between the high- and low-RS groups, the present study included Estimation of Stromal and Immune Cells in Malignant Tumor Tissues Using Expression Data and Cell Type Identification by Estimating Relative Subset of Known RNA Transcripts algorithms to investigate the state of the tumor microenvironment and immune infiltration inpatients in the cohort from TCGA. In summary, four genes associated with the TIME of PDAC were identified, which may provide a reference for clinical treatment.

Introduction

In 2010, it has been reported that pancreatic ductal adenocarcinoma (PDAC) is a gastrointestinal malignancy with a 5-year survival rate <5% in the United States, giving patients with pancreatic cancer the poorest prognosis among patients with malignant cancer types (1). PDAC is characterized by rapid progression and metastasis (2), and poses a great threat to human health. At present, TNM staging (3) is regarded as a convincing cancer staging system to guide clinical treatment and offer a method for predicting the prognosis of patients with cancer (4). However, the clinical disadvantages of TNM staging are increasingly obvious when further investigations are performed (5). Certain clinicopathological features, such as TNM stage, lymph node involvement and differentiation, have been demonstrated to be independent prognostic factors in patients with pancreatic cancer (6). Nonetheless, the majority of these prognostic markers are verified after surgery. Therefore, it is crucial to identify useful predictive factors for PDAC before surgery. With the development of human gene sequencing technology, gene-based biomarkers have been markedly improved, which provides an opportunity to investigate effective biomarkers for guiding diagnosis, treatment and evaluation of pancreatic cancer prognosis (7).

Emerging evidence has elucidated that local immune suppression of the tumor microenvironment (TME) exhibits a strong association with cancer growth, metastasis and even tumor immune escape (8-10). Immune cells do not merely act as tumor killers, but can also function as an activator of tumors. They are capable of disturbing molecular signals and exert vital roles in cancer biology, including those associated with tumor growth, invasion and metastasis (11). However, certain types of cancer cells are able to avoid being detected and escape immune...
attack in order to promote tumor growth (12). At present, although targeted therapies and immunotherapies are effective for numerous solid malignancies, such as metastatic melanoma and lung cancer, few clinical benefits can be observed in pancreatic cancer (13, 14). Therefore, the latent prognostic value of immune genes can be illustrated through further investigation of the association between immune genes and survival. In addition, this may aid the identification of novel biomarkers. Ongoing studies that intend to target the stroma and the immune microenvironment either alone or in combination are expected to present more sustained treatment outcomes (15, 16). A detailed description of the TME integrating tumor and host-related factors may result in the verification of novel biomarkers for a more targeted method for immunotherapy, and other combination therapies to fix the immunosuppressive mechanisms in the microenvironment of pancreatic cancer.

Immune scores and analyses derived from the Estimation of Stromal and Immune Cells in Malignant Tumor Tissues Using Expression Data (ESTIMATE) algorithm can contribute to the quantification of the immune and stromal components in a tumor (17), and the Cell Type Identification by Estimating Relative Subset of Known RNA Transcripts (CIBERSORT) algorithm is an approach for describing immune cell components of tissues from their immune gene expression profiles (18). Both can be used to predict the immune microenvironment of pancreatic cancer. Considering the association between immunity and cancer, the present study was designed to identify immune-gene biomarkers for the prediction of the prognosis of PDAC, and to examine the differences in the immune microenvironment between high- and low-risk patients with pancreatic cancer.

Materials and methods

Acquisition of gene expression profiles. Gene expression profiles for use in the present study were downloaded from four databases: Genotype-Tissue Expression (19) via UCSC Xena (GTEx; https://toil.xenahubs.net/download/gtex_RSEM_gene_fpkm.gz), The Cancer Genome Atlas (20) (TCGA; https://portal.gdc.cancer.gov/repository), International Cancer Genome Consortium (21) (IGCC; https://dcc.icgc.org/releases/current/Projects/PACA-AU) and Gene Expression Omnibus (22) (GEO; http://www.ncbi.nlm.nih.gov/geo). The fragments per kilobase of transcript per million mapped reads (FPKM) data of normal pancreatic tissues were downloaded from the GTEx-Pancreas database (Full metadata), and the RNA-sequencing FPKM data of pancreatic tumor samples from the publicly available TCGA-PAAD database. Additionally, normalized high-throughput sequencing and microarray data of mRNAs (21, 22) were downloaded from the ICGC-PACA-AU database and GEO (GSE62452). Cases with complete clinical information in the pancreatic cancer cohort were included. Furthermore, patients with an overall survival ≤60 days were excluded, as these patients may have succumbed to factors not associated with the tumor, such as hemorrhage and severe infection (23). As samples from GTEx obtained from normal pancreatic tissue samples, there is no disease-related information. Eventually, a total of 459 samples obtained from GTEx (n=156), TCGA (Table SI; n=158), ICGC (Table SII; n=79) and GEO (Table SIII; n=66) databases were extracted for further analysis.

Weighted gene co-expression network analysis (WGCNA) of immune-related genes (IRGs). The limma package (version 3.44.3; http://www.sthda.com/english/wiki/survminer-r-package-survival-data-analysis-and-visualization) was used to analyze the differentially expressed mRNAs in samples between the GTEx and TCGA databases, and the gene set was intersected with 2,214 IRGs, such as those for cytokines and cytokine receptors, downloaded from the ImmPort database (https://immport.niaid.nih.gov) (24). At the same time, PDAC samples from TCGA were obtained for WGCNA to obtain gene modules significantly associated with the survival time and status of the patient. Furthermore, genes with potential prognostic value in the module were extracted via univariate Cox regression analysis (P<0.01).

Construction of prognostic prediction model. In order to screen IRGs with the most predictive efficacy in univariate Cox regression analysis (survival package, version 2.41-3; https://cran.r-project.org/web/packages/survival/index.html), a least absolute shrinkage and selection operator (LASSO) analysis (glmnet package, version 3.0-1; https://cran.r-project.org/web/packages/glmnet/index.html) was adopted to deal with over-fitting of the model. Subsequently, a prognostic model was constructed using the IRGs obtained from the multivariate regression analysis. The risk score (RS) of each patient was calculated by multiplying the expression level of each IRG with its corresponding regression coefficients. The following computational algorithm was used for this analysis (25-27): RS=|βgene(1) x exprgene(1) + βgene(2) x exprgene(2) + ... + βgene(n) x exprgene(n)|. |β| represents the regression coefficient generated from univariate Cox analysis of IRGs, and ‘exprgene’ refers to the expression of IRGs in the sample. The optimal cut-off point of risk value that was most relevant to survival was determined using the surv_cutpoint function in the survminer package (version 0.4.3; http://www.sthda.com/english/wiki/survminer-r-package-survival-data-analysis-and-visualization), which was the standard to separate PDAC samples into high-and low-RS subgroups.

Comparison of the immunological microenvironment between high- and low-RS groups. The R package ESTIMATE (version 2.0.0; https://bioinformatics.mdanderson.org/estimate/package.html) was used to calculate the tumor purity, infiltrating stromal cells and immune cell levels that can reflect the overall situation of TME in the TCGA pancreatic cancer cohort (11). Additionally, the R package GSEABase (http://www.bioconductor.org/packages/release/bioc/html/GSEABase.html) and GSVA (http://www.bioconductor.org/packages/release/bioc/html/GSVA.html) were used to calculate the enrichment of the high-RS group and the low-RS group in 29 immune function-associated gene sets. The ssGSEA score (28) can mirror the standard of immune-associated cells, functions and signaling pathways in each TCGA pancreatic cancer sample in order to elucidate the differences between high- and low-RS groups in terms of the TME.

Comparison of immune cell subtypes between high- and low-RS groups. CIBERSORT is a bioinformatics method used to evaluate immune cell composition by transformation of standardized gene expression (29). Based on the CIBERSORT
package (version 1.03; http://cibersort.stanford.edu/), the mRNA expression matrix of patients from TCGA, corrected by the zoom function in the limma package, was transformed into 22 subtypes of immune cells, which quantify the cellular composition of the immune response (18). After removing samples with P>0.05 from the CIBERSORT analysis results, a Mann-Whitney U test was performed to compare the differences in immune cell subtypes between high- and low-RS groups. The association between immune cells and modeling IGRs was also analyzed.

Enrichment analysis of upregulated genes in the high-RS group. The R package ClusterProfiler (version 3.16.1; http://www.bioconductor.org/packages/release/bioc/html/clusterProfiler.html) was used to analyze the differentially expressed genes in high- and low-RS groups. The association between immune cells and modeling IGRs was also analyzed.

Identification of survival-associated modules using WGCNA. A total of 554 IRGs were identified to be differentially expressed between normal pancreatic tissues and pancreatic tumors (Fig. S1A and B). In the WGCNA of TCGA cohort, two important parameters, R² (R²>0.9) and average connectivity were fully considered, hence β=4 was selected as the soft-thresholding power (Fig. 1A). According to the weighted

Statistical analysis. The Shapiro-Wilk normality test was used to measure the normality of the variables between two groups. The statistical significance of the discrepancy between normally distributed variables was calculated via unpaired Student's t-test and the association was estimated by Pearson's correlation coefficient. Survival rates were measured by the Kaplan-Meier (KM) method, and the significance of disparity between survival curves was assessed via the log-rank test. Survival predictive accuracy of prognostic models was assessed based on a time-dependent receiver operating characteristic curve (ROC) analysis. Univariate ANOVA was used to calculate whether the clinical outcome (alive, dead with tumor and dead tumor free) was significantly affected by different grades of clinical parameters. Mean and standard deviation reflected the central trend and discrete trend of age in each clinical outcome. All statistical analyses were performed using R software (version 3.6.1; https://www.r-project.org/). P<0.01 was used as the threshold of genes considered with prognostic value calculated by univariate Cox analysis. P<0.05 was considered to indicate a statistically significant difference.

Results

Identification of survival-associated modules using WGCNA. A total of 554 IRGs were identified to be differentially expressed between normal pancreatic tissues and pancreatic tumors (Fig. S1A and B). In the WGCNA of TCGA cohort, two important parameters, R² (R²>0.9) and average connectivity were fully considered, hence β=4 was selected as the soft-thresholding power (Fig. 1A). According to the weighted
adjacency matrix constructed by the soft-thresholding power, the differential immune genes in the cohort from TCGA were classified into six diverse gene modules (Fig. 1B). The present study assessed the association between the module genes and characteristics including the survival time and status of patients with pancreatic cancer in the cohort from TCGA. Based on the correlation coefficient and the P-value, it was revealed that the blue module had the highest module significance (OS_time: cor=-0.38, P=1x10^-7; OS_status: cor=0.26, P=4x10^-4; Fig. 1C) and gene significance (OS_time: cor=0.61, P=2.9e-13; OS_status: cor=0.45, P=3.6e-7; Fig. 1D and E). Therefore, the blue module, which contained 117 IRGs, was selected for further analysis.

**Construction of the prognostic model based on IRGs.** Based on univariate Cox regression analysis of TCGA cohort, 42 genes with significant prognostic value were further extracted from the blue module (P<0.01; Fig. 2A). LASSO regression analysis was applied to avoid over-simulation of the model by adjusting the complexity of the classifier, and IRGs that correlated highly with one another were deleted (Fig. 2B and C). In the case of dimensionality reduction of the preliminary results using the LASSO method, the genes and their coefficients involved in the model were determined by multivariate Cox regression analysis. The coefficients of four IRGs are presented in Table SIV. Ultimately, four optimal IRGs [2’-5’-oligoadenylate synthetase 1 (OAS1), MET proto-oncogene, receptor tyrosine kinase (MET), interleukin 1 receptor type 2 (IL1R2) and interleukin 20 receptor subunit β (IL20RB)] were included in the prognostic model to calculate the RS, and the formula was as follows: RS=0.221350 x OAS1 + 0.515099 x MET + 0.179351 x IL1R2 + 0.141478 x IL20RB. The cut-off value was determined to be 4.15 using the surv_cutpoint function in the survminer R package, and the patients were divided into high- and low-RS groups. In the TCGA training set, the KM survival curve analysis demonstrated that the low-RS group had significantly improved survival compared with the high-RS group (P=3.88e-05; Fig. 3A), and the ROC curve revealed that the model had a good predictive function for the prognosis of patients with pancreatic cancer (ROC 1-year=0.738; Fig. 3B; ROC 2-years =0.691; Fig. 3C). Additionally, it was revealed that the RS could be used as an independent predictor [hazard ratio (HR), 1.958; P<0.001; Fig. 4] for the prognosis of patients with pancreatic cancer in the training cohort by multivariate regression analysis including RS, sex, age, TNM stage and tumor stage.

**Verification of the model for survival prediction in ICGC and GEO cohorts.** In order to confirm the robustness of the prognostic classifier used in the present study, the duplicate formula
for RS and cut-off value were used for the analysis of the ICGC and GEO validation sets. In each cohort, the 79 patients in the ICGC cohort and 66 patients in the GSE62452 dataset were divided into high- and low-RS groups, respectively. The KM survival curves (P=0.03692; Fig. 3D and G) and ROC curves (Fig. 3E, F, H and I) validated the reliability of the prognostic model used in the present study. Among them, although RS was not effective in predicting the 1-year survival rate (P>0.05; data not shown) of patients with pancreatic cancer in the GSE62452 dataset, it exerted a preferable effect on calculating the 3-year survival rate (Fig. 3G and I). Furthermore, to emphasize the superiority of the RS in predicting overall survival of patients with PDAC, time-dependent ROC curve analysis was used to compare the predictive effect of RS with classic disease classification parameters, and this revealed that RS was the best index in three independent databases (Fig. 3). The RS distribution, survival status and risk gene expression in the training cohort and validation cohort are presented in Fig. S2A-I.

Different function of immune infiltration between high- and low-RS group. The analysis results of the ESTIMATE algorithm revealed the difference of immune status between high- and low-RS groups in the cohort from TCGA (Fig. 5A-D). ssGSEA analysis revealed that the high-RS group had lower enrichment levels in multiple positively regulated immunity gene sets (‘Cytolytic activity’ and ‘Inflammation-promoting’) than the low-RS group (Fig. 5A), which suggested that various
immunity-associated biological processes and mechanisms were activated in both the aforementioned two groups, and increased immune activity may help patients with pancreatic cancer achieve improved outcomes. Notably, the high-RS group had a higher level of tumor purity and lower immune score than the low-RS group (P<0.05; Fig. 5B and D). The high-RS group also had a lower ESTIMATE score which reflected the comprehensive level of tumor purity, immune score and stromal score (P<0.05; Fig. 5C).

Different proportions of immune cell subsets between high- and low-RS groups. The CIBERSORT algorithm was used to evaluate the composition of 22 immune cell types in TCGA cohort. It was revealed that the resting CD4 memory cells occupied the highest proportion among all cells, while activated CD4 memory cells accounted for a small proportion (Fig. 6A). This suggested that the absence of immune cell components with immune activation functions in the immune microenvironment may serve an important role in accelerating the development of pancreatic cancer. By comparing the composition of immune cells in the high-RS group and the low-RS group of TCGA cohort, it was revealed that the high-RS group had a lower infiltration of CD8 T cells and a higher composition of M0 macrophages (Fig. 6B). Finally, the present study analyzed the correlation between the gene expression of four marker genes in the TCGA cohort and the composition of 22 immune cells. The results revealed that the expression levels of OAS1, MET and IL20RB genes were negatively correlated with the composition of M2 macrophages (Fig. 7A, B and D; P<0.01). With the increase of CD8 T cells, the gene expression levels of MET and IL1R2 were decreased (Fig. 7B and C; P<0.01), which suggested that CD8 T cells may repress the activation of the oncogenes observed in pancreatic cancer.

Enrichment analysis of upregulated genes in the high-RS group compared with the low-RS group. The results revealed that the upregulated genes were associated with 8 biological processes and 7 molecular functions, including ‘defense response to bacterium’, ‘defense response to other organisms’ and ‘receptor ligand activity’ (Fig. 8A). Through pathway analyses, the present study revealed that upregulated genes were associated with ‘cytokine-cytokine receptor interaction’ and ‘JAK-STAT signaling pathway’ (Fig. 8B).

Discussion

As a highly malignant tumor, over half of all PDAC cases are diagnosed at an advanced stage for which mortality rates closely parallel incidence rates. PDAC is also the fourth leading cause for cancer-associated mortality worldwide (30,31). With characteristics of insidious onset and early metastasis, ~80% of patients are diagnosed at a late stage of disease (32). At present, tumor immunotherapy results in marked effects in cancer treatment (33). The impact of the immune microenvironment on tumor cells has been demonstrated in numerous studies (34,35). Immunotherapy has emerged as an option for pancreatic cancer (36). Wartenberg et al (37) suggested that immunophenotypic classification is associated with tumor characteristics leading to pancreatic cancer with prognostic/predictive significance. Considering the significance of the immune microenvironment in cancer progression, finding immunity-associated biomarkers to predict the prognosis of patients with PDAC is necessary, and may also serve an important role in immunotherapy (38).

The present study revealed that a module including 117 IRGs was associated with pancreatic cancer prognosis using WGCNA. In the univariate Cox regression and LASSO analyses, four IRGs were included in the classifier (OAS1, MET, IL1R2 and IL20RB). The model could effectively predict the prognosis of patients with PDAC in the TCGA cohort [1-year area under the curve (AUC)=0.738; 2-years AUC=0.691]. In addition, the performance of the model was assessed using patients in the ICGC cohort (1-year AUC=0.702; 2-years AUC=0.638) and GSE62452 (2-years AUC=0.711; 3-years AUC=0.753). Using multivariate regression analyses, the RS of the model was demonstrated to be an independent factor for predicting the prognosis of patients with pancreatic cancer in the TCGA cohort (HR, 1.958; P=0.001).

The discovery of the following four genes provides guidance for searching for targeted genes for immunotherapy of pancreatic cancer. It has been reported that immune response by four IRGs (OAS1, MET, IL1R2 and IL20RB) is associated with the prognosis of pancreatic cancer (39-42). The OAS system is an antiviral signaling pathway induced by interferon and OAS genes that are described as interferon-stimulated genes (43,44). There are three types of OAS proteins in humans, OAS1, OAS2 and OAS3. OAS1 has been demonstrated to be associated with pancreatic cancer and prostate cancer (41,45). Oncolytic virus therapy is a promising treatment option for pancreatic cancer; however, high expression levels of OAS in cell lines are associated with resistance to this therapy (46). Therefore, OAS1 could be a potential immunotherapy target in pancreatic cancer and a prognostic biomarker to identify suitability of patients with pancreatic cancer to receive immunotherapy (45,47). The MET receptor, with a hepatocyte growth factor receptor ligand, is upregulated in several malignancies, including breast, lung and pancreatic cancer (48,49). Activation of several intracellular signaling pathways that is mediated by MET has led to the emergence of diverse cellular hallmarks of cancer, including cell proliferation, survival, invasion, migration, metastasis and inhibition.
of apoptosis (50‑52). MET may be involved in the malignant process of pancreatic cancer, and can serve as a biomarker for assessing the prognosis of patients with pancreatic cancer (53). IL1R2, as an endogenous inhibitor, can be highly expressed in the follicular helper T cells of mice, and the specific marker molecules, such as IL1R2, on the surface of tumor infiltrating Treg cells in colorectal cancer and non‑small cell lung cancer have also been demonstrated to be upregulated (54,55). Furthermore, another study on breast cancer revealed that IL1R2 is more abundant in tumor‑infiltrating T cells compared with Treg cells or peripheral blood T cells in normal breast tissues (56). In conclusion, IL1R2 is regarded as a negative regulatory factor that can be used as a potential target for immunotherapy.

IL20R, a heterodimeric receptor, is composed of two chains, interleukin‑20 receptor subunit α and IL20RB (57). IL20RB may affect the IL20 signaling pathway. Several studies have demonstrated that IL20RB serves a significant role in a number of different types of cancer, including breast cancer (58), lung cancer (59), nasopharyngeal carcinoma (60) and pancreatic cancer (40). Overall, the results of the present study may provide potential mRNA targets for immunotherapy and prognostic evaluation to help improve the clinical outcomes of pancreatic cancer.

TME serves a vital role in the development of pancreatic cancer (61). Additionally, it is the leading cause of chemotherapy resistance in patients with pancreatic cancer, as well as poor prognosis (62). TME is a complex integrated system comprising cancer cells and other components, such as extracellular matrix, surrounding blood vessels, inflammatory mediators and other non‑tumor cells (63). The mutational landscape of cancer cells can influence the phenotype and extent of immune cell infiltration, and more generally affect the entire TME (64,65). ESTIMATE is an algorithm based on gene expression signatures to estimate immune and stromal cells, as well as tumor purity (11). Immune and stromal cells have been authenticated as prognostic factors for tumors (66,67). The ESTIMATE algorithm has been utilized to analyze the large amount of data on glioblastoma (17), cholangiocarcinoma (68) and numerous other malignancies, while the immune and stromal scores of pancreatic cancer in

Figure 5. Comparison of immune response profile associated with the immune prognostic model in TCGA cohort. (A) Enrichment levels of immune cell types in high‑and low‑risk score groups of TCGA samples. (B) Distribution of tumor purity in the high‑and low‑risk groups. (C) Distribution of ESTIMATE scores in the high‑and low‑risk groups. (D) Distribution of immune scores in the high‑and low‑risk groups. *P<0.05. ESTIMATE, Estimation of Stromal and Immune cells in Malignant Tissues using Expression data; TCGA, The Cancer Genome Atlas.
Figure 6. Landscape of immune cell infiltration in TCGA cohort. (A) Proportion of 22 immune cell types in the TCGA cohort. Specific 22 immune cell types represented by various colors in each sample are shown in a barplot. (B) Comparison of proportions of 22 immune cell types in the low- and high-risk score groups. Blue represents the low risk group and red represents the high risk group. NK cells, natural killer cells; TCGA, The Cancer Genome Atlas.

Figure 7. Correlation between the expression levels of four genes included in the classifier and immune cells in samples from The Cancer Genome Atlas cohort (P<0.05). (A) OAS1. (B) MET. (C) IL1R2. (D) IL20RB. NK, natural killer; OAS1, 2’-5’-oligoadenylate synthetase 1; MET, MET proto-oncogene, receptor tyrosine kinase; IL1R2, interleukin 1 receptor type 2; IL20RB, interleukin 20 receptor subunit β; abs, absolute value.
prognostic value have not been adequately researched. Based on the ESTIMATE package, the present study demonstrated the difference in tumor purity and immune activity between the high- and low-RS groups. Compared with the high-RS group, the present study revealed that the low-RS group had a higher degree of enrichment in the gene sets reflecting positive immune function, such as cytolytic activity and inflammation-promoting, which revealed that immune activity had an impact on the survival and prognosis of patients with pancreatic cancer.

Figure 8. Enrichment analysis of upregulated genes in the high-RS group. (A) Gene Ontology enrichment analysis of upregulated genes in the high-RS group compared with the low-RS group. (B) Kyoto Encyclopedia of Genes and Genomes enrichment analysis of upregulated genes in the high-RS group compared with the low-RS group. RS, risk score; BP, biological process; MF, molecular function; P.adjust, adjusted P-value.
The CIBERSORT algorithm is able to assess 22 types of immune cells based on the gene expression data with high sensitivity and specificity (18). Several studies have suggested that the relative levels of various immune cells, analyzed by the CIBERSORT algorithm, could estimate the immune cell composition in tumors (69-71). In the present study, the difference in immune infiltration of pancreatic cancer cells in TCGA cohorts mainly consisted of macrophage M0 and T cells CD8, as determined by the CIBERSORT algorithm. In the immune infiltration analysis, the proportion of CD8 T cells in the high-RS group was relatively lower compared with the low-RS group. Additionally, the proportion of M0 macrophages in the high-RS group was relatively higher than low-RS group. This indicated that an imbalance in the CD8 T cells and M0 macrophage ratio may lead to a lower survival rate in the high-RS group. Tahkola et al. (72) revealed that the 5-year survival rates of the CD8 T cells low-, medium- and high-infiltration groups in pancreatic cancer were 4.2, 13.4 and 31.5%, respectively. Another study demonstrated that an imbalance in the proportion of immune cells was closely associated with low survival rate and poor clinical outcomes in patients with cancer (73). However, the specific effect of these differentially expressed chemotactic factors on immune infiltration of pancreatic cancer requires further investigation.

To summarize, in the present study, an IRG prognosis model of pancreatic cancer was constructed, and the effectiveness of the model was verified in the independent validation set. The difference in overall survival between high- and low-RS groups based on this model may be caused by the difference of immune infiltration and TME. The findings of the present study added certain guidance values to the analysis of pancreatic cancer pathogenesis and provide a reference for clinical treatment. There were some limitations to the present study, and thus, the validity of the conclusions should be confirmed by an investigation involving a larger number of cases, and further animal studies or cellular experiments are required to test the prognostic accuracy of the signatures.

Acknowledgements

Not applicable.

Funding

No funding was received.

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Authors’ contributions

SY, JF, YZ, YX and FF contributed to the study concept and design, and the acquisition, interpretation and analysis of the data. SY, JF and FF drafted the manuscript. FF was responsible for the integrity of the work as a whole. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References


