Identification of genes and pathways associated with pancreatic ductal adenocarcinoma by bioinformatics analyses

JIN LONG, ZHONGBO ZHANG, ZHE LIU, YUANHONG XU and CHUNLIN GE

Department of General Surgery, The First Hospital of China Medical University, Shenyang, Liaoning 110001, P.R. China

Received December 11, 2014; Accepted November 12, 2015

DOI: 10.3892/ol.2015.4042

Abstract. This study aimed to explore the underlying genes and pathways associated with pancreatic ductal adenocarcinoma (PDAC) by bioinformatics analyses. Gene expression profile GSE43795 was downloaded from the Gene Expression Omnibus database. The differentially expressed genes (DEGs) between six PDAC and five non-neoplastic pancreatic tissue samples were analyzed using the limma package. Gene ontology (GO) and pathway enrichment analyses of DEGs were performed, followed by functional annotation and protein-protein interaction (PPI) network construction. Finally, the sub-network was identified and pathway enrichment analysis was performed on the contained DEGs. A total of 374 downregulated and 559 upregulated DEGs were identified. The downregulated DEGs were enriched in GO terms associated with digestion and transport and pathways related to metabolism, while the upregulated DEGs were enriched in GO terms associated with the cell cycle and mitosis and pathways associated with the occurrence of cancer including the cell cycle pathway. Following functional annotation, the oncogene pituitary tumor-transforming 1 (PTTG1) was upregulated. In the PPI network and sub-network, cell division cycle 20 (CDC20) and BUB1 mitotic checkpoint serine/threonine kinase B (BUB1B) were hub genes with high connectivity degrees. Additionally, DEGs in the sub-network including cyclin B1 (CCNB1) were mainly enriched in the cell cycle and p53 signaling pathways. In conclusion, the cell cycle and p53 signaling pathways may play significant roles in PDAC, and DEGs including CDC20, BUB1B, CCNB1 and PTTG1 may be potential targets for PDAC diagnosis and treatment.

Introduction

Pancreatic ductal adenocarcinoma (PDAC) ranks among the most malignant of human cancers (1). Currently, the annual number of associated mortalities is similar to the disease’s annual incidence (2). The poor prognosis of PDAC is correlated with the nonspecificity of symptoms, advanced disease at presentation and lack of effective adjuvant and systemic therapy (3). Currently, surgical extirpation for localized disease offers the only chance of long-term survival (3). Therefore, understanding the pathological mechanisms to detect PDAC as early as possible is an urgent requirement to enable further advances in therapeutic modalities and agents.

Previous studies have identified that the development of PDAC may involve certain genetic factors including the overexpression of oncogenes, inactivation of tumor suppressor genes or the deregulation of various signaling pathways (4). For instance, kirsten rat sarcoma viral oncogene homolog mutations have been observed to occur with increasing frequency in progressively later stages of pancreatic adenocarcinoma (5). Iacobuzio-Donahue et al demonstrated that tumor antigen p97, cathepsin L2 and kallikrein 10 were differentially expressed among PDACs (6). Additionally, the phosphoinositide 3-kinase signaling pathway is known to be activated in pancreatic cancer, which is due to the aberrant expression of phosphatase and tensin homolog. Progress achieved in understanding the mechanism of PDAC is likely to contribute to the treatment of this disease. However, no breakthrough treatments have been identified, so the present knowledge would appear to be insufficient.

In the present study, we downloaded microarray data of GSE43795 and identified the differentially expressed genes (DEGs) between PDAC and non-neoplastic pancreatic tissue (NN) samples to explore the molecular mechanisms of PDAC. Park et al (7) used the dataset GSE43795 to study the characterization of gene expression and activated signaling pathways in solid pseudopapillary neoplasms of the pancreas. However, the functional annotation and protein-protein interaction (PPI) of DEGs are still far from being clear. In the present study, we performed functional enrichment analyses and functional annotation. Finally, PPI networks and sub-networks were constructed and analyzed to study and identify target genes for the diagnosis and treatment of PDAC. We aimed to explore the underlying genes and pathways associated with PDAC. The findings from this study are likely to play a significant role in improving the treatment of PDAC.
role in PDAC genesis and may potentially serve as biomarkers in the diagnosis and treatment of PDAC.

Materials and methods

**Affymetrix microarray data.** The microarray data of GSE43795 were downloaded from the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) database based on the platform of GPL10558 Illumina HumanHT-12 V4.0 expression beadchip. A total of six PDAC and five NN samples were used in this study to develop the Affymetrix microarray data (Affymetrix, Inc., Santa Clara, CA, USA).

**Data pre-processing and differential expression analysis.**

Background correction, quartile data normalization and probe summarization were performed for the original array data, then they were converted into expression measures by the robust multiarray average (8) algorithm in the R affy package (9) (http://www.bioconductor.org).

For the GSE43795 dataset, the limma eBayes (10) method in Bioconductor (http://www.bioconductor.org) was used to identify genes which were differentially expressed between PDAC and NN samples. The log₂-fold change (log₂FC) was calculated. llog₂FC ≥3 and false discovery rate (FDR) <0.01 were considered as the cutoff values for DEG screening.

**Gene ontology and pathway enrichment analyses.** Gene ontology (GO) (11) is a tool used for collecting a large number of gene annotation terms. The Kyoto Encyclopedia of Genes and Genomes (KEGG) knowledge database (12) is a collection of online databases dealing with genomes, enzymatic pathways and biological chemicals. The Database for Annotation, Visualization and Integrated Discovery (DAVID) (13), as a comprehensive set of functional annotation tools, has been developed for relating the functional terms with gene lists using a clustering algorithm. In order to analyze the DEGs at the functional level, we performed GO and KEGG pathway enrichment analyses using the DAVID online tool to obtain the enriched biological processes (BPs) and pathways. P<0.01 was set as the threshold value.

**Functional annotation of DEGs.** Based on the data information of transcription factors (TFs), the DEGs were selected and annotated to determine whether these genes had the function of transcriptional regulation. The tumor suppressor gene database (TSGene) (14) integrates TSGs with large-scale experimental evidence to provide a comprehensive resource for the further investigation of TSGs and their molecular mechanisms in cancer. The tumor-associated gene (TAG) database (15) is used to save new genes that play a role in carcinogenesis. In this study, we extracted all known oncogenes and TSGs from the TAG and TSG databases.

**PPI network construction and sub-network identification.** The Search Tool for the Retrieval of Interacting Genes (STRING) database (16) is a precomputed global resource which was designed to evaluate PPI information. In the PPI network, each node stands for a gene and the edges represent the interactions between nodes. The degree indicates the number of edges linked to a given node and the nodes with a high degree are defined as the hub genes that possess essential biological functions. In this study, the STRING online tool was applied to analyze the PPI network of DEGs and only those experimentally validated interactions with a combined score >0.9 were selected as significant.

The BioNet package (http://www.bioconductor.org/packages/release/bioc/html/BioNet.html) (17) provides a comprehensive set of methods for the integrated analysis of gene expression data and biological networks. In the current study, we used BioNet to identify the sub-network in the PPI network with FDR <1.0E-06.

Based on the DEGs in the sub-network, we performed KEGG pathway enrichment analysis with P<0.01.

Results

**Identification of DEGs.** For the dataset GSE43795, a total of 979 transcripts were obtained following data pre-processing. Among them, 393 were downregulated transcripts corresponding to 374 DEGs and 586 were upregulated transcripts which corresponded to 559 DEGs.

**GO and pathway enrichment analyses.** Following GO and pathway analyses for down- and upregulated DEGs, the top five GO BP terms were collected and are shown in Table I. The GO terms enriched by downregulated DEGs were mainly related to digestion, transport and signaling while the GO terms enriched by upregulated DEGs were mainly associated with the cell cycle and mitosis.

The pathways enriched by downregulated DEGs were mainly related to digestion, absorption and metabolism, and included protein digestion and absorption and metabolism of xenobiotics by cytochrome P450. The pathways enriched by upregulated DEGs were mainly related to the occurrence and spread of cancer, including the cell cycle and p53 signaling pathways (Table II).

**Functional annotation of DEGs.** After researching the expression of TFs and TAGs, 15 TFs were downregulated, including prospero homeobox 1 and PBX/knotted 1 homeobox 2, and 19 TFs were upregulated, including vitamin D receptor and upstream transcription factor 1. In addition, among the downregulated DEGs, 22 genes were TAGs. Of these, 3 were oncogenes, 16 were TSGs and the effect of other 3 genes was uncertain. In the upregulated DEGs, 12 were oncogenes, including pituitary tumor-transforming 1 (PTTG1), 32 were TSGs and the effect of other 11 genes was uncertain (Table III).

**PPI network construction and sub-network pathway enrichment analysis.** Based on data from the STRING database, the PPI network was constructed (Fig. 1). Ten nodes were selected as hub genes (degree ≥12), including cell division cycle 20 (CDC20, degree=18) and BUB1 mitotic checkpoint serine/threonine kinase B (BUB1B, degree=16).

Using the BioNet package, sub-networks with 34 nodes were detected. Fig. 2 shows that CDC20 was a hub gene with degree=16.

Following KEGG pathway enrichment analyses of the DEGs in the sub-network, we observed that these DEGs including cyclin B1 (CCNB1) were mainly enriched in the cell cycle, p53 signaling and oocyte meiosis pathways.
Discussion

The identification of DEGs in PDAC is critical to the development of novel strategies to detect and treat this highly malignant cancer. In the present study, a total of 933 DEGs were identified between PDAC and NN samples through gene expression profiling of GSE43795. The downregulated DEGs were mainly enriched in the BP terms associated with digestion, transport and signaling, and pathways associated with digestion, absorption and metabolism. The upregulated DEGs were enriched in the biological processes associated with nuclear division, mitosis, and cell division. These findings provide insights into potential therapeutic targets for PDAC.
were mainly enriched in the BP terms associated with cell cycle and mitosis, and in the cell cycle and p53 signaling pathways. The oncogene \textit{PTTG1} was upregulated following functional annotation. In the PPI network, the hub genes \textit{CDC20} and \textit{BUB1B} had higher connectivity degrees. Additionally, \textit{CCNB1}, \textit{CDC20} and \textit{BUB1B} were enriched in several pathways in the sub-networks. This result suggested that these genes and pathways may play significant roles in the progression of PDAC. Cancer is characterized by uncontrolled cell proliferation, and tumor cells have typically acquired damage to genes
that directly regulate their cell cycles (18). Research has revealed that the loss of appropriate cell cycle regulation leads to genomic instability (19). The cell cycle is believed to play a role in the etiology of spontaneous cancers (20). In the present study, the cell cycle pathway was observed to be enriched by several upregulated DEGs, including CDC20 and BUB1B, which were also hub genes in the PPI network and sub-network. CDC20 is one of the regulators of spindle checkpoints, and appears to act as a regulatory protein interacting with several other proteins at multiple points in the cell cycle (21). In mammals, CDC20 is involved in anaphase onset and late mitotic events (22) and its expression is essential for cell division (23). At present, CDC20 is frequently reported to be upregulated in numerous types of malignancies including pancreatic cancer (23). Chang et al suggested that CDC20 expression may play a role in facilitating PDAC cell mitosis. For the other DEG, BUB1B, its encoded protein is also a key protein in the mitotic spindle checkpoint (24). It has been reported that overexpression of spindle assembly checkpoint molecules may result in DNA aneuploidy and carcinogenesis in mice (25). The high expression of BUB1B has often been reported to be associated with chromosomal instability in several malignancies, including kidney carcinomas, breast cancer and bladder cancer (24,26,27). Our findings also revealed that CDC20 and BUB1B expression and cell cycle pathways in which the two DEGs participated may play key roles in PDAC tumorigenesis and progression and may thus be useful as therapeutic targets.

In this study, the p53 signaling pathway was also observed to be enriched by upregulated DEGs including CCNB1. The p53 protein inhibits malignant transformation through direct and indirect regulation of transcription of the genes associated with the cell cycle and apoptosis (21). Presently, TP53 is the most frequently mutated gene in human cancer and its mutation frequency is up to 96% in pancreatic adenocarcinoma (28,29). CCNB1 is a regulatory protein involved in mitosis, and plays an essential role in cell proliferation (30). In normal tissues, the expression level of CCNB1 is low; however, it was noted to be overexpressed in tumors with TP53 mutation, including colorectal, cervical and pancreatic cancer (31-33). In addition, TP53 has been demonstrated to regulate the promoter of CCNB1 in opposing ways (34). Briefly, the p53 signaling pathway and its enriched DEGs including CCNB1 were closely associated with PDAC; thus, this pathway and these genes may be used as potential targets for PDAC treatment.

The results of functional annotation of DEGs revealed that the PTTG1 oncogene was upregulated; in addition, it was noted to participate in the PPI network and sub-network. Our findings also revealed that PTTG1 was enriched in the BP terms associated with mitosis. PTTG1 is a regulatory protein, and plays a central role in chromosome stability, cell transformation and gene regulation (35,36). In particular, PTTG1 is a critical mitotic checkpoint protein that helps hold sister chromatids together before entering anaphase (37). Research has identified that PTTG1 expression is highly activated in rapidly proliferating cells (38). To date, the overexpression of PTTG1 has been identified in numerous cancers tissues as well as in cell lines, including colon, ovarian, breast and various other solid tumors (39-41). In our study, the overexpression of PTTG1 is consistent with the observations above. Taken

Figure 2. Sub-network of differentially expressed genes (DEGs). Color indicates log₂-fold change (from low to high: green, pink, red). Green nodes represent downregulated DEGs; red nodes represent upregulated DEGs. Circles indicate that the node is significant in the network while squares indicate less significance.
together, these data support the hypothesis that PTTG1 may be a candidate molecular marker associated with PDAC progression and prognosis.

Although bioinformatics technologies have the potential to identify and validate candidate agents for critical diseases, certain limitations remain in this study. Firstly, the sample size for microarray analysis was small, which may have caused a high rate of false positive results. Secondly, this study lacked experimental verification. Further genetic and experimental studies with a larger sample size are still required in the future to confirm the results. However, our data provide a comprehensive bioinformatics analysis of the DEGs and pathways which may be involved in PDAC. The findings of the present study may contribute to our understanding of the underlying molecular mechanisms of PDAC. DEGs including CDC20, BUB1B, CCNB1 and PTTG1 as well as the cell cycle pathway and p53 signaling pathway have the potential to be used as targets for PDAC diagnosis and treatment.

Acknowledgements

The authors wish to express thanks to Fenghe (Shanghai) Information Technology Co., Ltd. Their ideas and assistance provided a valuable added dimension to this research.

References

1. Aguirre AJ, Bardeesy N, Sinha M, Lopez L, Tuveson DA, Horner J, Redston MS and DePinho RA: Activated Kras and Ink4/Arf deficiency cooperate to produce metastatic pancreatic ductal adenocarcinoma. Genes Dev 17: 3102-3126, 2003.
2. Garcia M, Jemal A, Ward EM, Tomatis L, Mathers C, Foreman K, Ferlay J, Bray F, Parkin DM, et al: Cancer incidence and mortality patterns in the world: Estimates for the year 2008. Int J Cancer 121: 199-266, 2007.
3. Conlon KC, Klimstra DS and Brennan MF: Long-term survival after curative resection for pancreatic ductal adenocarcinoma. Clinico-pathological analysis of 5-year survivors. Ann Surg 223: 273-279, 1996.
4. Sarkar FH, Banerjee S and Li Y: Pancreatic cancer: pathogenesis, prevention and treatment. Toxicol Appl Pharmacol 224: 326-336, 2007.
5. Berthelemy P, Bouisson M, Escourrou J, Vayssie N, Rumeau JL and Pradayrol L: Identification of K-ras mutations in pancreatic juice in the early diagnosis of pancreatic cancer. Ann Intern Med 123: 188-191, 1995.
6. Iacobuzio-Donahue CA, Ashfaq R, Maitra A, Adsay NV, Shen-Ong GL, Berg K, Hollingsworth MA, Cameron JL, Yeo CJ, Kern SE, et al: Highly expressed genes in pancreatic ductal adenocarcinomas: a comprehensive characterization and comparison of the transcription profiles obtained from three major technologies. Cancer Res 63: 8614-8622, 2003.
7. Park M, Kim M, Hwang D, Park M, Kim WK, Kim SK, Shin J, Park ES, Yong CG, Paik VK and Kim H: Characterization of gene expression and activated signaling pathways in solid-pseudopapillary neoplasm of pancreas. Mod Pathol 20: 580-593, 2014.
8. Iriyazry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U and Speed TP: Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. Hum Mutat 28: 622-629, 2007.
9. Hingorani SR, Wang L, Multani AS, Combs C, Deramaudt TB, Horban RH, Rustgi AK, Chang S and Tuveson DA: Trp53R172H and KrasG12D cooperate to promote chromosomal instability and widely metastatic pancreatic adenocarcinoma in mice. Cancer Cell 7: 469-483, 2005.
10. Suzuki T, Urano T, Miki Y, Moriya T, Akahira J, Ishida T, Horie K, Inoue S and Sasano H: Nuclear cyclin B1 in human colorectal cancers. J Cancer Res Clin Oncol 123: 124-127, 1997.
11. El-Ghobashy AA, Shaaban AM, Herod J, et al: Overexpression of cyclins A and B as markers of neoplastic glandular lesions of the cervix. Gynecol Oncol 92: 628-634, 2004.
12. Wang A, Yoshimi N, Iino N, Tanaka T and Morii H: Overexpression of cyclin B1 in human colorectal cancers. J Cancer Res Clin Oncol 123: 124-127, 1997.
13. El-Ghobashy AA, Shaaban AM, Herod J, et al: Overexpression of cyclins A and B as markers of neoplastic glandular lesions of the cervix. Gynecol Oncol 92: 628-634, 2004.
14. Weissmueller S, Mancho E, Saborowski M, et al: Mutant p53 drives pancreatic cancer metastasis through cell-autonomous PDGF receptor β signaling. Cell 157: 382-394, 2014.
15. Yin XY, Grove L, Datta NS, Katula K, Long MW and Prochownik EV: Inverse regulation of cyclin B1 by c-Myc and p53 and induction of tetraploidy by cyclin B1 overexpression. Cancer Res 61: 6487-6493, 2001.
16. Wang X, Horwitz GA, Ponziani TR, et al: Structure, expression and function of human pituitary tumor-transforming gene (PTTG). Mol Endocrinol 13: 156-166, 1999.
36. Pei L and Melmed S: Isolation and characterization of a pituitary tumor-transforming gene (PTTG). Mol Endocrinol 11: 433-441, 1997.
37. Ying H, Furuya F, Zhao L, Araki O, West BL, Hanover JA, Willingham MC and Cheng SY: Aberrant accumulation of PTTG1 induced by a mutated thyroid hormone β receptor inhibits mitotic progression. J Clin Invest 116: 2972-2984, 2006.
38. Genkai N, Homma J, Sano M, Tanaka R and Yamanaka R: Increased expression of pituitary tumor-transforming gene (PTTG)-1 is correlated with poor prognosis in glioma patients. Oncol Rep 15: 1569-1574, 2006.
39. Heaney AP, Singson R, McCabe CJ, Nelson V, Nakashima M and Melmed S: Expression of pituitary-tumour transforming gene in colorectal tumours. Lancet 355: 716-719, 2000.
40. Solbach C, Roller M, Fellbaum C, Nicoletti M and Kaufmann M: PTTG mRNA expression in primary breast cancer: a prognostic marker for lymph node invasion and tumor recurrence. Breast 13: 80-81, 2004.
41. Puri R, Tousson A, Chen L and Kakar SS: Molecular cloning of pituitary tumor transforming gene 1 from ovarian tumors and its expression in tumors. Cancer Lett 163: 131-139, 2001.