Screening of Therapeutic Targets for Pancreatic Cancer by Bioinformatics Methods

Authors
Xiaojie Xiao¹, Zheng Wan¹, Xinmei Liu², Huaying Chen³, Xiaoyan Zhao³, Rui Ding³, Yajun Cao³, Fangyuan Zhou³, Enqi Qiu³, Wenrong Liang³, Juanjuan Ou³, Yifeng Chen³, Xueting Chen⁴, Hongjian Zhang¹

Affiliations
¹ Department of Oncology and Vascular Interventional Radiology, Zhongshan Hospital Xiamen University, Xiamen, China
² Animal and Plant Inspection and Quarantine Technology Center Shenzhen Customs, Shenzhen Haiguan, Shenzhen, China
³ Zhongshan Hospital of Xiamen University, Zhongshan Hospital Xiamen University, Xiamen, China
⁴ Wanbei Coal and Electricity Group General Hospital, Suzhou, China

Key words
pancreatic cancer, differentially expressed genes, targets, bioinformatics, treatment

ABSTRACT
Pancreatic cancer (PC) has the lowest survival rate and the highest mortality rate among all cancers due to lack of effective treatments. The objective of the current study was to identify potential therapeutic targets in PC. Three transcriptome datasets, namely GSE62452, GSE46234, and GSE101448, were analyzed for differentially expressed genes (DEGs) between cancer and normal samples. Several bioinformatics methods, including functional analysis, pathway enrichment, hub genes, and drugs were used to screen therapeutic targets for PC. Fisher’s exact test was used to analyze functional enrichments. To screen DEGs, the paired t-test was employed. The statistical significance was considered at p < 0.05. Overall, 60 DEGs were detected. Functional enrichment analysis revealed enrichment of the DEGs in “multicellular organismal process”, “metabolic process”, “cell communication”, and “enzyme regulator activity”. Pathway analysis demonstrated that the DEGs were primarily related to “Glycolipid metabolism”, “ECM-receptor interaction”, and “pathways in cancer”. Five hub genes were examined using the protein-protein interaction (PPI) network. Among these hub genes, 10 known drugs targeted to the CPA1 gene and CLPS gene were found. Overall, CPA1 and CLPS genes, as well as candidate drugs, may be useful for PC in the future.
Introduction

Pancreatic cancer (PC), a common tumor of the gastrointestinal tract, has a poor survival rate [1]. This is primarily because PC is hidden on the posterior side of the right upper abdomen [2]. Patients may be unaware of the initial symptoms such as upper abdominal discomfort, weight loss, yellowing of the skin, fatigue, and cognitive issues, making them easily overlooked. Moreover, the lack of precise biomarkers for PC aggravates the issue [3]. Lack of effective treatments are the second reason for poor survival rate. Various treatments such as surgery, radiotherapy, and chemotherapy are typically utilized. Nevertheless, those who have undergone surgery have a high chance of relapse, and are not as responsive to radiation or chemotherapy treatments [4]. Resultantly, more effective biomarkers or novel treatments for PC are warranted.

Bioinformatics methods have been used in numerous diseases, including cancers [5–7], further providing novel insights into cancer. A few bioinformatics studies could examine only one gene associated with PC [8–10]. Tumors are not exclusively caused by a single gene, but rather are the result of several genetic factors combined. Moreover, the above studies ignored targeted drugs for cancer. Hence, the diagnosis and management of PC is a difficult task and its comprehensive exploration has attracted intense curiosity.

Novel uses of earlier drugs can be a revolutionary development [11]. Drugs for non-cancerous have the potential to treat cancer. For instance, statins used for patients undergoing heart failure treatment have demonstrated anti-tumor activity [12–14]. Aspirin, an antiplatelet drug, has shown anti-tumor effects as well [15–17]. Hence, it is hypothesized that some existing drugs could be useful in the treatment of PC.

The objective of the current study was to identify target genes and drugs in PC using several bioinformatical methods. First, three pooled datasets were selected from the Gene Expression Omnibus (GEO) database. Second, differentially expressed genes (DEGs) were detected between PC patients and healthy individuals. Next, these DEGs were analyzed using several bioinformatics methods. Finally, the potential biomarkers and drugs targeted to PC were identified. Expectedly, the present study may offer a promising treatment for PC.

Materials and Methods

Data summary

Gene Expression Omnibus (GEO) database stores microarray and high-throughput gene expression data [18]. Three datasets, namely GSE62452, GSE46234, and GSE101448, were obtained from GPL6244, GPL570, and GPL10558 platforms, respectively, in the GEO database. GSE62452 had 61 cancer and 69 normal tissues; GSE46234 comprised four cancer and four normal samples; GSE101448 showed 19 cancer and 24 normal samples (Supplement ▶ Table 1S).

Ethics statement

As the data were re-analyzed from the public dataset, no ethical approval by the local ethics committee was necessary.

DEGs identification

GEO2R, an interactive web tool, was employed to identify the DEGs between PC and normal specimens [19]. The upregulated DEGs are logFC > 1 and p < 0.05. The opposite logFC are the downregulated DEGs. Venn diagram tool (http://bioinformatics.psb.ugent.be/webtools/Venn/) was applied to obtain common DEGs.

Functional and pathway enrichment analysis

Database for Annotation, Visualization, and Integrated Discovery (DAVID), an online bioinformatics tool, was used for Gene ontology (GO) function and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses [19].

Protein-protein interaction (PPI) network

To establish an association among the DEGs and construct the PPI network, the Search Tool for the Retrieval of Interacting Genes (STRING, http://string-db.org) was applied [20]. Subsequently, Cytoscape version 3.7.2 was used to visualize the PPI network. The MCODE (Molecular Complex Detection) plugin from Cytoscape analyzed the hub genes [21].

Drug screening

The Drug Gene Interaction Database (DGIdb) (https://www.dgidb.org) was used to search for drugs associated with hub genes.

Statistics analysis

Fisher’s exact test was employed to evaluate functional enrichments. The t-test was applied to screen DEGs. A value of p < 0.05 indicated statistical significance.

Results

DEGs Identification

Venn diagram depicts 60 genes, including 33 upregulated (▶ Fig. 1a) and 27 downregulated genes (▶ Fig. 1b) overlapping among three datasets. ▶ Table 1 lists the names of DEGs.

<table>
<thead>
<tr>
<th>Term</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upregulated</td>
<td>KIA1A324, CELA3A, CEL, EGF, AQP8, CLPS, TRHDE, CPB1, GP1, PDK4, RBPI1, PRSS3P2, PDIA2, CTRC, IAPP, PLA2G1B, CELA3B, ERP27, CELA2B, ERP44, CTRL, TMED6, ALB, AOX1, F11, CPA2, REG1B, PNLIPRP2, CPA1, NRS5A2, PNLIPRP1, KLK1, SERPINI2</td>
</tr>
<tr>
<td>Downregulated</td>
<td>SERPINB5, CEACAM6, COL1A1, FN1, LAMB3, DPCR1, SLPI, NOX4, CDH11, ITGA2, SLCA1A4, COL3A1, ANXA10, POSTN, CEACAMS, TMC5, CTSE, GABRP, THBS2, KRT19, SULF1, LAMC2, AHNAK2, TFF1, CLDN18, CP, AGR2</td>
</tr>
</tbody>
</table>

▶ Table 1 List of the differentially expressed genes (DEGs).
Functional enrichment analysis

For functional enrichment, biological process (BP) terms were clustered in the “multicellular organismal process”, “biological regulation”, “cell communication”, “response to stimulus”, and “metabolic process”. Besides, cellular component (CC) terms were associated with “endomembrane system”, “extracellular space”, “vesicle”, “membrane” and “protein-containing complex”. In Molecular Function (MF) annotation, functional enrichment was associated with “hydrolase activity”, “structural molecule”, “protein binding”, “ion binding,” and “enzyme regulator” (Fig. 2). KEGG pathway revealed enrichment in “small cell lung cancer”, “glycolipid metabolism”, “ECM-receptor interaction”, “pathways in cancer”, and “focal adhesion” (Table 2).

The construction of PPI

Forty-eight genes and 145 edges were clustered in the PPI network (Fig. 3a). Top genes were selected via the MCODE plugin. Fig. 3b shows nine top genes (CLPS, CELA3B, CPA2, CELA3A, CPA1, CPB1, CTRC, CTRL, and PRSS3P2).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa04512</td>
<td>ECM-receptor interaction</td>
<td>7.55E-07</td>
</tr>
<tr>
<td>hsa04510</td>
<td>Focal adhesion</td>
<td>9.85E-06</td>
</tr>
<tr>
<td>hsa05222</td>
<td>Small cell lung cancer</td>
<td>0.004666</td>
</tr>
<tr>
<td>hsa00561</td>
<td>Glycerolipid metabolism</td>
<td>0.01453</td>
</tr>
<tr>
<td>hsa05200</td>
<td>Pathways in cancer</td>
<td>0.0425</td>
</tr>
</tbody>
</table>

Table 2 Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of the differentially expressed genes (DEGs).
Screening the drugs

The top nine genes were employed to find drugs. CPA1 and CLPS genes matched with 10 drugs (▶ Table 3). In the KEGG pathway, these genes were associated with the “fat digestion and absorption pathway”, “pancreatic secretion” and “protein digestion and absorption” (▶ Fig. 4).

Discussion

PC has the highest mortality and lowest survival rates of all cancers due to its difficulty to be detected in the early stages and the lack of effective treatments. Therefore, identifying biomarkers to diagnose or treat PC becomes urgent. Data sequencing can reveal the underlying diagnostic and prognostic mechanisms of different diseases, especially cancer. The development of related medications has opened up a new way to examine cancer and hypothesize about its molecular causes.

In this study, GSE62452, GSE46234, and GSE101448 datasets were analyzed for DEGs between abnormal and normal tissues. Sixty DEGs were screened. BP terms were clustered in the “multicellular organismal process”, “biological regulation”, “cell communication”, “response to stimulus”, and “metabolic process”. Further, CC terms were associated with “endomembrane system”, “extracellular space”, “vesicle”, “membrane”, and “protein-containing complex”. MF annotation revealed an association with “hydrolase activity”, “structural molecule”, “protein binding”, “ion binding” and “enzyme regulator”. In the KEGG pathway, PC was enriched in “small cell lung cancer”, “glycolipid metabolism”, “ECM-receptor interaction”, “pathways in cancer”, and “focal adhesion”. These results revealed an association of abnormal lipid metabolism with PC. Numerous research papers have established a link between lipid metabolism disorders and PC, in agreement with our results [22–24].

A total of 48 genes with 145 edges were included in the PPI part. Thereafter, hub genes were selected by the MCODE algorithm. Nine top genes, namely CLPS, CELA3B, CPA2, CELA3A, CPA1, CPB1, CTRC, CTRL, and PRSS3P2 were employed to identify drugs. CPA1 and CLPS genes matched with 10 drugs. In the KEGG pathway, these genes showed association with “pancreatic secretion”, “protein digestion and absorption”, and “fat digestion and absorption pathway”.

The protein encoded by the co-enzyme colipase (CLPS), a co-factor for efficient dietary lipid hydrolysis, performs tissue-specific regulation of expression in pancreatic alveolar cells [25, 26]. CLPS is key to the development and progression of PC and is a likely target for treatment [27]. Furthermore, CLPS has been reported to contribute to type 2 diabetes development [28].

Carboxypeptidase A1 (CPA1), a zinc metalloprotease produced by pancreatic alveolar cells, plays a vital role in the cleavage of C-terminal branched chains from dietary proteins [29]. When comparing the differentiating marker between normal and neoplastic pancreatic alveolar cells, CPA1 displays high sensitivity [29, 30]. Besides, the CPA1 variant aggravates the risk of chronic pancreatitis [31]. Hence, CLPS and CPA1 genes were associated with PC. We found 10 medications that have been given the green light by the FDA, which could potentially be useful in treating PC, and are specifically targeted at CLPS and CPA1 genes.

Conclusion

Overall, CPA1 and CLPS genes as well as candidate drugs were identified by bioinformatics methods in this study. This study may offer a novel idea for the diagnosis and treatment of PC.

▶ Table 3 The known drugs associated with CAP1 and CLPS genes.

<table>
<thead>
<tr>
<th>Drug ID</th>
<th>Drug name</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB04058</td>
<td>d-[(Amino)carbonyl]phenylalanine</td>
<td>0.001036</td>
</tr>
<tr>
<td>DB03441</td>
<td>2-Benzyl-3-iodopropanoic Acid</td>
<td>0.001142</td>
</tr>
<tr>
<td>DB04316</td>
<td>d-[(N-Hydroxyamino)carbonyl]phenylalanine</td>
<td>0.001628</td>
</tr>
<tr>
<td>DB08222</td>
<td>Methoxyundecylphosphonic Acid</td>
<td>0.001753</td>
</tr>
<tr>
<td>DB04233</td>
<td>(Hydroxyethyl)octane</td>
<td>0.002442</td>
</tr>
<tr>
<td>DB06924</td>
<td>(2R)-2-Benzyl-3-nitropropanoic acid</td>
<td>0.002850</td>
</tr>
<tr>
<td>DB03012</td>
<td>Phenylalanine-N-sulfonamide</td>
<td>0.003800</td>
</tr>
<tr>
<td>DB02451</td>
<td>8-Nonylguloside</td>
<td>0.00455</td>
</tr>
<tr>
<td>DB03201</td>
<td>d-Cysteine</td>
<td>0.00570</td>
</tr>
<tr>
<td>DB02494</td>
<td>α-Hydroxy-β-phenylpropionic Acid</td>
<td>0.00995</td>
</tr>
</tbody>
</table>
Author Contributions
(I) Conception and design: Zhang Hongjian; (II) Administrative support: Wan Zheng; (III) Provision of study materials: Xiao Xiaojie; (IV) Collection and assembly of data: Xiao Xiaojie, Liu Xinmei, Chen Huaying, and Zhao Xiaoyan; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Funding Information
Natural Science Foundation of Fujian Province — 2022J05298

Acknowledgements
This study was supported by Natural Science Foundation of Fujian Province (Grant NO.2022J05298).

Conflict of Interest
The authors declare that they have no conflict of interest.

References

Fig. 4 Screening drugs for hub genes. a: The potential drugs targeted the CPA1 and CLPS. b: The pathway associated with CPA1 and CLPS genes.


