

Identify key genes and pathways in pancreatic ductal adenocarcinoma via bioinformatics analysis

Huatian Luo

Fujian Medical University <https://orcid.org/0000-0002-6505-825X>

Da-qiu Chen

Fujian Medical University

Jing-jing Pan

Fujian Medical University

Zhang-wei Wu

Fujian Medical University

Can Yang

Fujian Medical University

Hao-jie Xu

Fujian Medical University

Shang-hua Xu

Fujian Medical University

Hong-da Cai

Fujian Medical University

Shi Chen (✉ wawljwalj@163.com)

Fujian Provincial Hospital

Yao-dong Wang (✉ bill1054491834@fjmu.edu.cn)

Fujian Medical University

Research article

Keywords: Bioinformatical analysis, Microarray, Pancreatic cancer, Differentially expressed gene

Posted Date: April 7th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-19826/v1>

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

Abstract

Background: Pancreatic cancer has many pathologic types, among which pancreatic ductal adenocarcinoma (PDAC) is the most common one. Bioinformatics has become a very common tool for the selection of potentially pathogenic genes.

Methods: Three data sets containing the gene expression profiles of PDAC were downloaded from the gene expression omnibus (GEO) database. The limma package of R language was utilized to explore the differentially expressed genes (DEGs). To analyze functions and signaling pathways, the Database Visualization and Integrated Discovery (DAVID) was used. To visualize the protein-protein interaction (PPI) of the DEGs, Cytoscape was performed under the utilization of Search Tool for the Retrieval of Interacting Genes (STRING). With the usage of the plug-in cytoHubba in cytoscape software, the hub genes were found out. To verify the expression levels of hub genes, Gene Expression Profiling Interactive Analysis (GEPIA) was performed. Last but not least, UALCAN analysis online tool was implemented to analyze the overall survival.

Results: The 376 DEGs were highly enriched in biological processes including signal transduction, apoptotic process and several pathways, mainly associated with Protein digestion and absorption and Pancreatic secretion pathway. The expression levels of nucleolar and spindle associated protein 1 (NUSAP1) and SHC binding and spindle associated 1 (SHCBP1) were discovered highly expressed in pancreatic ductal adenocarcinoma tissues. NUSAP1 and SHCBP1 had a high correlation with prognosis.

Conclusions: The findings of this bioinformatics analysis indicate that NUSAP1 and SHCBP1 may be key factors in the prognosis and treatment of pancreatic cancer.

1. Background

Pancreatic cancer is famous for its high malignancy, rapid development and poor prognosis. The five-year survival rate of pancreatic cancer is no more than 6%.¹. It has many pathologic types, among which pancreatic ductal adenocarcinoma (PDAC) is the most common one and has a 5-year survival rate less than 3%.^{2,3}.

Even to this day, no treatment other than surgical excision such as chemotherapy or radiation has had a significant effect on pancreatic cancer. The specific mechanism of pancreatic cancer's strong ability to proliferate and invade has not been explored. Pancreatic cancer can easily invade and metastasize to surrounding tissues through blood vessels and lymph nodes.⁴. Many patients who have been diagnosed with pancreatic cancer have lost the opportunity for surgery because it is usually advanced.

Up to now, we still need more biomarkers which have high sensitivity and specificity rather than CA19-9, the only biomarker approved by the Food and Drug Administration. CA19-9 has many limitations, especially specificity and effectiveness.⁵. Differentially expressed genes play an important role in the proliferation, invasion and progression of pancreatic cancer, including *p16*, *SMAD4*, *K-ras* and *p53*.⁶. But they still couldn't fully explain how pancreatic cancer develops.

Bioinformatics has become a very common tool for the selection of potentially pathogenic genes. Many researchers have used bioinformatics to identify potential oncogenes. Despite the lack of in vivo or in vitro analysis, it is still a clever way to find potential oncogenes. For example, Look for potential oncogenes in ovarian cancer and pancreatic cancer.⁷⁻⁹.

Microarray analysis has been a very mature technology and has been widely used to screen differentially expressed genes in cancer. This present study is set out to investigate the DEGs between pancreatic ductal adenocarcinoma tissues and normal pancreatic tissues by determining gene expression profiles from GEO database. Besides, the functional enrichment and interaction network analyses of the DEGs were verified to help us explore the hub genes and the underlying molecular mechanisms of pancreatic ductal adenocarcinoma.

In result, 376 DEGs and 30 hub genes were identified, which may be potential biomarkers for pancreatic ductal adenocarcinoma. Among these genes, *SHCBP1* and *NUSAP1* were worthy of further experimental verification of their mechanism in pancreatic tumors and were likely to be a key factor in the prognosis and treatment of pancreatic cancer

2. Methods

2.1 Affymetrix Microarray Data

National Center for Biotechnology Information Search database created a gene expression database (<http://www.ncbi.nlm.nih.gov/geo>), which preserves the high-throughput gene expression data provided by research institutions around the world.

Utilizing the keywords "pancreatic ductal adenocarcinoma", we searched the GEO database. Three GEO datasets were downloaded, including GSE46234 contributed by Raeder H, GSE46385 contributed by Walters DM, and GSE71989 contributed by Schmittgen T. These gene expression profiles of pancreatic ductal adenocarcinoma were downloaded based on GPL570 Affymetrix Human Genome U133 Plus 2.0 Array. The GSE46234 dataset included 4 pancreatic ductal adenocarcinoma tissues and 4 normal pancreatic tissues. The GSE46385 consisted of 9 pancreatic ductal adenocarcinoma tissues and 3 normal pancreatic tissues. GSE71989 contained 13 pancreatic ductal adenocarcinoma tissues samples and 8 normal pancreatic tissues. We presented the detailed information about the datasets in Table 1.

2.2 Screening for DEGs

By using Affy package of R language, we manipulated the raw data by adjusting background, normalizing the data and log-transforming the raw data values.¹⁰. Using Perl language software, we merge these GEO datasets together and batch normalized the data set by utilizing the sva package in R language. These DEGs in pancreatic ductal adenocarcinoma and normal pancreatic tissues were analyzed by utilizing the limma package. Log2-foldchange(log2FC) > 2 and a corrected *p* value < 0.05 were set as the cut-off criteria of DEGs samples.

2.3 GO and KEGG Pathway Enrichment Analyses

We utilized the DAVID(version 6.8) online database to explore the Biological functions of the DEGs obtained from the integration of microarray data .¹¹.In order to verify the enrichment signaling pathways of DEGs, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and Gene Ontology (GO) enrichment analyses were performed with DAVID.

2.4 PPI Network Integration, Modules Analysis, and Selection of Hub Genes

We utilized the STRING (version11) online database to build a PPI network,in order to identify the interaction between protein-protein interaction(PPI) .¹². We set the highest confidence of the argument of interactions at > 0.7. To visualize and to analyze the PPI network, we used the Cytoscape (version 3.6.1) software. The plug-in cytoHubba was utilized for exploring key nodes and fragile motifs in the PPI network.¹³.

2.5 RNA sequencing expression and survival analysis and of core genes

UALCAN is a interactive web resource for analyzing cancer OMICS data (TCGA and MET500). It has high quality graphics using java script and CSS.¹⁴. To identify these DEGs, we applied the GEPIA website to analyze the data of RNA sequencing expression on the basis of thousands of samples from TCGA and Genotype-Tissue Expression(GTEx).¹⁵.

3. Results

3.1 DEGs in pancreatic ductal adenocarcinoma

In whole study, there were 26 pancreatic ductal adenocarcinoma tissues and 15 normal pancreatic tissues. 376 DEGs were identified, including 191 downregulated and 185 upregulated genes. We show the DEGs in Figs. 1, known as the volcano plots of DEGs. We present the Heatmap in Fig. 2. The gene names of them are exhibited in table 2.

3.2 GO and KEGG Analysis of DEGs.

As shown in Table 3 and Fig. 3,all 376 DEGs were analyzed by DAVID and the results of GO analysis revealed for biological processes (BP),DEGs were highly enriched in signal transduction, apoptotic process, proteolysis, cell adhesion and cell-cell signaling; for molecular function (MF),DEGs were mainly enriched in identical protein binding, calcium ion binding, protein homodimerization activity, serine-type endopeptidase activity, catalytic activity and receptor binding; for GO cell component (CC), DEGs were significantly enriched in extracellular exosome, extracellular space, extracellular region, integral component of plasma membrane and cell surface.

We presented the KEGG analysis results in Table 4, which indicated that DEGs were particularly enriched in Protein digestion and absorption, Pancreatic secretion, ECM-receptor interaction and some other pathways. (*P* < 0.05).

3.3 Protein–protein interaction network (PPI) and modular analysis

In order to explore the biological characteristics of these DEGs, we created a PPI network with the usage of the STRING database. We presented the result in Fig. 4. Meanwhile we utilized the cytoHubba plug-in to screen hub genes and the result is shown in Fig. 5 and table 5.

3.4 Survival Analysis

By comparing all the literature on pancreatic cancer to date, we found two genes that had not been explored and studied in pancreatic cancer especially pancreatic ductal adenocarcinoma.

In order to validate the expressions of these two hub genes, we utilized the GEPIA website to analyze the data of RNA sequencing expression on the basis of 350 samples (Fig. 6.Aand Fig. 6.B). Moreover, UALCAN website was used to analyze the hub genes survival in the 177 samples which is derived from the TCGA project (Fig. 6.Cand Fig. 6.D).

The result indicated the expressions of *NUSAP1* and *SHCBP1* were closely related to the survivals in pancreatic adenocarcinoma.

4. Discussion

In the past few decades, many causes and potential mechanisms for the formation and development of pancreatic adenocarcinoma have been identified by a lot of experts.

However, there was no significant improvement in 5-year survival. Moreover, the identification of reliable molecular markers with high prognostic value has not been found in pancreatic ductal adenocarcinoma. Therefore, we need to explore and discover a specific biomarker to detect early pancreatic adenocarcinoma, which is conducive to the treatment and survival of patients. Last but not least, we need to discover and validate new molecular targets to develop drugs that can effectively treat pancreatic cancer. Many studies concentrate on an independent genetic event, or the result is generated from independent studies which are inconsistent with each other by microarray analysis.¹⁶⁻¹⁸

Some GEO samples (GSM) of pancreatic adenocarcinoma were not grouped into pancreatic ductal adenocarcinoma, the result may lose some details or may even be wrong for the mixed disease subtypes. Pancreatic adenocarcinoma pathogenesis in different patients may depend on common changes of the expression of particular critical genes, and rather on personal particular changes of different genes. Comparing pancreatic adenocarcinoma instead of pancreatic ductal adenocarcinoma tissues with normal tissues may miss some important genes. Our study was done by comparing pancreatic ductal adenocarcinoma tissues with normal pancreatic tissues which may find some potential genes had not found in previous research. Our researcher found, in agreement with previous studies, most of the hub genes plays an important role in the development and progression of pancreatic ductal adenocarcinoma and they were reported and explored by previous research. That proves the reliability of our research in some ways.

TOP2A could activate EMT process which was highly associated with tumor metastasis in pancreatic cancer patients and could enhance tumor progression.¹⁹ PHGDH could down regulate the expression of *CCNB1*, *CCND1*, *MMP-2*, and *MMP-9* in pancreatic cancer cells and inhibited the cell invasion, migration, and proliferation.²⁰ The down-regulated expression of *CDC20* in pancreatic cancer cells could do the same work. Meanwhile it induced cell apoptosis in pancreatic cancer cells.²¹ The knock-down of *NDC80* in Panc-1 cells could inhibit cell proliferation and colony formation. Moreover, it infected Panc-1 cells resulted in induction of apoptosis and cell cycle arrest at G2/M phase.²² The interference suppression of *KIF20A* resulted in an inhibition of motility, invasion and proliferation of pancreatic cancer cell lines.²³ The upregulated *PBK/TOPK* in pancreatic cancer and upregulated expression levels was highly correlated with the invasive property of pancreatic cancer cells. By regulating *PBK/TOPK* pathway, the invasion ability of PDAC cells can be regulated.²⁴ The knock-down of *IMP3* in pancreatic cancer reduced PDAC cell viability, extracellular matrix adhesion and invasion. *IMP3*-depleted cells presented lower levels of *CD44* protein and *KIF11* mRNA.²⁵ Overexpression of *BUB1B*, *CDK1*, *CCNA2* and *CDC20* in pancreatic cancer was highly associated with disease-free survival in PDAC patients.²⁶

As for the genes we found, Nucleolar and Spindle Associated Protein 1 (*NUSAP1*) and SHC binding and spindle associated 1 (*SHCBP1*) were both hub genes and interact with the hub genes which we discussed above, which suggested that *NUSAP1* and *SHCBP1* may also play a very important role in pancreatic cancer.

We found that *NUSAP1* had been reported the expression of *NUSAP1* is up-regulated in non-small cell lung cancer, which is associated with the growth and development of non-small cell lung cancer and prognosis of the patients.²⁷ In non-small cell lung cancer, the inhibited expression of *NUSAP1* could result in the blocking the cell cycle on G1 phase.²⁸ In addition, *NUSAP1* Inhibited Cell Proliferation in Invasive Breast Cancer Cells.²⁹ Last but not least, *NUSAP1* was associated with HCC onset, progression, and prognosis, and exhibited higher expression in HCC compared with normal livers.³⁰ In all, *NUSAP1* was associated with different types of tumors. Until now, there is no such study in pancreatic cancer. The GEPIA website revealed that *NUSAP1* has a clear trend of increasing gene expression levels in Pancreatic adenocarcinoma compared to normal samples. The report was on the basis of 350 samples. Meanwhile, UALCAN website revealed that *NUSAP1* had a significantly worse survival. Above all, *NUSAP1* may play an important role in pancreatic cancer.

Another gene we found in this study was *SHCBP1*. In gastric cancer, *SHCBP1* suppressed apoptotic and promoted cell cycle progression which accelerated tumor growth and invasiveness.³¹ Meanwhile, *SHCBP1* accelerated cisplatin induced invasion, migration and apoptosis resistance by activating Wnt pathway in lung cancer.³² Besides, the interference suppression of *SHCBP1* could significantly promote the apoptosis of lung cancer cells.³³ In non-small cell lung carcinoma, the release of *SHCBP1* from SHC adaptor protein 1 (*SHC1*) could result in development of non-small cell lung carcinoma malignant progression and cell stemness.³⁴ In gliomas, the overexpression of *SHCBP1* promotes migration and invasion by activating the *NF-KB* signaling pathway.³⁵ Last but not least, the overexpression of *SHCBP1* could significantly promote HCC cell proliferation, survival and colony formation in HCC cell lines and the knockdown of *SHCBP1* resulted in cell cycle delay and suppressed cell proliferation in human hepatocellular carcinoma cells.³⁶ The GEPIA website reported that *SHCBP1* has a clear trend of increasing gene expression levels in Pancreatic adenocarcinoma compared to normal samples. The report was on the basis of 350 samples. Meanwhile, UALCAN website revealed that *SHCBP1* had a significantly worse survival. Above all, *SHCBP1* may play an important role in pancreatic cancer.

5. Conclusions

All in all, our data analysis study found *SHCBP1* and *NUSAP1* which has not been published in pancreatic ductal adenocarcinoma. They were identified on the base of mixed and normalized three different microarray datasets. Just as other researchers have used bioinformatics to explore potential oncogenes, we demonstrate the value of these two molecules for further exploration and verification from different perspectives. Results revealed that *NUSAP1* and *SHCBP1* were worthy of further experimental verification of their mechanism in pancreatic tumors and are likely to be a key factor in the prognosis and treatment of pancreatic cancer especially pancreatic ductal adenocarcinoma.

Abbreviations

Abbreviations	Full Name
PDAC	pancreatic ductal adenocarcinoma
GEO	the gene expression omnibus
DEGs	the differentially expressed genes
DAVID	the Database Visualization and Integrated Discovery
STRING	Search Tool for the Retrieval of Interacting Genes
GEPIA	Gene Expression Profiling Interactive Analysis
KEGG	Kyoto Encyclopedia of Genes and Genomes
GO	Gene Ontology
PPI	protein-protein interaction
GTE _x	Genotype-Tissue Expression
BP	biological processes
MF	molecular function
CC	cell component
GSM	GEO samples
HCC	hepatocellular carcinoma

Declarations

Ethics approval and consent to participate: Not applicable. Our manuscript does not report on or involve the use of any animal or human data or tissue, this section is not applicable to our submission.

Consent for publication: Individual person's data contained in our manuscript is shared by GEO which is opened to publication.

Availability of data and material: Cytoscape (version 3.6.1) software/ R language(version 3.6.1) is used in this manuscript

Competing interests: The authors declare that they have no competing interests.

Funding: This study was supported in part by a grant from National Natural Science Foundation of China (No.81772560 to Y.D. Wang and No.81502008 to S. Chen) and the Natural Science Foundation for Distinguished Young Scholars of Fujian Province (No.2018J06020 to S. Chen). Professor Yao-dong Wang and professor Shi Chen are the holder of these funding, their conception and design of experiments are key component of this study.

Authors' contributions: The authors would like to thank his tutors, professor YD and professor SC, for their conception and design of experiments. Moreover, they also drafted the work and revised it critically for important content. The authors also would like to thank DQ for statistical analysis. Meanwhile, the analysis tools was provided by JP and ZW. The figures and tables was prepared by CY. All authors have read and approved the manuscript.

Acknowledgements: The authors would like to thank his tutors, professor Yao-dong Wang and professor Shi Chen, for their conception and design of experiments. Moreover, they also drafted the work and revised it critically for important content. The authors also would like to thank Mr.Da-Qiu Chen for statistical analysis. Meanwhile, the analysis tools was provided by Ms.Jing-jing Pan and Mr. Zhang-wei Wu. The figures and tables was prepared by Mr. Can Yang. All authors have read and approved the manuscript.

7. References

- Kamisawa, T., Wood, L. D., Itoi, T. & Takaori, K. Pancreatic cancer. *Lancet* **388**, 73-85, doi:10.1016/s0140-6736(16)00141-0 (2016).
- Dodson, L. F., Hawkins, W. G. & Goedegebuure, P. Potential targets for pancreatic cancer immunotherapeutics. *Immunotherapy* **3**, 517-537, doi:10.2217/imt.11.10 (2011).
- Momeny, M. *et al.* Anti-tumor activity of cediranib, a pan-vascular endothelial growth factor receptor inhibitor, in pancreatic ductal adenocarcinoma cells. *Cell Oncol (Dordr)*, doi:10.1007/s13402-019-00473-9 (2019).
- Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).
- Crawley, A. S. & O'Kennedy, R. J. The need for effective pancreatic cancer detection and management: a biomarker-based strategy. *Expert Rev Mol Diagn* **15**, 1339-1353, doi:10.1586/14737159.2015.1083862 (2015).
- Wolfgang, C. L. *et al.* Recent progress in pancreatic cancer. *CA Cancer J Clin* **63**, 318-348, doi:10.3322/caac.21190 (2013).

- 7 Feng, H. *et al.* Identification of significant genes with poor prognosis in ovarian cancer via bioinformatical analysis. *J Ovarian Res* **12**, 35, doi:10.1186/s13048-019-0508-2 (2019).
- 8 Lu, W., Li, N. & Liao, F. Identification of Key Genes and Pathways in Pancreatic Cancer Gene Expression Profile by Integrative Analysis. *Genes (Basel)* **10**, doi:10.3390/genes10080612 (2019).
- 9 Chen, Q. *et al.* Screening and identification of hub genes in pancreatic cancer by integrated bioinformatics analysis. *J Cell Biochem* **120**, 19496-19508, doi:10.1002/jcb.29253 (2019).
- 10 Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. affy-analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307-315, doi:10.1093/bioinformatics/btg405 (2004).
- 11 Dennis, G. *et al.* DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biology* **4**, R60, doi:10.1186/gb-2003-4-9-r60 (2003).
- 12 von Mering, C. *et al.* STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res* **31**, 258-261, doi:10.1093/nar/gkg034 (2003).
- 13 Chin, C. H. *et al.* cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol* **8 Suppl 4**, S11, doi:10.1186/1752-0509-8-S4-S11 (2014).
- 14 Chandrashekar, D. S. *et al.* UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia* **19**, 649-658, doi:10.1016/j.neo.2017.05.002 (2017).
- 15 Tang, Z. *et al.* GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res* **45**, W98-W102, doi:10.1093/nar/gkx247 (2017).
- 16 Crnogorac-Jurcevic, T. *et al.* Molecular analysis of precursor lesions in familial pancreatic cancer. *PLoS One* **8**, e54830, doi:10.1371/journal.pone.0054830 (2013).
- 17 Zhang, G. *et al.* DPEP1 inhibits tumor cell invasiveness, enhances chemosensitivity and predicts clinical outcome in pancreatic ductal adenocarcinoma. *PLoS One* **7**, e31507, doi:10.1371/journal.pone.0031507 (2012).
- 18 Zhang, G. *et al.* Integration of metabolomics and transcriptomics revealed a fatty acid network exerting growth inhibitory effects in human pancreatic cancer. *Clin Cancer Res* **19**, 4983-4993, doi:10.1158/1078-0432.Ccr-13-0209 (2013).
- 19 Pei, Y. F., Yin, X. M. & Liu, X. Q. TOP2A induces malignant character of pancreatic cancer through activating beta-catenin signaling pathway. *Biochim Biophys Acta Mol Basis Dis* **1864**, 197-207, doi:10.1016/j.bbadis.2017.10.019 (2018).
- 20 Song, Z., Feng, C., Lu, Y., Lin, Y. & Dong, C. PHGDH is an independent prognosis marker and contributes cell proliferation, migration and invasion in human pancreatic cancer. *Gene* **642**, 43-50, doi:10.1016/j.gene.2017.11.014 (2018).
- 21 Guo, W., Zhong, K., Wei, H., Nie, C. & Yuan, Z. Long non-coding RNA SPRY4-IT1 promotes cell proliferation and invasion by regulation of Cdc20 in pancreatic cancer cells. *PLoS One* **13**, e0193483, doi:10.1371/journal.pone.0193483 (2018).
- 22 Meng, Q. C. *et al.* Overexpression of NDC80 is correlated with prognosis of pancreatic cancer and regulates cell proliferation. *Am J Cancer Res* **5**, 1730-1740 (2015).
- 23 Stangel, D. *et al.* Kif20a inhibition reduces migration and invasion of pancreatic cancer cells. *J Surg Res* **197**, 91-100, doi:10.1016/j.jss.2015.03.070 (2015).
- 24 Hinzman, C. P. *et al.* Aberrant expression of PDZ-binding kinase/T-LAK cell-originated protein kinase modulates the invasive ability of human pancreatic cancer cells via the stabilization of oncoprotein c-MYC. *Carcinogenesis* **39**, 1548-1559, doi:10.1093/carcin/bgy114 (2018).
- 25 Pasiliao, C. C. *et al.* The involvement of insulin-like growth factor 2 binding protein 3 (IMP3) in pancreatic cancer cell migration, invasion, and adhesion. *BMC Cancer* **15**, 266, doi:10.1186/s12885-015-1251-8 (2015).
- 26 Dong, S., Huang, F., Zhang, H. & Chen, Q. Overexpression of BUB1B, CCNA2, CDC20, and CDK1 in tumor tissues predicts poor survival in pancreatic ductal adenocarcinoma. *Biosci Rep* **39**, doi:10.1042/bsr20182306 (2019).
- 27 Yu, Z. *et al.* [Expression of NUSAP1 and its relationship with prognosis in non-small cell lung cancer]. *Zhonghua Zhong Liu Za Zhi* **41**, 522-526, doi:10.3760/cma.j.issn.0253-3766.2019.07.007 (2019).
- 28 Li, X. *et al.* Knockdown of SP1/Syncytin1 axis inhibits the proliferation and metastasis through the AKT and ERK1/2 signaling pathways in non-small cell lung cancer. *Cancer Med*, doi:10.1002/cam4.2448 (2019).
- 29 Zhang, X., Pan, Y., Fu, H. & Zhang, J. Nucleolar and Spindle Associated Protein 1 (NUSAP1) Inhibits Cell Proliferation and Enhances Susceptibility to Epirubicin In Invasive Breast Cancer Cells by Regulating Cyclin D Kinase (CDK1) and DLGAP5 Expression. *Med Sci Monit* **24**, 8553-8564,

doi:10.12659/msm.910364 (2018).

30 Shan, S., Chen, W. & Jia, J. D. Transcriptome Analysis Revealed a Highly Connected Gene Module Associated With Cirrhosis to Hepatocellular Carcinoma Development. *Front Genet* **10**, 305, doi:10.3389/fgene.2019.00305 (2019).

31 Dong, Y. D., Yuan, Y. L., Yu, H. B., Tian, G. J. & Li, D. Y. SHCBP1 is a novel target and exhibits tumorpromoting effects in gastric cancer. *Oncol Rep* **41**, 1649-1657, doi:10.3892/or.2018.6952 (2019).

32 Zou, A. *et al.* SHCBP1 promotes cisplatin induced apoptosis resistance, migration and invasion through activating Wnt pathway. *Life Sci* **235**, 116798, doi:10.1016/j.lfs.2019.116798 (2019).

33 Wang, F., Li, Y., Zhang, Z., Wang, J. & Wang, J. SHCBP1 regulates apoptosis in lung cancer cells through phosphatase and tensin homolog. *Oncol Lett* **18**, 1888-1894, doi:10.3892/ol.2019.10520 (2019).

34 Liu, L. *et al.* EGF-induced nuclear localization of SHCBP1 activates beta-catenin signaling and promotes cancer progression. *Oncogene* **38**, 747-764, doi:10.1038/s41388-018-0473-z (2019).

35 Zhou, Y. *et al.* Overexpression of SHCBP1 promotes migration and invasion in gliomas by activating the NF-kappaB signaling pathway. *Mol Carcinog* **57**, 1181-1190, doi:10.1002/mc.22834 (2018).

36 Tao, H. C. *et al.* Targeting SHCBP1 inhibits cell proliferation in human hepatocellular carcinoma cells. *Asian Pac J Cancer Prev* **14**, 5645-5650, doi:10.7314/apjcp.2013.14.10.5645 (2013).

Tables

Table I: Details for GEO data

Contributor	Organism	GEO	Platform	PDAC tissues	Normal pancreatic tissues
Raeder H	Homo sapiens	GSE46234	GPL570	4	4
Walters DM	Homo sapiens	GSE46385	GPL570	3	9
Schmittgen T.	Homo sapiens	GSE71989	GPL570	8	13

Abbreviation: GEO, Gene Expression Omnibus. PDAC, pancreatic ductal adenocarcinoma

Table II: Screening DEGs in PDAC by integrated microarray

DEGs	Gene names
Upregulated	GPX8 DHFR MBOAT1 TMSB10 ITGA2 AHR LGALS3BP TMEM154 MTD1 DDX6 EOGT LINC00342 S100A6 LOC100506119 S100A11 PCED1B OSBPL3 IL1RAP KIF20B MYOF IL18 GPCR5A NHS2 COL5A2 ANLN C15orf48 LIPG CXCL8 AKR1B10 CDC20 EGR3 CD52 TMEM45B COMP CXCR4 CDC42EP5 MMP12 IFI44L ID1 COL11A1 COL3A1 TNFRSF11B SLPI AOC1 SFTA2 TMPRSS4 COL1A
Downregulated	RAB4B C1orf167 HCG22 MAGEF1 COX6A2 RTN4RL2 PLXNB3 ZNF503-AS1 HGFAC INH CHGA LOC100129476 NRTN GDF10 PDIA2 CTNND2 TTY15 PRDM12 EPB41L4B LOC285095 PAIP2B P
	LOC100422737 TMED6 SLC3A1 GAS2 IMPA2 SST7 CPB1 SLC43A1 ACADL DPP10-AS1 GMNN REG3G ARSE LOC102723493 FAM3B PAK3 TRHDE TPST2 REG1A SERPINA4 MT1M P2RX1 REG3A

Abbreviation: DEGs, differentially expressed genes

Table III. Top 10 GO analysis of genes associated with pancreatic ductal adenocarcinoma

BP	GO:0007165	signal transduction	35	9.433962	0.011999	18.57046
BP	GO:0006915	apoptotic process	27	7.277628	5.27E-05	0.089673
BP	GO:0006508	proteolysis	26	7.008086	1.79E-05	0.030391
BP	GO:0007155	cell adhesion	21	5.660377	7.34E-04	1.242051
BP	GO:0007267	cell-cell signaling	19	5.121294	2.71E-06	0.004608
BP	GO:0042493	response to drug	19	5.121294	3.15E-05	0.053611
BP	GO:0006954	inflammatory response	18	4.851752	0.001307	2.201399
BP	GO:0043066	negative regulation of apoptotic process	18	4.851752	0.008352	13.30212
BP	GO:0030198	extracellular matrix organization	17	4.58221	1.57E-06	0.002673
BP	GO:0008284	positive regulation of cell proliferation	17	4.58221	0.021356	30.74535
CC	GO:0070062	extracellular exosome	96	25.87601	6.34E-09	8.39E-06
CC	GO:0005615	extracellular space	91	24.5283	6.77E-27	8.97E-24
CC	GO:0005576	extracellular region	88	23.71968	9.07E-20	1.20E-16
CC	GO:0005887	integral component of plasma membrane	36	9.703504	0.066999	60.0868
CC	GO:0009986	cell surface	22	5.929919	0.001741	2.28161
CC	GO:0005578	proteinaceous extracellular matrix	20	5.390836	9.94E-07	0.001316
CC	GO:0048471	perinuclear region of cytoplasm	19	5.121294	0.051997	50.69765
CC	GO:0005788	endoplasmic reticulum lumen	16	4.312668	4.26E-06	0.005648
CC	GO:0031012	extracellular matrix	16	4.312668	5.81E-04	0.767204
CC	GO:0005581	collagen trimer	11	2.96496	9.94E-06	0.013167
MF	GO:0042802	identical protein binding	27	7.277628	0.003165	4.540907
MF	GO:0005509	calcium ion binding	26	7.008086	0.003553	5.084766
MF	GO:0042803	protein homodimerization activity	24	6.469003	0.015913	20.95821
MF	GO:0004252	serine-type endopeptidase activity	19	5.121294	2.66E-06	0.003893
MF	GO:0003824	catalytic activity	13	3.504043	3.17E-04	0.463466
MF	GO:0005102	receptor binding	13	3.504043	0.043661	48.03253
MF	GO:0008083	growth factor activity	12	3.234501	3.32E-04	0.486339
MF	GO:0005179	hormone activity	11	2.96496	1.43E-05	0.020907
MF	GO:0004867	serine-type endopeptidase inhibitor activity	10	2.695418	1.13E-04	0.165216
MF	GO:0008201	heparin binding	10	2.695418	0.004112	5.863225

Note: BP, biological process; CC, cell component; MF, molecular function; GO, gene ontology;

Table IV. KEGG pathway enrichment analyses of DEGs associated with pancreatic ductal adenocarcinoma

Category	Description	Count	%	P Value	Genes
KEGG_PATHWAY	Protein digestion and absorption	18	4.851752	1.19E-11	<i>FXYP2, CELA3A, CELA3B, SLC16A10, COL3A1, SLC3A1, COL5A2, COL5A1, PRSS2, PRSS3, COL1A2, CPA2, CELA2B, CPA1, CPB1, COL11A1, CTRL, COL10A1</i>
KEGG_PATHWAY	Pancreatic secretion	16	4.312668	2.90E-09	<i>PNLIP, FXYD2, CELA3A, PNLIPRP1, CELA3B, PNLIPRP2, CEL, PRSS2, PRSS3, PLA2G1B, CPA2, CELA2B, CPA1, SLC4A4, CPB1, CTRL</i>
KEGG_PATHWAY	ECM-receptor interaction	9	2.425876	9.21E-04	<i>LAMB3, COMP, COL3A1, COL1A2, ITGA2, LAMC2, COL11A1, COL5A2, COL5A1</i>
KEGG_PATHWAY	Fat digestion and absorption	6	1.617251	0.001997	<i>PNLIP, CEL, CLPS, PNLIPRP1, PNLIPRP2, PLA2G1B</i>
KEGG_PATHWAY	Glycerolipid metabolism	7	1.886792	0.002217	<i>PNLIP, CEL, PNLIPRP1, PNLIPRP2, AKR1B10, MBOAT1, LIPG</i>
KEGG_PATHWAY	Maturity onset diabetes of the young	5	1.347709	0.002891	<i>IAPP, INS, NEUROD1, NR5A2, NKX2-2</i>
KEGG_PATHWAY	Amoebiasis	9	2.425876	0.003269	<i>GNA15, LAMB3, COL3A1, COL1A2, CXCL8, LAMC2, COL11A1, COL5A2, COL5A1</i>
KEGG_PATHWAY	Mineral absorption	6	1.617251	0.003429	<i>FXYD2, MT1M, MT1H, STEAP1, MT1G, MT1F</i>
KEGG_PATHWAY	Proximal tubule bicarbonate reclamation	4	1.078167	0.015803	<i>FXYD2, CA4, SLC4A4, AQP1</i>
KEGG_PATHWAY	Focal adhesion	11	2.96496	0.022393	<i>LAMB3, PAK3, COMP, COL3A1, COL1A2, ITGA2, LAMC2, EGF, COL11A1, COL5A2, COL5A1</i>
KEGG_PATHWAY	Legionellosis	5	1.347709	0.036982	<i>IL18, CXCL3, PYCARD, CXCL8, BNIP3</i>
KEGG_PATHWAY	Cytokine-cytokine receptor interaction	11	2.96496	0.058396	<i>INHBA, TNFRSF11B, IL22RA1, CXCL5, CCL20, CXCR4, IL18, CXCL3, IL1RAP, CXCL8, CCL18</i>
KEGG_PATHWAY	Bladder cancer	4	1.078167	0.070274	<i>TYMP, CXCL8, EGF, MMP1</i>
KEGG_PATHWAY	p53 signaling pathway	5	1.347709	0.071255	<i>CCNB1, RRM2, PMAIP1, SFN, IGFBP3</i>
KEGG_PATHWAY	Axon guidance	7	1.886792	0.076909	<i>PAK3, CXCR4, ROBO1, PLXNA2, EFNB2, PLXNB3, SEMA3C</i>
KEGG_PATHWAY	Bile secretion	5	1.347709	0.077599	<i>FXYD2, NCEH1, AQP8, SLC4A4, AQP1</i>
KEGG_PATHWAY	Complement and coagulation cascades	5	1.347709	0.077599	<i>CD55, SERPINA5, C6, C5, PLAU</i>
KEGG_PATHWAY	Platelet activation	7	1.886792	0.083922	<i>P2RX1, COL3A1, COL1A2, ITGA2, COL11A1, COL5A2, COL5A1</i>

Table V. Top 30 hub genes

Gene	Description
<i>CCNB1</i>	cyclin B1 [Homo sapiens (human)]
<i>CDC20</i>	cell division cycle 20 [Homo sapiens (human)]
<i>NDC80</i>	NDC80 kinetochore complex component [Homo sapiens (human)]
<i>TOP2A</i>	DNA topoisomerase II alpha [Homo sapiens (human)]
<i>KIF20A</i>	kinesin family member 20A [Homo sapiens (human)]
<i>PBK</i>	PDZ binding kinase [Homo sapiens (human)]
<i>RRM2</i>	ribonucleotide reductase regulatory subunit M2 [Homo sapiens (human)]
<i>KIF11</i>	kinesin family member 11 [Homo sapiens (human)]
<i>BUB1B</i>	BUB1 mitotic checkpoint serine/threonine kinase B [Homo sapiens (human)]
<i>TPX2</i>	TPX2 microtubule nucleation factor [Homo sapiens (human)]
<i>PRC1</i>	protein regulator of cytokinesis 1 [Homo sapiens (human)]
<i>NUSAP1</i>	nucleolar and spindle associated protein 1 [Homo sapiens (human)]
<i>DLGAP5</i>	DLG associated protein 5 [Homo sapiens (human)]
<i>RACGAP1</i>	Rac GTPase activating protein 1 [Homo sapiens (human)]
<i>CEP55</i>	centrosomal protein 55 [Homo sapiens (human)]
<i>DTL</i>	denticleless E3 ubiquitin protein ligase homolog [Homo sapiens (human)]
<i>RAD51AP1</i>	RAD51 associated protein 1 [Homo sapiens (human)]
<i>ZWINT</i>	ZW10 interacting kinetochore protein [Homo sapiens (human)]
<i>SHCBP1</i>	SHC binding and spindle associated 1 [Homo sapiens (human)]
<i>ANLN</i>	anillin actin binding protein [Homo sapiens (human)]
<i>CKS2</i>	CDC28 protein kinase regulatory subunit 2 [Homo sapiens (human)]
<i>UHRF1</i>	ubiquitin like with PHD and ring finger domains 1 [Homo sapiens (human)]
<i>PPY</i>	pancreatic polypeptide [Homo sapiens (human)]
<i>CXCL8</i>	C-X-C motif chemokine ligand 8 [Homo sapiens (human)]
<i>SST</i>	somatostatin [Homo sapiens (human)]
<i>ANXA1</i>	annexin A1 [Homo sapiens (human)]
<i>NMU</i>	neuromedin U [Homo sapiens (human)]
<i>CXCR4</i>	C-X-C motif chemokine receptor 4 [Homo sapiens (human)]
<i>C5</i>	complement C5 [Homo sapiens (human)]
<i>CCL20</i>	C-C motif chemokine ligand 20 [Homo sapiens (human)]

Figures

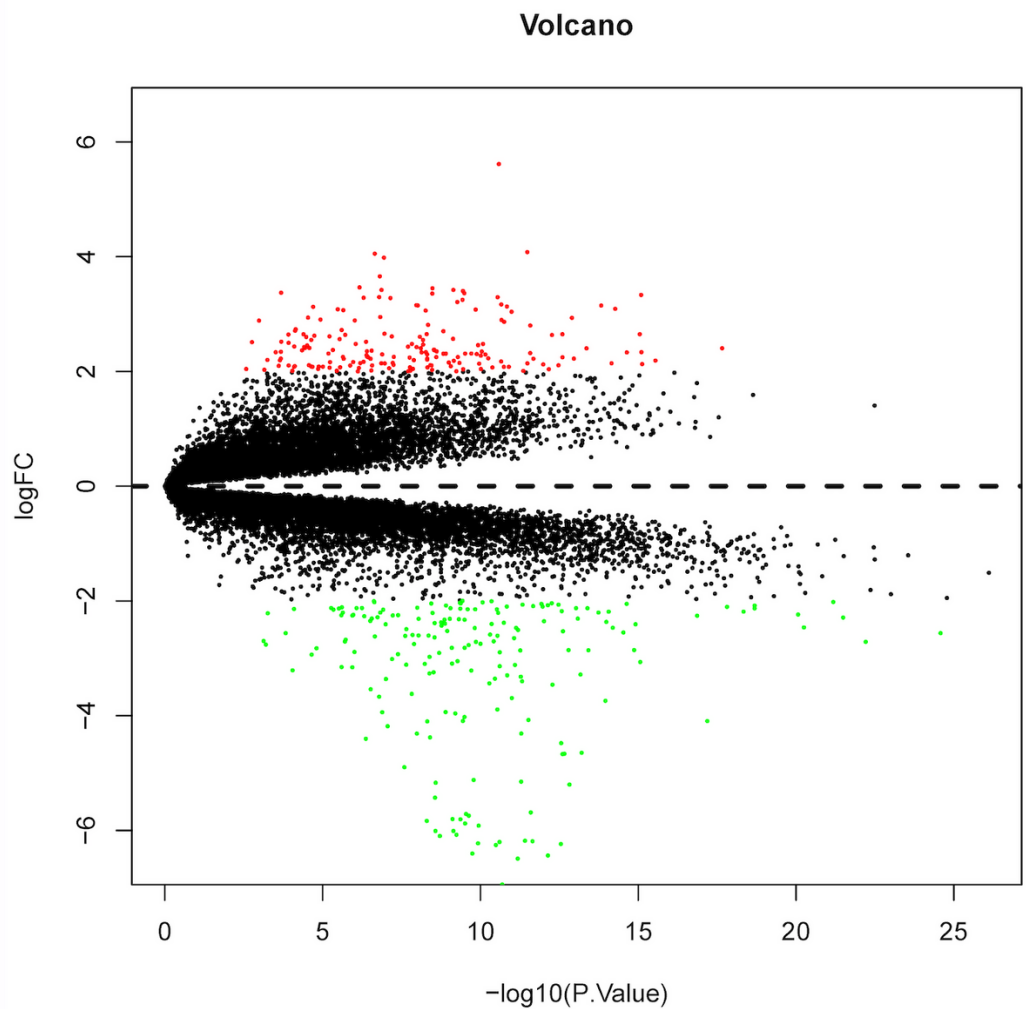


Figure 1

Volcano.Volcano plot of gene expression profile data in pancreatic ductal adenocarcinoma tissues and the normal pancreatic tissues. The red, green, and gray points represent upregulated genes, downregulated genes, and none differentially expressed genes, respectively.

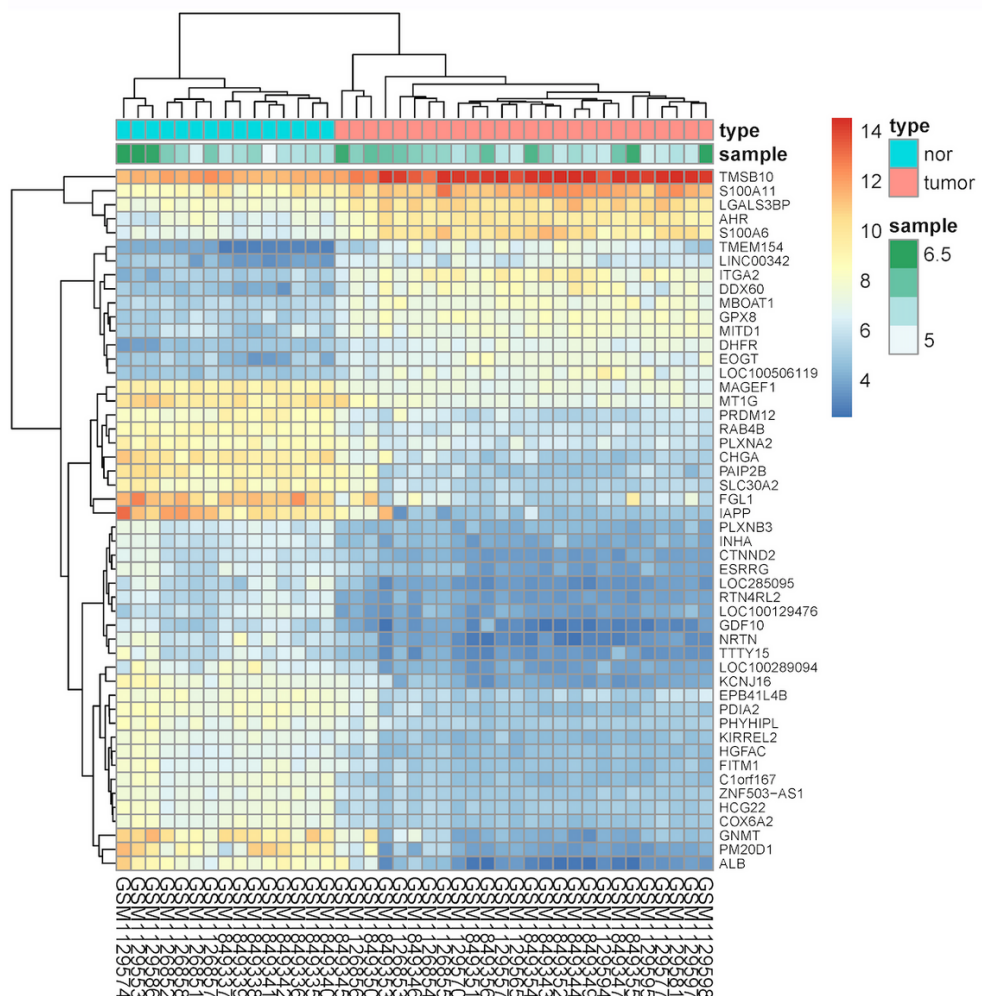


Figure 2

Heat map. Heat map showing top 50 up-regulated and down-regulated differentially expressed genes in pancreatic ductal adenocarcinoma tissues compared to the normal tissues. They are screened on the basis of $\log_2FC > 2$ and a corrected $p < 0.05$.

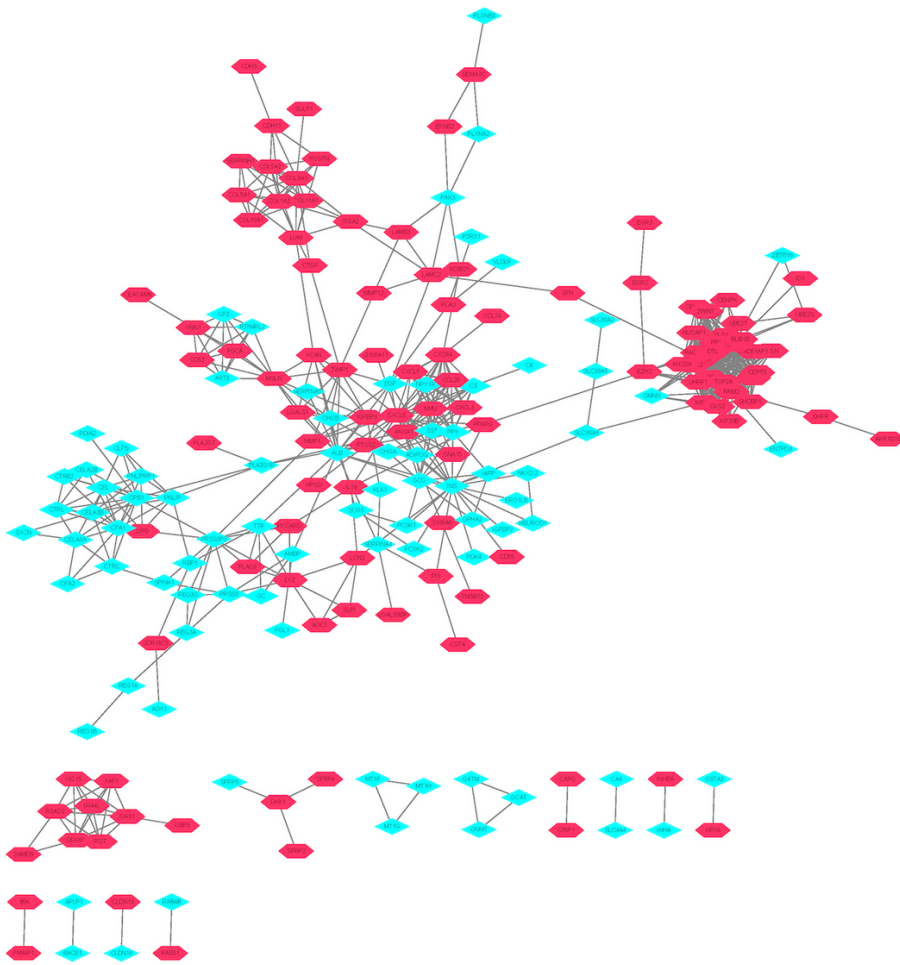


Figure 4

PPI network. Common DEGs PPI network constructed by STRING online database and Module analysis. The nodes meant proteins; the edges meant the interaction of proteins; green meant down-regulated DEGs and red meant up-regulated DEGs.

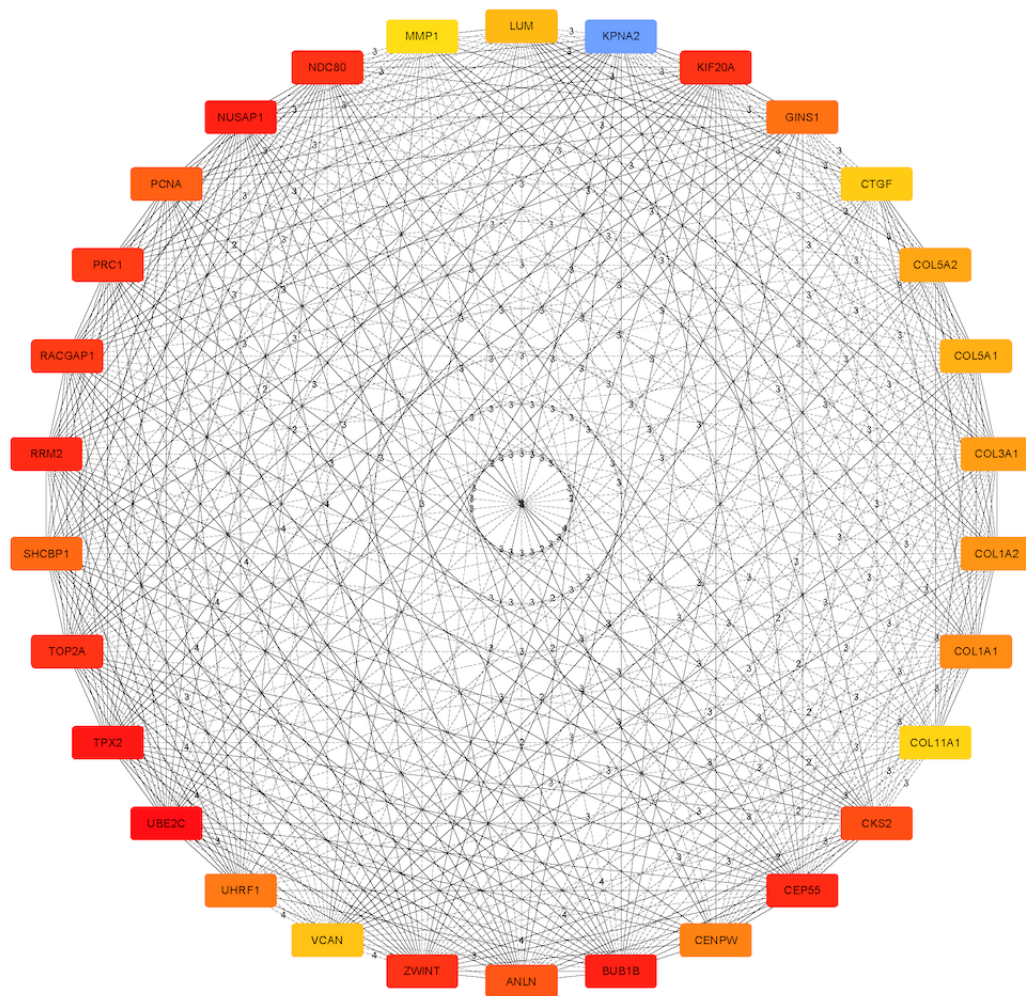


Figure 5

The hub genes. The nodes meant proteins; the edges meant the interaction of proteins; green meant down-regulated DEGs and red meant up-regulated DEGs.

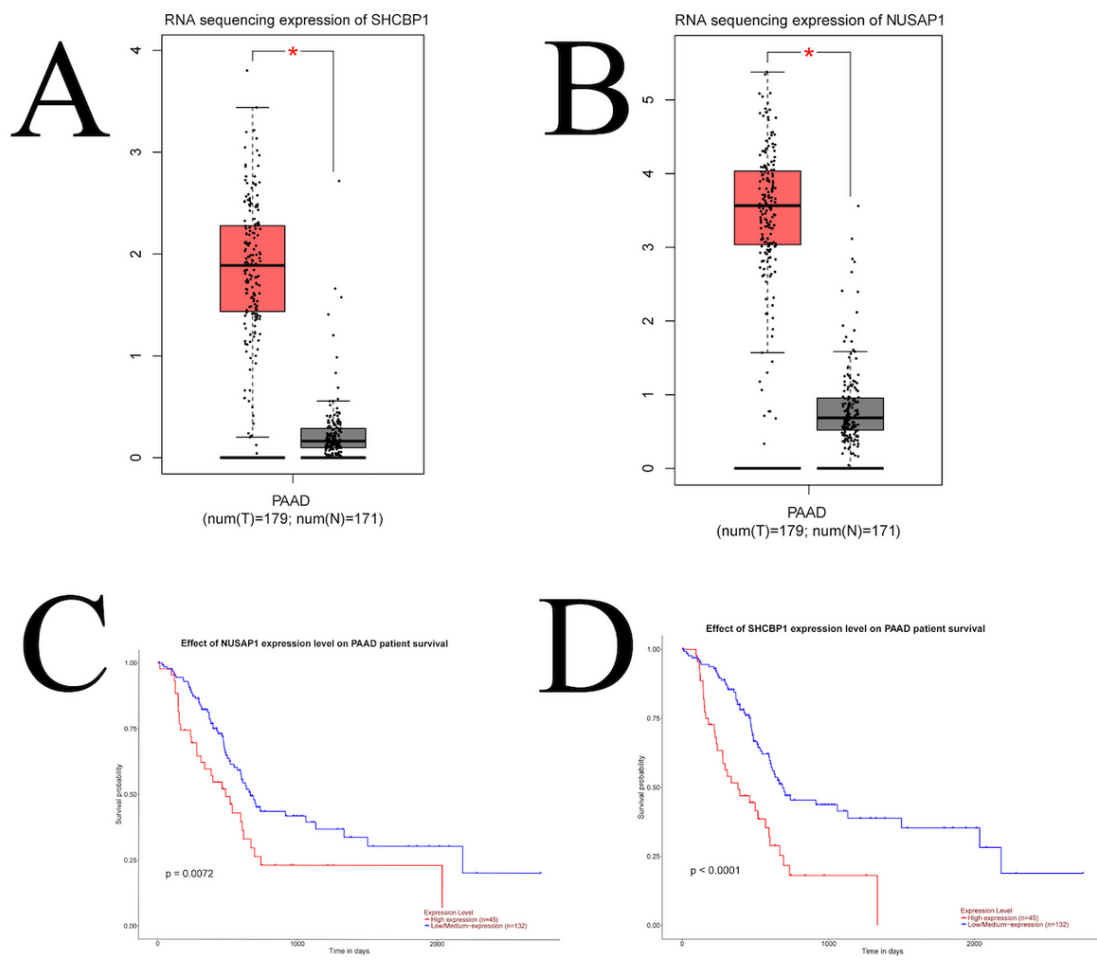


Figure 6

A. RNA sequencing expression of SHCBP1.B.RNA sequencing expression of NUSAP1.C.NUSAP1 expression level on PAAD patient survival,D.SHCBP1 expression level on PAAD patient survival