Integrated Analysis of Glutathione Metabolic Pathway in Pancreatic Cancer

Xingui Wu1,2†, Ruyuan Yu1,2†, Meisongzhu Yang1,2, Yameng Hu1,2, Miaoling Tang1,2, Shuxia Zhang1,2, Aninaerjiang Abudourousuli1,2, Xincheng Li1,2, Ziwen Li1,2, Xinyi Liao1,2, Yingru Xu1,2, Man Li1,2, Suwen Chen1,2, Wanying Qian1,2, Rongni Feng1,2, Jun Li1,2† and Fenjie Li1,3*

1Program of Cancer Research, Affiliated Guangzhou Women and Children’s Hospital, Zhongshan School of Medicine, Sun Yat-Sen University, Guangzhou, China, 2Department of Biochemistry, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, China, 3Department of Pediatric Surgery, Guangdong Provincial Key Laboratory of Research in Structural Birth Defect Disease, Guangzhou Women and Children’s Medical Center, Zhongshan School of Medicine, Sun Yat-Sen University, Guangzhou, China

Metabolic enzyme-genes (MEs) play critical roles in various types of cancers. However, MEs have not been systematically and thoroughly studied in pancreatic cancer (PC). Global analysis of MEs in PC will help us to understand PC progressing and provide new insights into PC therapy. In this study, we systematically analyzed RNA sequencing data from The Cancer Genome Atlas (TCGA) (n = 180 + 4) and GSE15471 (n = 36 + 36) and discovered that metabolic pathways are disordered in PC. Co-expression network modules of MEs were constructed using weighted gene co-expression network analysis (WGCNA), which identified two key modules. Both modules revealed that the glutathione signaling pathway is disordered in PC and correlated with PC stages. Notably, glutathione peroxidase 2 (GPX2), an important gene involved in glutathione signaling pathway, is a hub gene of the key modules. Analysis of immune microenvironment components reveals that PC stage is associated with M2 macrophages, the marker gene of which is significantly correlated with GPX2. The results indicated that GPX2 is associated with PC progression, providing new insights for future targeted therapy.

Keywords: pancreatic cancer, metabolic enzyme genes, glutathione metabolism, GPX2, WGCNA

INTRODUCTION

PC remains a highly fatal malignancy and is the fourth leading cause of cancer-related mortality in both sexes in the United States (Rahib et al., 2014; Morrison et al., 2018; Mizrahi et al., 2020), while it is the sixth leading cause of cancer-related mortality in China (Song et al., 2013; Zhu et al., 2018; Zhao et al., 2019). The incidence and mortality of PC have increased year on year. Surgical resection remains the only potentially curative treatment for PC (Sala Elarre et al., 2019). However, nearly two-thirds of patients lose the chance of...
surgery because of rapid progression and metastasis of PC. Even with early surgically resection, the recurrence rate is still high. For patients with locally advanced PC or distal metastasis, the 5-survival rate is also poor (Mizrahi et al., 2020). Hence, identifying new biological mechanisms and developing new strategies for PC treatment are urgently required.

**FIGURE 1** | In PC, metabolic enzyme-genes are abnormally expressed. (A) Venn diagram of metabolic enzyme-genes in human pancreatic cancer. (B,C) Comparison of the expression levels of other highly expressed genes and metabolic enzyme genes in the TCGA and GSE15471 data (**, Pvalue < 0.01). (D) Venn diagram of differentially expressed genes and metabolic enzyme genes in the TCGA and GSE15471 data. The KEGG pathway enrichment analysis for the differentially expressed genes. (E) The TCGA and GSE15471 data showed that metabolic pathways, including the glutathione pathway, were enriched in both sets of data (p < 0.05).
FIGURE 2 | Co-expression analysis of MEs from the GSE15471 data. (A) Clustering tree of PC samples in the GSE15471 data. (B) Gene distribution in the WGCNA network analysis. (C) Heatmap plot of the topological overlap in the MEs network. (D) Analysis of the relationships between genes in modules between tumor and normal samples. (E) In the black module, for most genes, expression in the tumor samples was higher than that in the normal samples.
FIGURE 3 | Co-expression analysis of metabolic enzymes genes from the TCGA data (A) Clustering tree of PC samples in the TCGA data. (B) Gene distribution in the WGCNA network analysis. (C) Heatmap plot of the topological overlap in the MEs network. (D) Analysis of relationships between genes in modules between tumor and normal samples. (E) Distribution of pathologic_stage-related genes in all modules. Genes are presented on the X-axis, and the enrichment significance is shown on the Y-axis.
Aberrant metabolism, especially redox homeostasis, is a major hallmark of cancer, and these changes promote the acquisition and maintenance of malignant properties (Gray et al., 2014). Glutathione is the most important regulator of the redox homeostasis, which can be altered by the activity of glutathione peroxidase (Roth et al., 2002; Vander Heiden et al., 2009; Levine and Puzio-Kuter, 2010; Lunt and Vander Heiden, 2011). Glutathione metabolism plays an important role in many cellular processes, including cell differentiation, proliferation and apoptosis (Ballatori et al., 2009; Cacciare et al., 2010; Traverso et al., 2013). It was also reported that the expression of genes related to glutathione metabolism are increased in many tumors. Among them, GPX2 is an important antioxidant enzyme in...
FIGURE 5 | Glutathione signaling pathway genes are associated with PC stage. (A,B) Heatmap of differentially expressed glutathione pathway genes in the GSE15471 and TCGA datasets. (C) Four quadrant diagram of the glutathione genes that differ between the GSE15471 and TCGA datasets. (D) The expression level of RRM2, GSTP1 and GPX2 in our own RNA-seq data (**p value < 0.01). (E) In the TCGA data, the RRM2, GSTP1 and GPX2 genes were significantly differently expressed in different stages of PC (**p < 0.05). (F) Survival analysis and Relapse-free Survival analysis of RRM2 (left and middle) in TCGA dataset; survival analysis of GSTP1 in TCGA dataset (right).
FIGURE 6 | GPX2 is associated with M2 macrophages, which predicts cancer immune heterogeneity. (A) Comparison of immune cell fractions among subtypes. Immune cell fraction distribution in 36 normal samples and 36 pancreatic cancer groups. (B) Comparison of immune cell fractions between subtypes. Among them, 21 samples in Stage I, 151 samples in Stage II, 4 samples in Stage III and 5 samples in Stage IV, with a simple row score >0.05. (C) Scatter plots showing the correlation between GPX2 and CD163.
glutathione metabolism, which can scavenge a variety of peroxides (Brigelius-Flohe and Kipp, 2012; Brigelius-Flohe and Maiorino, 2013; Sakamoto et al., 2014; Huang et al., 2018). Additionally, GPX2 is highly expressed in many tumors and can promote tumor growth (Naiki et al., 2018; Du et al., 2020; Li et al., 2020). However, little is known about the role of GPX2 in PC.

In this study, we analyzed the RNA sequencing data from TCGA (n = 180 + 4) and GSE15471 (n = 36 + 36). We identified the differentially expressed genes (DEGs), which were then subjected to Gene Ontology (GO) (Ashburner et al., 2000) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto, 2000) pathway enrichment analysis. It revealed that metabolic pathways are significantly activated in PC. Co-expression network modules of MEs were constructed using WGCNA (Langfelder and Horvath, 2008), which identified two key modules associated with PC. The glutathione signaling pathway is included in the two modules and correlated with PC stages. GPX2, identified as a hub gene of the modules, is associated with M2 macrophages and is significantly upregulated during PC progression. The findings suggest GPX2 as a good marker of tumor heterogeneity and provides a basis for future targeted therapy.

MATERIALS AND METHODS

The Metabolic enzymes Datasets and Patient Information Acquisition
We organized the 1689 MEs from the HMDB database (Supplementary Table S1) (Wishart et al., 2018). We used 180 cases RNA-seq of pancreatic ductal adenocarcinoma (PAAD and 4 cases of normal tissue), with clinical information for WGCNA analysis in TCGA (http://www.cancer.gov/tcga). Meanwhile, we also organized 36 normal samples and 36 tumor samples (GSE15471) of microarray from the GEO database.

Microarray Analysis
The microarray data GSE15471 were downloaded from the Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) (Barrett et al., 2005). GSE15471 dataset contained 36 PC tissues and 36 paired adjacent normal tissues. For further analysis, the data was profiled on the GPL570 platform (Affymetrix, Human Genome U133 plus 2.0 Array). Differentially expressed genes (DEGs) were identified using the R software (version 4.0.5) package limma. The cutoff criteria were \( p_{\text{value}} < 0.05 \) and \( \log_{2}(\text{fold change})(FC) > 0.586 \), while RNAs with low expression values was filtered out.

TCGA-PAAD RNA-Seq Analysis
The R software (version 4.0.5) package DEGseq (Wang et al., 2010) were used to analyze the RNA-seq from the TCGA-PAAD datasets and to identify DEGs. The cutoff criteria were \( p_{\text{value}} < 0.05 \) and \( \log_{2}(\text{fold change})(FC) > 0.586 \), and filtered out RNAs with low expression values (the number of samples with normalized gene expression < 1 was more than half of the total number of samples).

Weighted Gene Co-Expression Network Analysis
RNA-seq and microarray datasets used above was analyzed by WGCNA to construct a co-expression network (Langfelder and Horvath, 2008). The expression profiles of MEs were obtained from the TCGA dataset and GSE15471 datasets, respectively. Then the samples were clustered through the systematic cluster tree to determine any outliers.

The soft-thresholding function was used calculate the power parameter. The dynamic tree cut method was used to identify the modules of co-expressed gene. Then a dendrogram of genes was obtained using a hierarchical clustering approach, which was based on dissimilarity of the unsigned topological overlap matrix (TOM). Lastly, genes with similar expression profiles were grouped and network modules was produced.

Protein-Protein Interaction Network Construction
We chose two modules from the TCGA and GSE15471 by screening edges and the criterion of weight value was >0.02 for the brown module (TCGA) and the black module (GSE15471) respectively. To visualize the co-expression network and identify the nodes and hub genes, the results were input into Cytoscape to (v3.8.0; downloaded from the website https://cytoscape.org/) (Shannon et al., 2003).

Functional Enrichment Analysis of Genes
We used the KEGG Orthology-Based Annotation System (KOBAS) (Xie et al., 2011) for GO and KEGG pathway enrichment analysis. A \( p_{\text{value}} < 0.05 \) was set as the cutoff criterion.

mRNA Expression Data Analysis Using CIBERSORT
CIBERSORT (https://cibersort.stanford.edu/index.php) (Newman et al., 2015) was used to analyze the mRNA expression data for immune microenvironment components determination.

RESULTS

Metabolism Enzyme Genes Show Significantly Higher Expression Than Other Genes in Pancreatic Cancer
We used the HMDB database and identified a total of 1689 metabolism-related genes (Supplementary Table S1). Then we extracted data for 180 samples of PAAD and 4 samples of normal tissue from the TCGA dataset and used the GSE15471 dataset comprising 36 PAAD samples and 36 normal samples. We found that 92.5% (1563) of the metabolism-related genes were expressed in both datasets (Figure 1A). The expression levels of metabolism-related genes were significantly higher than those of the other expressed genes in
Co-Expression Network Modules of MEs Identified by Weighted Gene Co-Expression Network Analysis in the GSE15471 Data

We performed WGCNA analysis on metabolic genes in GSE15471. We found that tumor tissue could be well differentiated from normal tissue using the created metabolic gene clusters (Figure 2A). We detected 7 modules in the network (Figure 2B). Genes in the seven modules showed a high correlation with each other (Figure 2C). The black module was also found to have the strongest correlation with pancreatic cancer patients, with a coefficient of correlation of 0.56 ($p = 1 \times 10^{-3}$) (Figure 2D). It shows that in the black module, gene expression was higher in most tumor samples than in normal tissue (Figure 2E). Therefore, this module was selected as the most clinically important screening module for further analysis.

Co-Expression Network Modules of Metabolic Enzyme-Genes Identified by Weighted Gene Co-Expression Network Analysis in the The Cancer Genome Atlas Data

We also subjected metabolic genes in the TCGA to WGCNA analysis. Since there were only four cases of normal samples of pancreatic cancer in the RNA-seq data of TCGA, we used other clinical indicators, such as sex, age, neoplasm histologic grade (NHG), pathologic N, and pathologic stage in the analysis (Figure 3A). Subsequently, we detected nine modules in the network (Figure 3B). Genes in the nine modules showed a high correlation with each other (Figure 3C). The results indicated that the brown module had the most significant correlation with pathologic stage (correlation coefficient = −0.21, $p = 0.008$, Figure 3D). We examined the modules’ average gene significance related to pathologic stage, and the brown module was also found to have the most negative association with the pathologic stage of PC (Figure 3E). Therefore, the brown module was selected as the most clinically important screening module for further analysis.

Protein-Protein Interaction Network of Important Modules Reveals That the Glutathione Signaling Pathway is Dysregulated

Then the black module of the GSE15471 data was used to construct a PPI network (Figure 4A), in which GPX2 and GSTP1 (encoding glutathione S-Transferase Pi 1) were identified as the hub genes. The black module was analyzed by KEGG, which showed that the glutathione pathway was significantly enriched (Figure 4B). Then we also used the brown module of TCGA to create a PPI network in which the hub genes was MANIC1 (mannosidase alpha class 1C member 1) gene (Figure 4C). The brown module was then analyzed by KEGG, which showed that the glutathione pathway was also significantly enriched (Figure 4D). We then compared the expression levels of the metabolic enzyme genes and found that the gene expression levels of members of the glutathione pathway were significantly higher than those of other metabolic genes (Figure 4E). These results indicated that the glutathione pathway are significantly dysregulated in pancreatic cancer.

Glutathione Signaling Pathway Genes are Associated With PC Stages

Since glutathione pathway has a great impact on the clinical evaluation of PC, we examined the expression level of genes involved in glutathione pathway. In the GSE15471 data, 14 genes in the glutathione pathway were differentially expressed (|log2(FC)| >0.586, $p_{\text{value}} < 0.05$) (Figure 5A). In the TCGA, 12 glutathione pathway genes were found to be significantly differentially expressed (Figure 5B). We found that glutathione pathway genes, including GPX2, GSTP1, RRM2, are differentially expressed in both data (Figure 5C). We also verified the expression of GPX2, GSTP1 and RRM2 in glutathione pathway in our own RNA sequencing data and the results was consistent with that of public data we used (Figure 5D). Among them, GPX2, GSTP1 and RRM2 (Ribonucleotide Reductase Regulatory Subunit M2) showed different expression levels in different stages of PAAD (Figure 5E) (**$p_{\text{value}} < 0.01$). In stage II, the expression levels of these three genes were relatively high, suggesting they could be clinical indicators and providing a preliminary basis for later tumor heterogeneity. Then we performed the survival analysis of the glutathione pathway genes in patients with PC. The results indicated that patients with higher expression of glutathione pathway genes, including GSTP1 and RRM2, have shorter survival term (Figure 5F). Collectively, these data indicated that glutathione pathway genes may play critical roles in PC.

GPX2 is Associated With M2 Macrophages, Which Predicts Cancer Immune Heterogeneity

Since glutathione pathway genes, including GPX2, showed different expression levels in different stages of PAAD, we tried to predict immune heterogeneity based on gene expression profiles in PC at different stages. To comprehensively determine the cellular components of the tumor microenvironment across different stages of pancreatic cancer, we used the CIBERSORT in silico cytometry method (Newman et al., 2015) to evaluate 22 different immune cell types in 36 normal tissues and 36 tumor tissues in
GSE15471, and 153 tumor tissues at different stages in TCGA. In the results for the GSE15471 data, PC appeared to promote an M2 macrophage gene signature compared with normal pancreatic tissue (Figure 6A). In the results of the TCGA analysis, M2 macrophages were more abundant in patients with stage IV disease compared with the patients at other stages (Figure 6B). We evaluated the correlation between the expression of GPX2 and CD163, the marker of M2, and it showed they are significantly correlated (Figure 6C). These results suggested that GPX2 may play important roles in cancer immune heterogeneity.

**DISCUSSION**

In the present study, we analyzed TCGA and GSE15471 data and found significant differences in glutathione pathway gene expression in PC. In normal tissues, the expression levels of GPX2, GSTP1 and RRM2 were lower than those in PC cells. Then, using the TCGA data, we found that the expression levels of GPX2, GSTP1 and RRM2 were also different in different stages of PC. In addition, these three genes showed higher expression in stage II than stage I samples. These results are also consistent with the findings of previous studies.

Transcriptome profiling and microarray of tumor samples are widely used to interrogate pathway functionality and for phenotype-based patient classification. In our study, we found that there are many M2 macrophages around PC cells. To develop markers for PC staging, further single-cell sequencing is required to finely dissect the expression of glutathione pathway genes in PC cells.

In conclusion, bioinformatic analyses revealed that eight genes of the glutathione pathway are differentially expressed between PC and normal pancreas tissues, and three of them show significant expression differences in different stages of PC staging. These results help to refine the cellular heterogeneity of PC. Exploiting such cellular heterogeneity, targeted drugs are used to kill early PC cells at a later stage. This can prevent pancreatic cancer from progressing, thereby prolonging the survival of patients with PC and buying time for the development of further treatment.

**REFERENCES**

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene Ontology: Tool for the Unification of Biology. Nat. Genet. 25 (1), 25–29. doi:10.1038/75556

Ballatori, N., Krance, S. M., Notenboom, S., Shi, S., Tieu, K., and Hammond, C. L. (2009). Glutathione Dysregulation and the Etiology and Progression of Human Diseases. Biol. Chem. 390 (3), 191–214. doi:10.1515/BC.2009.033

Barrett, T., Suzek, T. O., Troup, D. B., Wilhite, S. E., Ngage, W. C., Ledoux, P., et al. (2004). NCBI GEO: Mining Millions of Expression Profiles-Database and Tools. Nucleic Acids Res. 33, D562–D566. Database issue. doi:10.1093/nar/gkh022

Brigelius-Flohé, R., and Kipp, A. P. (2012). Physiological Functions of GPx2 and its Role in Inflammation-Triggered Carcinogenesis. Ann. N. Y Acad. Sci. 1259, 19–25. doi:10.1111/j.1749-6632.2012.06574.x

**DATA AVAILABILITY STATEMENT**

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

**ETHICS STATEMENT**

Data used in this project has been publicly published, which has obtained patient consent and approval.

**AUTHOR CONTRIBUTIONS**

FL and JL designed the project. XW performed data analysis and statistical calculation. XW and RY prepared the original manuscript. All authors reviewed and edited the final manuscript. FL and JL supervised the whole study.

**FUNDING**

This work was supported by Natural Science Foundation of China (No. 82030078, 81830082, 82003128, 82072609, 81621004 and 31900519).

**ACKNOWLEDGMENTS**

We would like to thank the native English speaking scientists of Elixigen Company (Huntington Beach, California) for editing our manuscript.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2022.896136/full#supplementary-material

Brigelius-Flohé, R., and Maiorino, M. (2013). Glutathione Peroxidases. Biochim. Biophys. Acta (Bba) - Gen. Subjects 1830 (5), 3289–3303. doi:10.1016/j.bbagen.2012.11.020

Cacciatore, I., Cornacchia, C., Pinnen, F., Mollica, A., and Di Stefano, A. (2010). Prodrug Approach for Increasing Cellular Glutathione Levels. Molecules 15 (3), 1242–1264. doi:10.3390/molecules15031242

Du, H., Chen, B., Jiao, N.-L., Liu, Y.-H., Sun, S.-Y., and Zhang, Y.-W. (2020). Elevated Glutathione Peroxidase 2 Expression Promotes Cisplatin Resistance in Lung Adenocarcinoma. Oxidative Med. Cell Longevity 2020, 1–14. doi:10.1155/2020/7370157

Gao, X., Jiang, M., Chu, Y., Han, Y., Jin, Y., Zhang, W., et al. (2022). ETV4 Promotes Pancreatic Ductal Adenocarcinoma Metastasis through Activation of the CXCL13/CXCR5 Signaling axis. Cancer Lett. 524, 42–56. doi:10.1016/j.canlet.2021.09.026

Gray, L. R., Tompkins, S. C., and Taylor, E. B. (2014). Regulation of Pyruvate Metabolism and Human Disease. Cell. Mol. Life Sci. 71 (14), 2577–2604. doi:10.1007/s00018-013-1539-2
