

# Prognostic value of Glypican family genes in early-stage pancreatic ductal adenocarcinoma after pancreaticoduodenectomy and possible mechanisms

**Jun-qi Liu**

Guangxi Medical University First Affiliated Hospital

**Xi-wen Liao**

Guangxi Medical University First Affiliated Hospital

**Xiang-kun Wang**

Guangxi Medical University First Affiliated Hospital

**Cheng-kun Yang**

Guangxi Medical University First Affiliated Hospital

**Xin Zhou**

Guangxi Medical University First Affiliated Hospital

**Zheng-qian Liu**

Guangxi Medical University First Affiliated Hospital

**Quan-fa Han**

Guangxi Medical University First Affiliated Hospital

**Tian-hao Fu**

Guangxi Medical University First Affiliated Hospital

**Guang-zhi Zhu**

Guangxi Medical University First Affiliated Hospital

**Chuang-ye Han**

Guangxi Medical University First Affiliated Hospital

**Hao Su**

Guangxi Medical University First Affiliated Hospital

**Jian-lu Huang**

Guangxi Medical University First Affiliated Hospital

**Guo-tian Ruan**

Guangxi Medical University First Affiliated Hospital

**Ling Yan**

Guangxi Medical University First Affiliated Hospital

**Xin-ping Ye**

Guangxi Medical University First Affiliated Hospital

**Tao Peng** (✉ [pengtaogmu@163.com](mailto:pengtaogmu@163.com))

the first affiliated hospital of Guangxi Medical University

---

## Research article

**Keywords:** GPC family genes, pancreatic ductal adenocarcinoma, prognostic indicator, mechanism

**Posted Date:** November 13th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-48421/v2>

**Version of Record:** A version of this preprint was published on December 10th, 2020. See the published version at <https://doi.org/10.1186/s12876-020-01560-0>.

## Abstract

**Background:** This study explored the prognostic significance of Glypican (GPC) family genes in patients with pancreatic ductal adenocarcinoma (PDAC) after pancreaticoduodenectomy using data from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO).

**Methods:** A total of 112 PDAC patients from TCGA and 48 patients from GEO were included in the analysis. The relationship between overall survival and the expression of GPC family genes as well as basic clinical characteristics was analyzed using the Kaplan-Meier method with the log-rank test. Joint effects survival analysis was performed to further examine the relationship between GPC genes and prognosis. A prognosis nomogram was established based on clinical characteristics and prognosis-related genes. Prognosis-related genes were investigated by genome-wide co-expression analysis and gene set enrichment analysis (GSEA) was carried out to identify potential mechanisms of these genes affecting prognosis.

**Results:** In TCGA database, high expression of *GPC2*, *GPC3*, and *GPC5* was significantly associated with favorable survival (log-rank  $P = 0.031$ ,  $0.021$ , and  $0.028$ , respectively; adjusted  $P$  value =  $0.005$ ,  $0.022$ , and  $0.020$ , respectively), and joint effects analysis of these genes was effective for prognosis prediction. The prognosis nomogram was applied to predict the survival probability using the total scores calculated. Genome-wide co-expression and GSEA analysis suggested that the *GPC2* may affect prognosis through sequence-specific DNA binding, protein transport, cell differentiation and oncogenic signatures (KRAS, RAF, STK33, and VEGFA). *GPC3* may be related to cell adhesion, angiogenesis, inflammatory response, signaling pathways like Ras, Rap1, PI3K-Akt, chemokine, GPCR, and signatures like cyclin D1, p53, PTEN. *GPC5* may be involved in transcription factor complex, TFRC1, oncogenic signatures (HOXA9 and BMI1), gene methylation, phospholipid metabolic process, glycerophospholipid metabolism, cell cycle, and EGFR pathway.

**Conclusion:** *GPC2*, *GPC3*, and *GPC5* expression may serve as prognostic indicators in PDAC, and combination of these genes showed a higher efficiency for prognosis prediction.

## Background

Pancreatic cancer (PC) is related to an unfavorable prognosis, and its mortality rate is close to its incidence rate (1). The incidence of PC is predicted to rise 40% in the next 10 years in North America and Europe (2), and according to the latest statistics, PC ranks fourth among cancers directly causing death for men and women in the United States (3), moreover, by 2030, its rank may increase to second (4). In China, the prognostic status of PC patients is also severe, and 5-year survival rate of patients with PC after age standardization is approximately 11.7% (5). Due to the unique biological behaviors of PC, metastasis is present when patients are diagnosed and only 9.7% patients can be diagnosed at an early stage (6). Furthermore, the 5-year survival rate is 9% for PC at all stages and 3% at advanced stages (3). So far, surgical resection remains the best therapy for PC at the early stage (7). Therefore, identifying reliable early molecular markers to improve prognosis of PC is important.

Glypican (GPC) family genes include six members (*GPC1*, *GPC2*, *GPC3*, *GPC4*, *GPC5*, *GPC6*), and all of the GPC family are expressed in human (8). Glypicans are attached to the cell membrane and function in biological processes such as cell and tissue growth, embryo development, and cell movement (9, 10). They are reported to be related to multiple diseases including various cancers. *GPC1* is upregulated in pancreatic cancer (11), esophageal cancer (12), and prostate cancer (13). Li et al. report that *GPC1* contributes to the proliferation and motility of esophageal cancer cells through the PTEN/Akt/ $\beta$ -catenin pathway (14). Increased level of *GPC3* in serum could serve as a marker for hepatoblastoma (15) as well as hepatocellular carcinoma (HCC) (16, 17). *GPC3* deletion mutation can help in diagnosis of Simpson-Golabi-Behmel syndrome type 1 (SGBS1), which is a serious genetic disease (18, 19). Overexpression of *GPC5* may accelerate tumor progression of lymphoma (20). In addition, *GPC5* may play a role in strengthening the interaction between Patched 1 and Hedgehog signaling in rhabdomyosarcoma (21). *GPC5* may serve as a key gene affecting the cell cycle of podocytes in kidneys, finally causing nephrotic syndrome (22).

Pancreatic ductal adenocarcinoma (PDAC) accounts for more than 80% of pancreatic neoplasms (1, 23). However, there are few studies on the prognostic value of GPC family genes in early-stage PDAC after pancreaticoduodenectomy despite the poor prognosis of this tumor type. In this study, we explored the relationship between GPC family genes expression and prognosis of PDAC patients.

# Methods

## Patient data

The RNA-sequencing dataset used in this study and the corresponding clinical data were acquired from The Cancer Genome Atlas (TCGA) (<https://portal.gdc.cancer.gov/>; accessed September 25, 2019), and *DESeq* was applied to normalize the initial material (24). To increase reliability of data analysis, previously established inclusion and exclusion criteria were used (25). The inclusion criteria were as follows: (X) survival information was complete; (X) histology result was confirmed as PDAC; (X) pathologic stage was X or X; (X) pancreaticoduodenectomy was carried out on patients. PDAC patients with pathologic stage X or X and those who underwent other surgical procedures were excluded from the study. According to the above criteria, 112 patients were included in the analysis. The clinical characteristics included in the analysis were age, sex, alcohol history, pathologic stage, histologic grade, radical resection, radiation therapy, targeted molecular therapy, overall survival (OS) time, and survival status. Dataset GSE62452 was downloaded from Gene Expression Omnibus (GEO) database to validate the prognostic value of survival-related genes (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE62452>; accessed October 5, 2020). Following the same criteria described above, we included 48 cases in this study.

## Analysis using public database

The expression status of GPC family genes in different normal tissues was analyzed by the Genotype-Tissue Expression (GTEx, <https://www.gtexportal.org/>, accessed October 9, 2019) website (26, 27). The Gene Expression Profiling Interactive Analysis (GEPIA, <http://gepia.cancer-pku.cn/>, accessed October 9, 2019), an online tool containing 9,736 tumors and 8,587 normal samples from the TCGA and the GTEx projects, was used to show expression level of each gene in both tumor and normal tissues of PDAC (28). The Database for Annotation, Visualization, and Integrated Discovery (DAVID) v6.8 (<https://david.ncifcrf.gov/>, accessed November 6, 2019) (29, 30) was chosen to carry out gene enrichment analysis containing Gene Ontology (GO) function analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. The possible functioning pathways of the genes were also investigated by Biological Network Gene Ontology (BiNGO) in Cytoscape (version 3.7.1) (31).

## Survival analysis

Two groups of patients were set up based on 50% cutoff expression value of each gene both in TCGA database and GEO database. The relationship between OS and gene expression level as well as basic clinical characteristics was analyzed using Kaplan-Meier method with the log-rank test. Log-rank  $P < 0.05$  was considered statistically significant. Multivariate Cox proportional hazards regression analysis was used to adjust for prognosis-significant factors. Hazard ratio (HR) and 95% confidence interval (CI) were considered to estimate the relative risk. Stratified analysis was carried out based on certain clinical characteristics of the patients for survival-related genes to explore their significance in prognosis. To understand the relationship between GPC genes and prognosis at a deeper level, joint effects survival analysis was taken into consideration. The survival-significant clinical characteristics, clinical factors usually related to prognosis of patients with malignant tumors clinically and prognosis-related genes were included to establish a prognosis nomogram. Better survival prediction could be made according to the total points.

## Genome-wide co-expression analysis

Genome-wide co-expression analysis of prognosis-related genes was performed to investigate their potential biological mechanisms based on TCGA database. A gene with Pearson correlation coefficient  $> 0.5$  and  $P < 0.05$  was considered as a co-expression gene. A co-expression network was built for each gene related to prognosis and its co-expressed genes using Cytoscape software (version 3.7.1) (32). GO function analysis and KEGG pathway analysis of these genes were also completed using DAVID (29, 30).

## Gene set enrichment analysis

To understand the underlying mechanisms of GPC genes affecting prognosis, we used Gene Set Enrichment Analysis (GSEA, <http://software.broadinstitute.org/gsea/index.jsp>, November 6, 2019) (33, 34). Databases c2 (c2.all.v7.0.symbols.gmt) and c6 (c6.all.v7.0.symbols.gmt) in the Molecular Signatures Database (MSigDB) (35) were used to search for possible pathways based

on TCGA database. Enrichment results were considered statistically significant if the nominal  $P$ -value was  $< 0.05$  and the false discovery rate (FDR) was  $< 0.25$ .

### Statistical analysis

Survival analysis was performed using Kaplan-Meier method with log-rank test. Univariate and multivariate survival analyses were performed with Cox proportional hazards regression model to calculate crude and adjusted HRs and 95% CIs. Survival curves were plotted using GraphPad Prism v.7.0 (GraphPad Software Inc., La Jolla, CA).

The unpaired  $t$  test was used to compare gene expression levels between normal and tumor tissues. The expression relationship of each GPC gene and its co-expressed genes was quantified by Pearson's correlation coefficient. The correlation plot was constructed using Cytoscape software (version 3.7.1). All statistical analyses were performed using SPSS v.25.0 software (IBM, Chicago, IL, USA). A  $P$  value  $< 0.05$  was considered statistically significant.

## Results

### Analysis using public database

The expression status of GPC family genes in tissues derived from various normal human organs was analyzed using GTEx (Fig. 1). The expression level of GPC family genes was lower in human pancreas than in other organs. The results of GEPIA analysis showed that expression of *GPC1*, *GPC3*, *GPC4*, and *GPC6* was significantly higher in PDAC tumor tissues than in normal tissues ( $P < 0.05$ ) (Fig. 2). GO functional enrichment analysis indicated that GPC family genes were mainly involved in composition of cell membrane, organelles and anchored components of the membrane, heparan sulfate proteoglycan binding, and glycosaminoglycan metabolic process (Fig. 3, Supplementary Table 1). The results of BiNGO analysis (Fig. 4) confirmed those of GO analysis.

### Survival analysis

The Kaplan-Meier method and log-rank test were used to investigate the association between basic clinical characteristics and OS in TCGA database. Supplementary Table 2 shows that histologic grade, extent of surgery, treatment with radiation and targeted molecular therapy were significant in OS. GPC family genes were divided into two groups based on expression level, and survival analysis was performed between the two groups. The results (Fig. 5A–F) demonstrated that expression of *GPC2*, *GPC3*, and *GPC5* was significantly associated with survival. The median survival time (MST) was significantly longer in patients with high expression of *GPC2*, *GPC3*, and *GPC5* than the low expression group (log-rank  $P = 0.031$ ,  $0.021$ , and  $0.028$ , respectively; MST, 634 days vs. 481 days, 614 days vs. 473 days, and 593 days vs. 485 days, respectively, Figure 5B, 5C, 5E and Fig. 6). After adjusting for survival-significant clinical parameters in a multivariate Cox proportional hazards regression model, *GPC2*, *GPC3*, and *GPC5* were still significantly associated with OS (Table 1) (adjusted  $P = 0.005$ , adjusted HR = 0.449, 95% CI = 0.258–0.782; adjusted  $P = 0.022$ , adjusted HR = 0.531, 95% CI = 0.309–0.914; and adjusted  $P = 0.020$ , adjusted HR = 0.525, 95% CI = 0.306–0.902, respectively). Results of stratified analysis for *GPC2*, *GPC3*, and *GPC5* are shown in Table 2. High expression of *GPC2* was significantly associated with better OS in patients who were male, were  $>60$  years old, had histologic grade G1 or G2, had R1 or Rx resection or whether received radiation therapy. *GPC3* expression was related to patients who were female, were  $>60$  years old, had histologic grade G1 or G2, or did not receive radiation or targeted molecular therapy. Moreover, *GPC5* could influence prognosis of patients who were  $\leq 60$  years old, had histologic grade G3 or G4, had R1 or Rx resection, or did not receive radiation or targeted molecular therapy.

### Joint effects analysis

Based on the prognostic significance of each GPC family gene, we combined every two genes among *GPC2*, *GPC3*, and *GPC5* to investigate their significance in PDAC prognosis. The combination of *GPC2* and *GPC3* was associated with worse survival outcome in group 1 (MST = 278 days, adjusted  $P$  value  $< 0.001$ ). The group of *GPC2* and *GPC5* was associated with the highest risk of death in group 2 (MST = 278 days, adjusted  $P$  value  $< 0.001$ ) and the group combining *GPC3* and *GPC5* showed the poorest prognosis in group 3 (MST = 278 days, adjusted  $P$  value  $< 0.001$ ).

We also analyzed survival associated with the three genes simultaneously. Group A showed the worst in survival status (MST = 219 days, adjusted  $P$  value = 0.018), whereas the best survival was observed in group D (MST = 702 days, adjusted  $P$  value < 0.001). These data are shown in Table 3 and Fig. 7A–D showed the survival curves.

### Prognosis nomogram

Based on the status of each clinical parameter and expression levels of *GPC2*, *GPC3*, and *GPC5*, a score for each variable was calculated. The total score could be calculated to predict 1-, 2-, and 3- year survival probabilities. The nomogram (Fig. 8) indicated that *GPC2*, *GPC3*, and *GPC5* affected the prognosis of PDAC to different degrees.

### Validation dataset to demonstrate the prognostic value of survival-related genes

To further understand the prognostic value of *GPC2*, *GPC3*, and *GPC5*, we acquired the GSE62452 dataset from GEO database. As shown in Supplementary Table 3, histologic grade was significantly associated with OS. GPC family genes were also divided into two groups by the median expression level of each gene and survival analysis between the two groups was carried out. Table 4 and Fig. 9A–F show that higher expression of *GPC3* was significantly related to better survival (log-rank  $P$  = 0.038) and higher expression of *GPC2* and *GPC5* was also related to better survival, though not significantly (log-rank  $P$  = 0.337 and 0.090, respectively). Multivariate Cox proportional hazards regression analysis adjusted for prognosis-related clinical characteristics showed that none of these genes was significantly correlated to overall survival (all adjusted  $P$  > 0.05).

### Genome-wide co-expression analysis of *GPC2*, *GPC3* and *GPC5* in PDAC

Genome-wide co-expression analysis was performed for each of these genes to investigate their related functional pathways through TCGA database. For *GPC2* and its co-expressed genes, a correlation network was established as shown in Fig. 10A (Supplementary Table 4). GO analysis indicated that *GPC2* and its co-expressed genes functioned mainly in sequence-specific DNA binding, protein transport, cell differentiation, and anterior/posterior pattern specification (Fig. 10B, Supplementary Table 5).

The correlation network for *GPC3* and its co-expressed genes (Fig. 11A, Supplementary Table 6) identified 511 positively co-expressed genes and 25 negatively co-expressed genes. GO analysis of these genes indicated that they were enriched in cell adhesion, angiogenesis, and inflammatory response (Fig. 11B, Supplementary Table 7). And KEGG analysis indicated that these genes were related to several biological processes, mainly in Ras, Rap1, PI3K-Akt, and chemokine signaling pathways (Fig. 11C, Supplementary Table 8).

The correlation network for *GPC5* and its co-expressed genes was shown in Fig. 12A and Supplementary Table 9. The results of GO analysis showed that these genes were associated with transcription factor complex and phospholipid metabolic process (Fig. 12B, Supplementary Table 10). KEGG analysis showed that these genes were involved in pancreatic secretion and glycerophospholipid metabolism (Fig. 12C, Supplementary Table 11).

### Gene set enrichment analysis

GSEA was carried out to explore possible mechanisms of GPC family genes affecting prognosis of PDAC patients through TCGA database. The results of c6 reference indicated that low *GPC2* expression was closely related to oncogenic signatures such as KRAS, RAF1, STK33, and VEGFA (Fig. 13A–F; Supplementary Table 12). GSEA results of c2 enrichment showed that high *GPC3* expression was associated with neuroactive ligand receptor interaction and GPCR ligand binding (Fig. 14A–C; Supplementary Table 13), and c6 enrichment suggested that high *GPC3* expression was correlated to cyclin D1, p53, and PTEN (Fig. 13D–F; Supplementary Table 14). For *GPC5*, c2 reference indicated that low expression of *GPC5* was related to the EGFR pathway, gene methylation status, TFRC1, and the cell cycle (Fig. 15A–D; Supplementary Table 15) and c6 reference indicated that low *GPC5* expression was related to HOXA9 and BMI1 (Fig. 15E–F; Supplementary Table 16).

## Discussion

In this research, we studied the relationship between GPC family gene expression and prognosis of early-stage PDAC patients after pancreaticoduodenectomy both in TCGA database and GEO database. We concluded that high expression of *GPC2*, *GPC3*, and

*GPC5* was significantly related to favorable prognosis in TCGA database, suggesting the value of these genes as biomarkers for predicting the prognosis of PDAC patients. Moreover, combination of the three genes showed a better predictive value for PDAC prognosis.

GPC family genes may contribute to the malignant behaviors of tumors and they are closely related to the development and prognosis of various cancers. Li et al. demonstrated that *GPC1* is enriched in exosomes produced by colorectal cancer cells HT-29 and HCT-116, and increased expression level of miR-96-5p and miR-149 can restrain both *GPC1* expression and cell proliferation of the tumor, suggesting that *GPC1* can be used as a marker for diagnosis and therapy of colorectal cancer (36). It is reported that *GPC2* could promote the proliferation of neuroblastoma cells as a result of MYCN binding to a motif of the promoter of *GPC2* and gain of chromosome 7q (37). *GPC2* can also be used as an effective prognostic indicator for prostate cancer and neuroblastoma (37-39). *GPC3* blocks the cell cycle in renal cancer cells 786-O and ACHN at G1 phase (40). Overexpression of *GPC3* reduces progression and metastasis of breast cancer cells LM3 through targeting canonical Wnt pathway (41). The *GPC5* rs2352028 variant and lower expression of this gene may contribute to increased risk of lung cancer (42, 43). Sun et al have shown that *GPC5* regulates epithelial-mesenchymal transition to reduce invasion of prostate cancer cells (44). Its expression can serve as a prognostic indicator in a cohort of prostate cancer patients in China (45). In this study, we demonstrated the relationship between OS and expression levels of *GPC2*, *GPC3*, and *GPC5*. Combined with results of GEPIA, it demonstrates their roles as tumor suppressor genes in PDAC.

To explore potential mechanisms of GPC genes affecting prognosis, we conducted GSEA and genome-wide co-expression analyses. The results showed that *GPC2* was associated with sequence-specific DNA binding, protein transport, cell differentiation and oncogenic signatures (KRAS, RAF, STK33, and VEGFA). In pancreatic cancer, mutation of *TP53* at codon 249 can alter the structure of p53, thus affecting its binding to a specific region of DNA and enhancing the risk of cancer (46, 47). A study showed that GDF11 regulates the biological behaviors of pancreatic cancer cells to influence their differentiation and high expression of GDF11 is associated with favorable OS in pancreatic cancer (48). RAF1 accelerates migration and invasion of pancreatic cancer and disorders of the RAF1 pathway are related to worse prognosis in pancreatic cancer patients (49, 50). Moreover, microRNA-216a may downregulate RAF1 in pancreatic cancer and increase cell apoptosis (51). VEGFA expression can increase as a result of the long non-coding RNA (lncRNA) 00511 in PDAC, which finally promotes tumor progression. The expression level of lnc00511 can be used as an indicator of prognosis in PDAC (52).

*GPC3* is related to cell adhesion, angiogenesis, inflammatory response, signaling pathways like Ras, Rap1, PI3K-Akt, chemokine, GPCR, and signatures like cyclin D1, p53, PTEN. For pancreatic cancer patients, the degree of inflammatory response can be measured by serum lactate dehydrogenase level and it is associated with the outcome of patients (53). Angiogenesis is dysregulated in PDAC, and it contributes to proliferation and deterioration of the tumor, making survival of patients worse (54, 55). Certain mutations of *KRAS* are associated with the response to drugs in PDAC cells (56). In PDAC associated with the *KRAS* mutation, decitabine therapy inhibits tumor growth (57). ARF6 is reported to be in close relationship with the Ras pathway and its overexpression is related to unfavorable prognosis of PDAC patients (58). PTEN plays a role in pancreatic cancer growth. The function of PTEN is regulated by HNF1A and finally affects the survival of pancreatic cancer patients (59, 60).

*GPC5* is involved in the transcription factor complex TFRC1, oncogenic signatures HOXA9 and BMI1, gene methylation, phospholipid metabolic process, glycerophospholipid metabolism, cell cycle, and the EGFR pathway. In pancreatic cancer, the transcription factor hif-2 $\alpha$  can speed up metabolism and promote tumor proliferation and high level of hif-2 $\alpha$  correlates with worse OS (61, 62). The methylation status of *GRAP2*, *ICAM3*, *A2ML1*, *MUC1*, and *MUC4* can influence the expression of these genes, which is associated with survival of pancreatic cancer (63, 64). Phosphatidylserine is related to apoptosis of pancreatic cancer cells with the involvement of microparticles (65). Stimuli such as oxidative stress can make phosphatidylserine appear outside on the pancreatic cancer cell membrane, finally leading to dysregulation of factors and cells such as VEGF and macrophages, making prognosis of patients unfavorable (66-68). The EGFR pathway contributes to pancreatic cancer growth and accelerates invasion of the cancer as a result of lnc00976 overexpression, which can deteriorate the outcome of patients (69, 70).

The present study had several limitations. First, clinical data acquired from TCGA and GEO databases did not include all the relevant information, and there may be some factors that needed to be adjusted. Second, because the study included PDAC patients who underwent pancreaticoduodenectomy, the sample size was relatively small. Third, the results of genome-wide

analysis and GSEA analysis were based on online databases to predict potential processes influencing prognosis, and further studies at molecular and genomic levels are necessary to confirm the results.

Despite these limitations, we identified *GPC2*, *GPC3*, and *GPC5* as biomarkers for prognosis of PDAC patients and showed that joint effects analysis was more effective for prediction of prognosis. We also explored possible mechanisms of survival-significant genes affecting PDAC prognosis through genome-wide analysis and GSEA analysis. These results could all improve prognostic prediction for PDAC and provide information valuable for the management of PDAC patients and making better clinical decisions in this population.

## Conclusion

We identified *GPC2*, *GPC3*, and *GPC5* as potential prognostic indicators for PDAC patients and showed that combination of these genes was more effective for prognosis prediction. Possible mechanisms of *GPC2* influencing prognosis may involve sequence-specific DNA binding, protein transport, cell differentiation and oncogenic signatures (KRAS, RAF, STK33, and VEGFA). *GPC3* may be related to cell adhesion, angiogenesis, inflammatory response, signaling pathways such as Ras, Rap1, PI3K-Akt, chemokine, and GPCR, and signatures including cyclin D1, p53, and PTEN. *GPC5* may be involved in the transcription factor complex TFRC1, the oncogenic signatures HOXA9 and BMI1, gene methylation, the phospholipid metabolic process, glycerophospholipid metabolism, cell cycle, and the EGFR pathway.

## Abbreviation

GPC: Glypican; PDAC: pancreatic ductal adenocarcinoma; TCGA: The Cancer Genome Atlas; GEO: Gene Expression Omnibus; GSEA: Gene Set Enrichment Analysis; DAVID: Database for Annotation, Visualization, and Integrated Discovery; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; BiNGO: Biological Network Gene Ontology; GTEX: Genotype-Tissue Expression; GEPIA: Gene Expression Profiling Interactive Analysis

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Availability of data and materials

Datasets generated and analyzed during the current study are available from The Cancer Genome Atlas, <https://portal.gdc.cancer.gov> and Gene Expression Omnibus, <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE62452>.

### Competing interests

The authors declare that they have no competing interests.

### Funding

This work was supported in part by the National Natural Science Foundation of China (No.: 81560535, 81802874, 81072321, 30760243, 30460143 and 30560133), Natural Science Foundation of Guangxi Province of China (Grant No.2017JJB140189y, 2018GXNSFAA050119), 2009 Program for New Century Excellent Talents in University (NCET), Guangxi Natural Sciences Foundation (No.: GuiKeGong 1104003A-7), and Guangxi Health Ministry Medicine Grant (Key-Scientific Research-Grant Z201018). The present study was also partly supported by Scientific Research Fund of the Health and Family Planning Commission of Guangxi Zhuang Autonomous Region (Z2016318, Z2016307), the Guangxi Key R & D Program (GKEAB18221019), The Basic Ability Improvement Project for Middle-aged and Young Teachers in Colleges and Universities in Guangxi (2018KY0110),

Innovation Project of Guangxi Graduate Education (JGY2018037), and 2018 Innovation Project of Guangxi Graduate Education (YCBZ2018036). In addition, the present study was also partly supported by Guangxi Key Laboratory for the Prevention and Control of Viral Hepatitis (No. GXCDCKL201902) and Research Institute of Innovative Think-tank in Guangxi Medical University (The gene-environment interaction in hepatocarcinogenesis in Guangxi HCCs and its translational applications in the HCC prevention). We would also like to acknowledge the support from the Key laboratory of High-Incidence-Tumor Prevention & Treatment (Guangxi Medical University), Ministry of Education (GKE2018-01, GKE2019-11).

### Authors' contributions

JL, XL, XY and TP designed this study. JL, WX, CY, XZ, ZL, QH, and TF analyzed the data. JL, GZ, CH, HS, JH, GR, and LY prepared figures and tables. JL, XL, XW and XZ finished the manuscript. XY and TP guided and supervised the whole research and TP revised the final manuscript. All authors read and approved the final manuscript.

### Acknowledgement

We would like to sincerely thank the contributors of The Cancer Genome Atlas and Gene Expression Omnibus for sharing the data with the public.

## References

1. Kamisawa T, Wood LD, Itoi T, Takaori K. Pancreatic cancer. *Lancet (London, England)*. 2016;388(10039):73-85.
2. Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res*. 2014;74(11):2913-21.
3. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA: a cancer journal for clinicians*. 2019;69(1):7-34.
4. Chen H, Kong Y, Yao Q, Zhang X, Fu Y, Li J, et al. Three hypomethylated genes were associated with poor overall survival in pancreatic cancer patients. *Aging*. 2019;11(3):885-97.
5. Zeng H, Zheng R, Guo Y, Zhang S, Zou X, Wang N, et al. Cancer survival in China, 2003-2005: a population-based study. *International journal of cancer*. 2015;136(8):1921-30.
6. Zhang L, Sanagapalli S, Stoita A. Challenges in diagnosis of pancreatic cancer. *World J Gastroenterol*. 2018;24(19):2047-60.
7. Vincent A, Herman J, Schulick R, Hruban RH, Goggins M. Pancreatic cancer. *Lancet (London, England)*. 2011;378(9791):607-20.
8. Li N, Gao W, Zhang YF, Ho M. Glypicans as Cancer Therapeutic Targets. *Trends in cancer*. 2018;4(11):741-54.
9. Kaur SP, Cummings BS. Role of glypicans in regulation of the tumor microenvironment and cancer progression. *Biochem Pharmacol*. 2019;168:108-18.
10. Filmus J, Capurro M, Rast J. Glypicans. *Genome biology*. 2008;9(5):224.
11. Kayed H, Kleeff J, Keleg S, Jiang X, Penzel R, Giese T, et al. Correlation of glypican-1 expression with TGF-beta, BMP, and activin receptors in pancreatic ductal adenocarcinoma. *International journal of oncology*. 2006;29(5):1139-48.
12. Hara H, Takahashi T, Serada S, Fujimoto M, Ohkawara T, Nakatsuka R, et al. Overexpression of glypican-1 implicates poor prognosis and their chemoresistance in oesophageal squamous cell carcinoma. *British journal of cancer*. 2016;115(1):66-75.
13. Suhovskih AV, Mostovich LA, Kunin IS, Boboev MM, Nepomnyashchikh GI, Aidagulova SV, et al. Proteoglycan expression in normal human prostate tissue and prostate cancer. *ISRN oncology*. 2013;2013:680136.
14. Li J, Chen Y, Zhan C, Zhu J, Weng S, Dong L, et al. Glypican-1 Promotes Tumorigenesis by Regulating the PTEN/Akt/beta-Catenin Signaling Pathway in Esophageal Squamous Cell Carcinoma. *Digestive diseases and sciences*. 2019;64(6):1493-502.
15. Zhou S, O'Gorman MR, Yang F, Andresen K, Wang L. Glypican 3 as a Serum Marker for Hepatoblastoma. *Scientific reports*. 2017;7:45932.
16. Tangkijvanich P, Chanmee T, Komtong S, Mahachai V, Wisedopas N, Pothacharoen P, et al. Diagnostic role of serum glypican-3 in differentiating hepatocellular carcinoma from non-malignant chronic liver disease and other liver cancers. *Journal of gastroenterology and hepatology*. 2010;25(1):129-37.

17. Jia X, Liu J, Gao Y, Huang Y, Du Z. Diagnosis accuracy of serum glypican-3 in patients with hepatocellular carcinoma: a systematic review with meta-analysis. *Archives of medical research*. 2014;45(7):580-8.
18. Sajorda BJ, Gonzalez-Gandolfi CX, Hathaway ER, Kalish JM. Simpson-Golabi-Behmel Syndrome Type 1. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, et al., editors. *GeneReviews*(R). Seattle (WA): University of Washington, Seattle University of Washington, Seattle. GeneReviews is a registered trademark of the University of Washington, Seattle. All rights reserved.; 1993.
19. Veugelers M, Cat BD, Muyltermans SY, Reekmans G, Delande N, Frints S, et al. Mutational analysis of the GPC3/GPC4 glypican gene cluster on Xq26 in patients with Simpson-Golabi-Behmel syndrome: identification of loss-of-function mutations in the GPC3 gene. *Human molecular genetics*. 2000;9(9):1321-8.
20. Yu W, Inoue J, Imoto I, Matsuo Y, Karpas A, Inazawa J. GPC5 is a possible target for the 13q31-q32 amplification detected in lymphoma cell lines. *Journal of human genetics*. 2003;48(6):331-5.
21. Li F, Shi W, Capurro M, Filmus J. Glypican-5 stimulates rhabdomyosarcoma cell proliferation by activating Hedgehog signaling. *The Journal of cell biology*. 2011;192(4):691-704.
22. Okamoto K, Tokunaga K, Doi K, Fujita T, Suzuki H, Katoh T, et al. Common variation in GPC5 is associated with acquired nephrotic syndrome. *Nature genetics*. 2011;43(5):459-63.
23. Seufferlein T, Bachet JB, Van Cutsem E, Rougier P. Pancreatic adenocarcinoma: ESMO-ESDO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of oncology : official journal of the European Society for Medical Oncology*. 2012;23 Suppl 7:vii33-40.
24. Anders S, McCarthy DJ, Chen Y, Okoniewski M, Smyth GK, Huber W, et al. Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nature protocols*. 2013;8(9):1765-86.
25. Liao X, Huang K, Huang R, Liu X, Han C, Yu L, et al. Genome-scale analysis to identify prognostic markers in patients with early-stage pancreatic ductal adenocarcinoma after pancreaticoduodenectomy. *OncoTargets and therapy*. 2017;10:4493-506.
26. The Genotype-Tissue Expression (GTEx) project. *Nature genetics*. 2013;45(6):580-5.
27. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science (New York, NY)*. 2015;348(6235):648-60.
28. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic acids research*. 2017;45(W1):W98-w102.
29. Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome biology*. 2003;4(5):P3.
30. Said-Salman IH, Jebaï FA, Yusef HH, Moustafa ME. Global gene expression analysis of *Escherichia coli* K-12 DH5alpha after exposure to 2.4 GHz wireless fidelity radiation. *Scientific reports*. 2019;9(1):14425.
31. Maere S, Heymans K, Kuiper M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics (Oxford, England)*. 2005;21(16):3448-9.
32. Munoz-San Martin M, Reverter G, Robles-Cedeno R, Buxo M, Ortega FJ, Gomez I, et al. Analysis of miRNA signatures in CSF identifies upregulation of miR-21 and miR-146a/b in patients with multiple sclerosis and active lesions. *Journal of neuroinflammation*. 2019;16(1):220.
33. Mandelbroum S, Manber Z, Elroy-Stein O, Elkon R. Recurrent functional misinterpretation of RNA-seq data caused by sample-specific gene length bias. *PLoS biology*. 2019;17(11):e3000481.
34. Li Y, Xue Z, Dong X, Liu Q, Liu Z, Li H, et al. Mitochondrial dysfunction and oxidative stress in bone marrow stromal cells induced by daunorubicin leads to DNA damage in hematopoietic cells. *Free radical biology & medicine*. 2019.
35. Frost HR. Computation and application of tissue-specific gene set weights. *Bioinformatics (Oxford, England)*. 2018;34(17):2957-64.
36. Li J, Chen Y, Guo X, Zhou L, Jia Z, Peng Z, et al. GPC1 exosome and its regulatory miRNAs are specific markers for the detection and target therapy of colorectal cancer. *Journal of cellular and molecular medicine*. 2017;21(5):838-47.
37. Bosse KR, Raman P, Zhu Z, Lane M, Martinez D, Heitzeneder S, et al. Identification of GPC2 as an Oncoprotein and Candidate Immunotherapeutic Target in High-Risk Neuroblastoma. *Cancer Cell*. 2017;32(3):295-309.e12.

38. Li N, Fu H, Hewitt SM, Dimitrov DS, Ho M. Therapeutically targeting glypican-2 via single-domain antibody-based chimeric antigen receptors and immunotoxins in neuroblastoma. *Proceedings of the National Academy of Sciences of the United States of America*. 2017;114(32):E6623-e31.
39. Xu N, Wu YP, Yin HB, Xue XY, Gou X. Molecular network-based identification of competing endogenous RNAs and mRNA signatures that predict survival in prostate cancer. *Journal of translational medicine*. 2018;16(1):274.
40. Valsechi MC, Oliveira AB, Conceicao AL, Stuqui B, Candido NM, Provazzi PJ, et al. GPC3 reduces cell proliferation in renal carcinoma cell lines. *BMC cancer*. 2014;14:631.
41. Fernandez D, Guerenio M, Lago Huvelle MA, Cercato M, Peters MG. Signaling network involved in the GPC3-induced inhibition of breast cancer progression: role of canonical Wnt pathway. *Journal of cancer research and clinical oncology*. 2018;144(12):2399-418.
42. Zheng Y, Kan M, Yu L, Niu X, Zhou D, He L, et al. GPC5 rs2352028 polymorphism and risk of lung cancer in Han Chinese. *Cancer investigation*. 2012;30(1):13-9.
43. Li Y, Sheu CC, Ye Y, de Andrade M, Wang L, Chang SC, et al. Genetic variants and risk of lung cancer in never smokers: a genome-wide association study. *The Lancet Oncology*. 2010;11(4):321-30.
44. Sun Y, Xu K, He M, Fan G, Lu H. Overexpression of Glypican 5 (GPC5) Inhibits Prostate Cancer Cell Proliferation and Invasion via Suppressing Sp1-Mediated EMT and Activation of Wnt/beta-Catenin Signaling. *Oncology research*. 2018;26(4):565-72.
45. Zhang C, Liu Z, Wang L, Qiao B, Du E, Li L, et al. Prognostic significance of GPC5 expression in patients with prostate cancer. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*. 2016;37(5):6413-8.
46. Mohamadkhani A, Naderi E, Sharafkhah M, Fazli HR, Moradzadeh M, Pourshams A. Detection of TP53 R249 Mutation in Iranian Patients with Pancreatic Cancer. *Journal of oncology*. 2013;2013:738915.
47. Subekti DRG, Murata A, Itoh Y, Fukuchi S, Takahashi H, Kanbayashi S, et al. The Disordered Linker in p53 Participates in Nonspecific Binding to and One-Dimensional Sliding along DNA Revealed by Single-Molecule Fluorescence Measurements. *Biochemistry*. 2017;56(32):4134-44.
48. Liu Y, Shao L, Chen K, Wang Z, Wang J, Jing W, et al. GDF11 restrains tumor growth by promoting apoptosis in pancreatic cancer. *OncoTargets and therapy*. 2018;11:8371-9.
49. Qi ZH, Xu HX, Zhang SR, Xu JZ, Li S, Gao HL, et al. RIPK4/PEBP1 axis promotes pancreatic cancer cell migration and invasion by activating RAF1/MEK/ERK signaling. *International journal of oncology*. 2018;52(4):1105-16.
50. Veit C, Genze F, Menke A, Hoeffert S, Gress TM, Gierschik P, et al. Activation of phosphatidylinositol 3-kinase and extracellular signal-regulated kinase is required for glial cell line-derived neurotrophic factor-induced migration and invasion of pancreatic carcinoma cells. *Cancer research*. 2004;64(15):5291-300.
51. Zhang Y, Tang X, Shi M, Wen C, Shen B. MiR-216a decreases MALAT1 expression, induces G2/M arrest and apoptosis in pancreatic cancer cells. *Biochemical and biophysical research communications*. 2017;483(2):816-22.
52. Zhao X, Liu Y, Li Z, Zheng S, Wang Z, Li W, et al. Linc00511 acts as a competing endogenous RNA to regulate VEGFA expression through sponging hsa-miR-29b-3p in pancreatic ductal adenocarcinoma. *Journal of cellular and molecular medicine*. 2018;22(1):655-67.
53. Yu SL, Xu LT, Qi Q, Geng YW, Chen H, Meng ZQ, et al. Serum lactate dehydrogenase predicts prognosis and correlates with systemic inflammatory response in patients with advanced pancreatic cancer after gemcitabine-based chemotherapy. *Scientific reports*. 2017;7:45194.
54. Li S, Xu HX, Wu CT, Wang WQ, Jin W, Gao HL, et al. Angiogenesis in pancreatic cancer: current research status and clinical implications. *Angiogenesis*. 2019;22(1):15-36.
55. Zhang Z, Ji S, Zhang B, Liu J, Qin Y, Xu J, et al. Role of angiogenesis in pancreatic cancer biology and therapy. *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie*. 2018;108:1135-40.
56. Hobbs GA, Baker NM, Miermont AM, Thurman RD, Pierobon M, Tran TH, et al. Atypical KRASG12R Mutant Is Impaired in PI3K Signaling and Macropinocytosis in Pancreatic Cancer. *Cancer discovery*. 2019.
57. Mottini C, Tomihara H, Carrella D, Lamolinara A, Iezzi M, Huang JK, et al. Predictive Signatures Inform the Effective Repurposing of Decitabine to Treat KRAS-Dependent Pancreatic Ductal Adenocarcinoma. *Cancer research*. 2019;79(21):5612-

25.

58. Hashimoto S, Furukawa S, Hashimoto A, Tsutaho A, Fukao A, Sakamura Y, et al. ARF6 and AMAP1 are major targets of KRAS and TP53 mutations to promote invasion, PD-L1 dynamics, and immune evasion of pancreatic cancer. *Proceedings of the National Academy of Sciences of the United States of America*. 2019;116(35):17450-9.
59. Yu Y, Liang S, Zhou Y, Li S, Li Y, Liao W. HNF1A/CASC2 regulates pancreatic cancer cell proliferation through PTEN/Akt signaling. *Journal of cellular biochemistry*. 2019;120(3):2816-27.
60. Zhang H, Feng X, Zhang M, Liu A, Tian L, Bo W, et al. Long non-coding RNA CASC2 upregulates PTEN to suppress pancreatic carcinoma cell metastasis by downregulating miR-21. *Cancer cell international*. 2019;19:18.
61. Zhang Q, Lou Y, Zhang J, Fu Q, Wei T, Sun X, et al. Hypoxia-inducible factor-2alpha promotes tumor progression and has crosstalk with Wnt/beta-catenin signaling in pancreatic cancer. *Molecular cancer*. 2017;16(1):119.
62. Li W, Chen C, Zhao X, Ye H, Zhao Y, Fu Z, et al. HIF-2alpha regulates non-canonical glutamine metabolism via activation of PI3K/mTORC2 pathway in human pancreatic ductal adenocarcinoma. *Journal of cellular and molecular medicine*. 2017;21(11):2896-908.
63. Kong L, Liu P, Zheng M, Xue B, Liang K, Tan X. Multi-omics analysis based on integrated genomics, epigenomics and transcriptomics in pancreatic cancer. *Epigenomics*. 2020.
64. Yokoyama S, Hamada T, Higashi M, Matsuo K, Maemura K, Kurahara H, et al. Predicted Prognosis of Patients with Pancreatic Cancer by Machine Learning. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2020.
65. Schneider J, Chromik AM, Uhl W, Mugge A, Bulut D. Apoptosis in esophagus and pancreas carcinoma cells induced by circulating microparticles is related to phosphatidyl serine and microparticle-associated caspases. *Medical oncology (Northwood, London, England)*. 2012;29(2):962-9.
66. Sharma B, Kanwar SS. Phosphatidylserine: A cancer cell targeting biomarker. *Seminars in cancer biology*. 2018;52(Pt 1):17-25.
67. Wang H, Zhan M, Yang R, Shi Y, Liu Q, Wang J. Elevated expression of NFE2L3 predicts the poor prognosis of pancreatic cancer patients. *Cell cycle (Georgetown, Tex)*. 2018;17(17):2164-74.
68. Di Caro G, Cortese N, Castino GF, Grizzi F, Gavazzi F, Ridolfi C, et al. Dual prognostic significance of tumour-associated macrophages in human pancreatic adenocarcinoma treated or untreated with chemotherapy. *Gut*. 2016;65(10):1710-20.
69. Lei S, He Z, Chen T, Guo X, Zeng Z, Shen Y, et al. Long noncoding RNA 00976 promotes pancreatic cancer progression through OTUD7B by sponging miR-137 involving EGFR/MAPK pathway. *Journal of experimental & clinical cancer research : CR*. 2019;38(1):470.
70. Kim J, Jo YH, Jang M, Nguyen NNY, Yun HR, Ko SH, et al. PAC-5 Gene Expression Signature for Predicting Prognosis of Patients with Pancreatic Adenocarcinoma. *Cancers*. 2019;11(11).

## Tables

Table 1 Prognostic value of Glypican family genes in The Cancer Genome Atlas database.

| No. of events (%) | MST (days) | Crude HR<br>95% CI | Crude<br>P value | Adjusted HR<br>95% CI <sup>a</sup> | Adjusted<br>P value <sup>a</sup> |
|-------------------|------------|--------------------|------------------|------------------------------------|----------------------------------|
| 36/56(64.3%)      | 518        | 1                  |                  | 1                                  |                                  |
| 33/56(58.9%)      | 511        | 0.986(0.610-1.594) | 0.955            | 1.120 (0.661-1.896)                | 0.674                            |
| 41/56(73.2%)      | 481        | 1                  |                  | 1                                  |                                  |
| 28/56(50.0%)      | 634        | 0.589(0.362-0.959) | 0.031            | 0.449(0.258-0.782)                 | 0.005                            |
| 38/56(67.9%)      | 473        | 1                  |                  | 1                                  |                                  |
| 31/56(55.4%)      | 614        | 0.568(0.349-0.925) | 0.021            | 0.531(0.309-0.914)                 | 0.022                            |
| 35/56(62.5%)      | 517        | 1                  |                  | 1                                  |                                  |
| 34/56(60.7%)      | 592        | 1.066(0.659-1.723) | 0.794            | 1.252(0.744-2.105)                 | 0.397                            |
| 35/56(62.5%)      | 485        | 1                  |                  | 1                                  |                                  |
| 34/56(60.7%)      | 593        | 0.577(0.351-0.948) | 0.028            | 0.525(0.306-0.902)                 | 0.020                            |
| 37/56(66.1%)      | 485        | 1                  |                  | 1                                  |                                  |
| 32/56(57.1%)      | 603        | 0.647(0.393-1.063) | 0.083            | 0.891(0.501-1.585)                 | 0.693                            |

Notes: <sup>a</sup>Adjusted for histologic grade, targeted molecular therapy, radiation therapy and radical resection. Abbreviations: MST, median survival time; HR, hazard ratio; CI, confidence interval.

Table 2 Stratified analysis of Glypican genes in The Cancer Genome Atlas database. Abbreviations: HR, hazard ratio; CI, confidence interval.

Notes: a Adjusted for histologic grade, targeted molecular therapy, radiation therapy and radical resection. b Information of radical resection was unavailable in 2 patients. c Information of radiation therapy was unavailable in 12 patients. d Information of targeted molecular therapy was unavailable in 10 patients.

Table 3 Joint effects analysis of combination of Glypican genes in The Cancer Genome Atlas database.

|   | <i>GPC2</i> |      |                        |                       | <i>GPC3</i> |      |                        |                       | <i>GPC5</i> |      |                        |                       |
|---|-------------|------|------------------------|-----------------------|-------------|------|------------------------|-----------------------|-------------|------|------------------------|-----------------------|
|   | Low         | High | Adjusted HR<br>95% CI  | Adjusted P<br>95%CI a | Low         | High | Adjusted HR<br>95% CI  | Adjusted P<br>95%CI a | Low         | High | Adjusted HR<br>95% CI  | Adjusted P<br>95%CI a |
| Age (years)                             |             |      |                        |                       |             |      |                        |                       |             |      |                        |                       |
| ≤60                                     | 16          | 22   | 0.433<br>(0.119-1.569) | 0.202                 | 20          | 18   | 1.061<br>(0.238-4.729) | 0.938                 | 18          | 20   | 0.197<br>(0.041-0.951) | 0.043                 |
| >60                                     | 40          | 34   | 0.415<br>(0.218-0.790) | 0.007                 | 36          | 38   | 0.517<br>(0.280-0.954) | 0.035                 | 38          | 36   | 0.637<br>(0.342-1.185) | 0.154                 |
| Sex                                     |             |      |                        |                       |             |      |                        |                       |             |      |                        |                       |
| Female                                  | 30          | 23   | 0.481<br>(0.189-1.221) | 0.124                 | 28          | 25   | 0.381<br>(0.160-0.909) | 0.030                 | 26          | 27   | 0.464<br>(0.200-1.076) | 0.073                 |
| Male                                    | 26          | 33   | 0.302<br>(0.140-0.652) | 0.002                 | 28          | 31   | 0.880<br>(0.372-2.086) | 0.772                 | 30          | 29   | 0.477<br>(0.211-1.078) | 0.075                 |
| Histologic grade                        |             |      |                        |                       |             |      |                        |                       |             |      |                        |                       |
| G1+G2                                   | 39          | 41   | 0.374<br>(0.178-0.787) | 0.010                 | 41          | 39   | 0.442<br>(0.222-0.880) | 0.020                 | 39          | 41   | 0.635<br>(0.322-1.254) | 0.191                 |
| G3+G4                                   | 17          | 15   | 0.495<br>(0.177-1.383) | 0.180                 | 15          | 17   | 0.644<br>(0.238-1.743) | 0.386                 | 17          | 15   | 0.303<br>(0.105-0.870) | 0.026                 |
| Radical resection <sup>b</sup>          |             |      |                        |                       |             |      |                        |                       |             |      |                        |                       |
| R0                                      | 36          | 30   | 0.532<br>(0.249-1.136) | 0.103                 | 29          | 37   | 0.589<br>(0.293-1.184) | 0.137                 | 31          | 35   | 0.712<br>(0.346-1.466) | 0.357                 |
| R1+Rx                                   | 20          | 24   | 0.378<br>(0.162-0.880) | 0.024                 | 26          | 18   | 0.368<br>(0.135-1.003) | 0.051                 | 25          | 19   | 0.305<br>(0.120-0.770) | 0.012                 |
| Radiation therapy <sup>c</sup>          |             |      |                        |                       |             |      |                        |                       |             |      |                        |                       |
| No                                      | 42          | 28   | 0.462<br>(0.241-0.885) | 0.020                 | 37          | 33   | 0.504<br>(0.271-0.938) | 0.031                 | 37          | 33   | 0.479<br>(0.254-0.904) | 0.023                 |
| Yes                                     | 11          | 19   | 0.258<br>(0.068-0.977) | 0.046                 | 11          | 19   | 0.612<br>(0.166-2.260) | 0.461                 | 12          | 18   | 0.725<br>(0.210-2.498) | 0.610                 |
| Targeted molecular therapy <sup>d</sup> |             |      |                        |                       |             |      |                        |                       |             |      |                        |                       |
| No                                      | 15          | 14   | 0.453<br>(0.187-1.095) | 0.079                 | 19          | 10   | 0.346<br>(0.125-0.956) | 0.041                 | 16          | 13   | 0.397<br>(0.158-0.995) | 0.049                 |
| Yes                                     | 38          | 35   | 0.477<br>(0.222-1.027) | 0.059                 | 30          | 43   | 0.668<br>(0.328-1.359) | 0.265                 | 33          | 40   | 0.571<br>(0.282-1.154) | 0.118                 |

| Group                 | No. of events | MST (days) | Crude HR<br>95% CI | Crude<br>P value | Adjusted HR<br>95% CI <sup>a</sup> | Adjusted<br>P value <sup>a</sup> |
|-----------------------|---------------|------------|--------------------|------------------|------------------------------------|----------------------------------|
| <i>GPC2+GPC3</i>      |               |            |                    |                  |                                    |                                  |
| 1                     | 21/26(80.8%)  | 278        | 1                  | 0.001            | 1                                  | <0.001                           |
| 2                     | 37/60(61.7%)  | 568        | 0.441(0.252-0.772) | 0.004            | 0.350(0.187-0.653)                 | 0.001                            |
| 3                     | 11/26(42.3%)  | 702        | 0.285(0.135-0.598) | 0.001            | 0.173(0.072-0.418)                 | <0.001                           |
| <i>GPC2+GPC5</i>      |               |            |                    |                  |                                    |                                  |
| I                     | 18/23(78.3%)  | 278        | 1                  | 0.001            | 1                                  | <0.001                           |
| II                    | 40/66(60.6%)  | 517        | 0.424(0.237-0.758) | 0.004            | 0.283(0.145-0.554)                 | <0.001                           |
| III                   | 11/23(47.8%)  | 702        | 0.253(0.116-0.552) | 0.001            | 0.141(0.057-0.353)                 | <0.001                           |
| <i>GPC3+GPC5</i>      |               |            |                    |                  |                                    |                                  |
| i                     | 26/38(68.4%)  | 393        | 1                  | 0.028            | 1                                  | 0.018                            |
| ii                    | 21/36(58.3%)  | 498        | 0.685(0.383-1.226) | 0.203            | 0.575(0.292- 1.129)                | 0.108                            |
| iii                   | 22/38(57.9%)  | 691        | 0.443(0.243-0.805) | 0.008            | 0.394(0.206-0.756)                 | 0.005                            |
| <i>GPC2+GPC3+GPC5</i> |               |            |                    |                  |                                    |                                  |
| A                     | 13/16(81.2%)  | 219        | 1                  | 0.028            | 1                                  | <0.001                           |
| B                     | 26/39(66.7%)  | 517        | 0.494(0.250-0.974) | 0.042            | 0.446(0.212-0.938)                 | 0.033                            |
| C                     | 23/42(54.8%)  | 592        | 0.269(0.130-0.560) | <0.001           | 0.176(0.076-0.406)                 | <0.001                           |
| D                     | 7/15(46.7%)   | 702        | 0.204(0.079-0.526) | 0.001            | 0.135(0.045-0.403)                 | <0.001                           |

Notes: Group 1:low *GPC2*+low *GPC3*; Group 2:low *GPC2*+high *GPC3* or high *GPC2*+ low *GPC3*; Group 3: high *GPC2*+high *GPC3*

Group I:low *GPC2*+low *GPC5*; Group II:low *GPC2*+high *GPC5* or high *GPC2*+low *GPC5*; Group III: high *GPC2*+high *GPC5*

Group i:low *GPC3*+low *GPC5*; Group ii:low *GPC3*+high *GPC5* or high *GPC3*+low *GPC5*; Group iii: high *GPC3*+high *GPC5*

Group A:low *GPC2*+low *GPC3*+low *GPC5*; Group B: high *GPC2*+low *GPC3*+low *GPC5* or low *GPC2*+high *GPC3*+low *GPC5* or low *GPC2*+low *GPC3*+high *GPC5*; Group C:high *GPC2*+high *GPC3*+low *GPC5* or high *GPC2*+low *GPC3*+high *GPC5* or low *GPC2*+high *GPC3*+high *GPC5*; Group D:high *GPC2*+high *GPC3*+high *GPC5*

<sup>a</sup>Adjusted for histologic grade, targeted molecular therapy, radiation therapy and radical resection.

Abbreviations: MST, median survival time; HR, hazard ratio; CI, confidence interval.

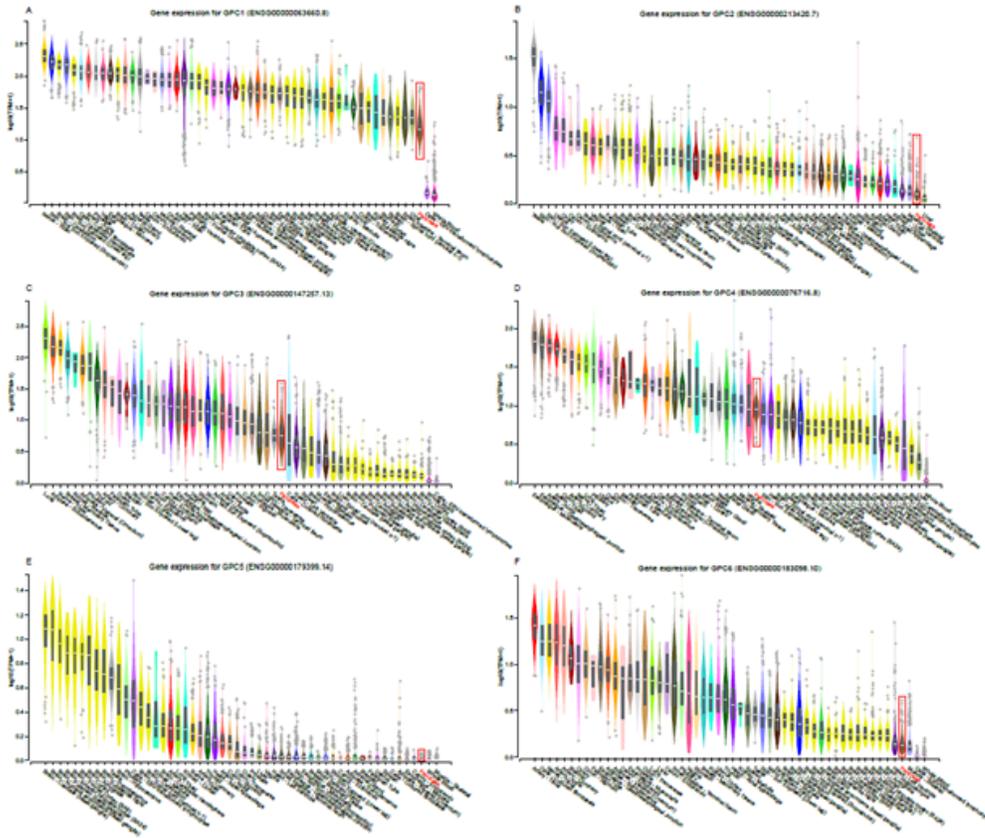
Table 4 Prognostic value of Glypican family genes in Gene Expression Omnibus database.

| Gene        | Samples<br>(n=48) | Crude HR<br>95% CI | Crude<br>P value | Adjusted HR<br>95% CI <sup>a</sup> | Adjusted<br>P value <sup>a</sup> |
|-------------|-------------------|--------------------|------------------|------------------------------------|----------------------------------|
| <i>GPC1</i> |                   |                    |                  |                                    |                                  |
| Low         | 24                | 1                  |                  | 1                                  |                                  |
| High        | 24                | 0.888(0.455-1.735) | 0.728            | 0.675(0.330-1.383)                 | 0.283                            |
| <i>GPC2</i> |                   |                    |                  |                                    |                                  |
| Low         | 24                | 1                  |                  | 1                                  |                                  |
| High        | 24                | 0.717(0.362-1.420) | 0.337            | 0.852(0.418-1.738)                 | 0.660                            |
| <i>GPC3</i> |                   |                    |                  |                                    |                                  |
| Low         | 24                | 1                  |                  | 1                                  |                                  |
| High        | 24                | 0.468(0.225-0.973) | 0.038            | 0.556(0.259-1.197)                 | 0.134                            |
| <i>GPC4</i> |                   |                    |                  |                                    |                                  |
| Low         | 24                | 1                  |                  | 1                                  |                                  |
| High        | 24                | 1.722(0.869-3.413) | 0.115            | 1.616(0.812-3.216)                 | 0.172                            |
| <i>GPC5</i> |                   |                    |                  |                                    |                                  |
| Low         | 24                | 1                  |                  | 1                                  |                                  |
| High        | 24                | 0.556(0.279-1.106) | 0.090            | 0.600(0.298-1.206)                 | 0.151                            |
| <i>GPC6</i> |                   |                    |                  |                                    |                                  |
| Low         | 24                | 1                  |                  | 1                                  |                                  |
| High        | 24                | 1.727(0.882-3.381) | 0.107            | 1.554(0.783-3.083)                 | 0.207                            |

<sup>a</sup>Adjusted for histologic grade.

Abbreviations: HR, hazard ratio; CI, confidence interval.

## Figures



**Figure 1**

Gene expression levels of Glypican family genes in normal human organ tissues. (A-F) Gene expression levels of Glypican1-6, respectively.

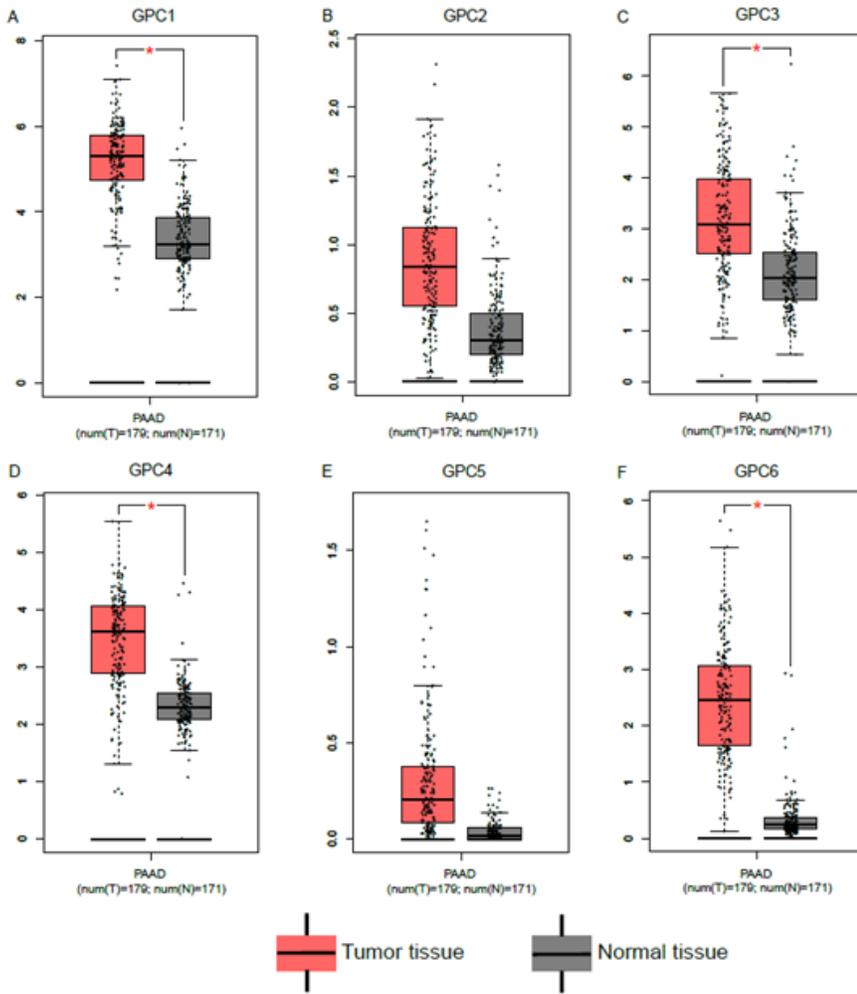


Figure 2

Gene level distribution of Glypican family genes in pancreatic ductal adenocarcinoma between tumor and normal tissues. (A-F) Gene level distribution of Glypican1-6 in pancreatic ductal adenocarcinoma between tumor and normal tissues, respectively. Notes: \*P < 0.05.

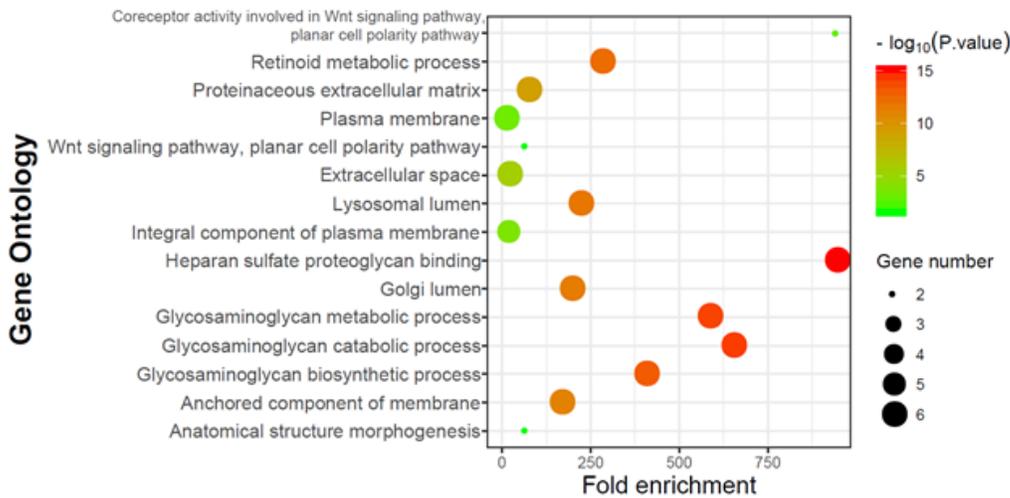


Figure 3

Function enrichment analysis of Gene Ontology for Glypican family genes completed by the Database for Annotation, Visualization, and Integrated Discovery.

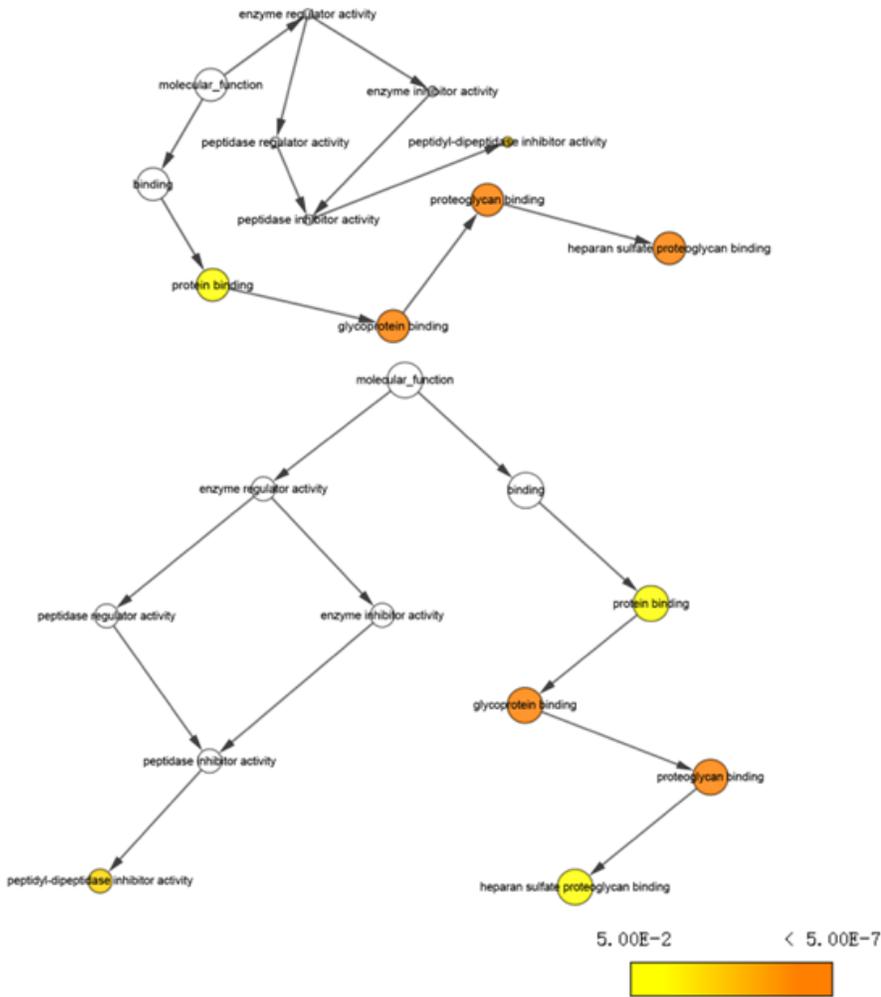
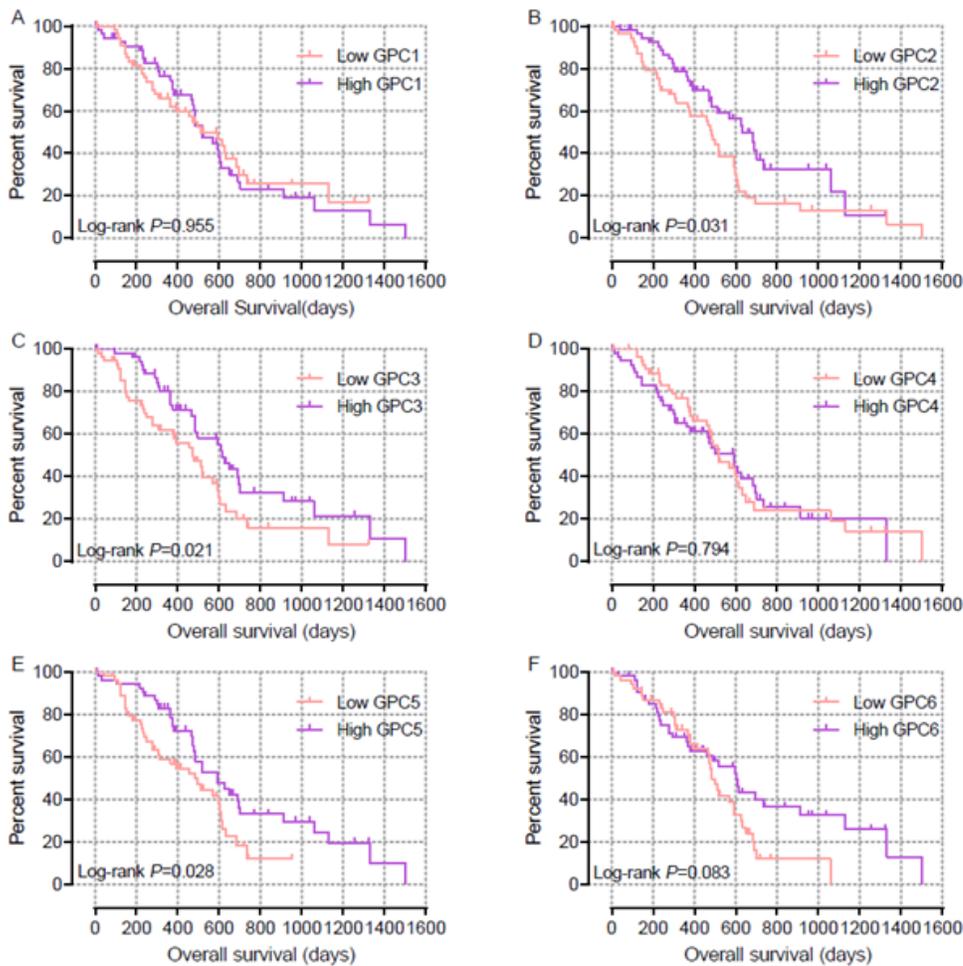


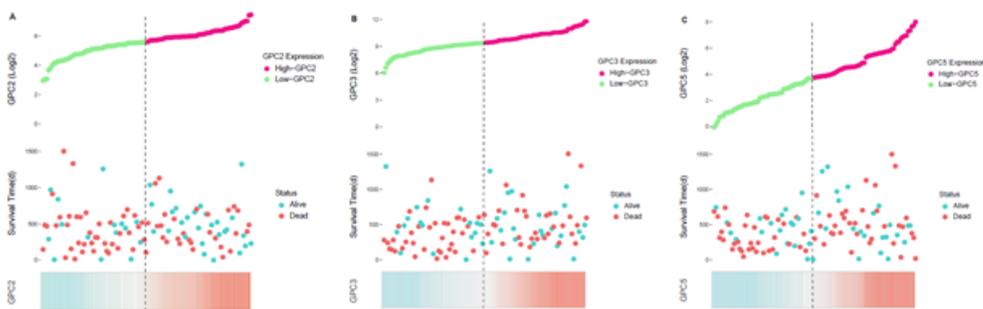
Figure 4

Functioning pathways of Glypican family genes carried out by Biological Network Gene Ontology in Cytoscape software.



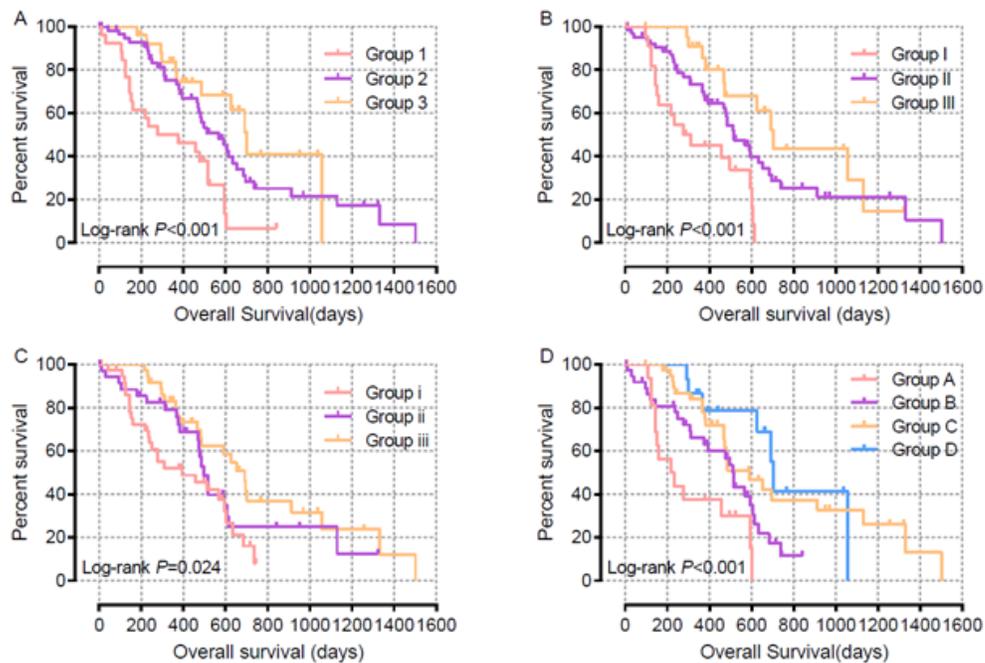
**Figure 5**

Kaplan-Meier survival curves of Glypican family genes for pancreatic ductal adenocarcinoma in The Cancer Genome Atlas database. (A-F) Kaplan-Meier survival curves of Glypican1-6, respectively.



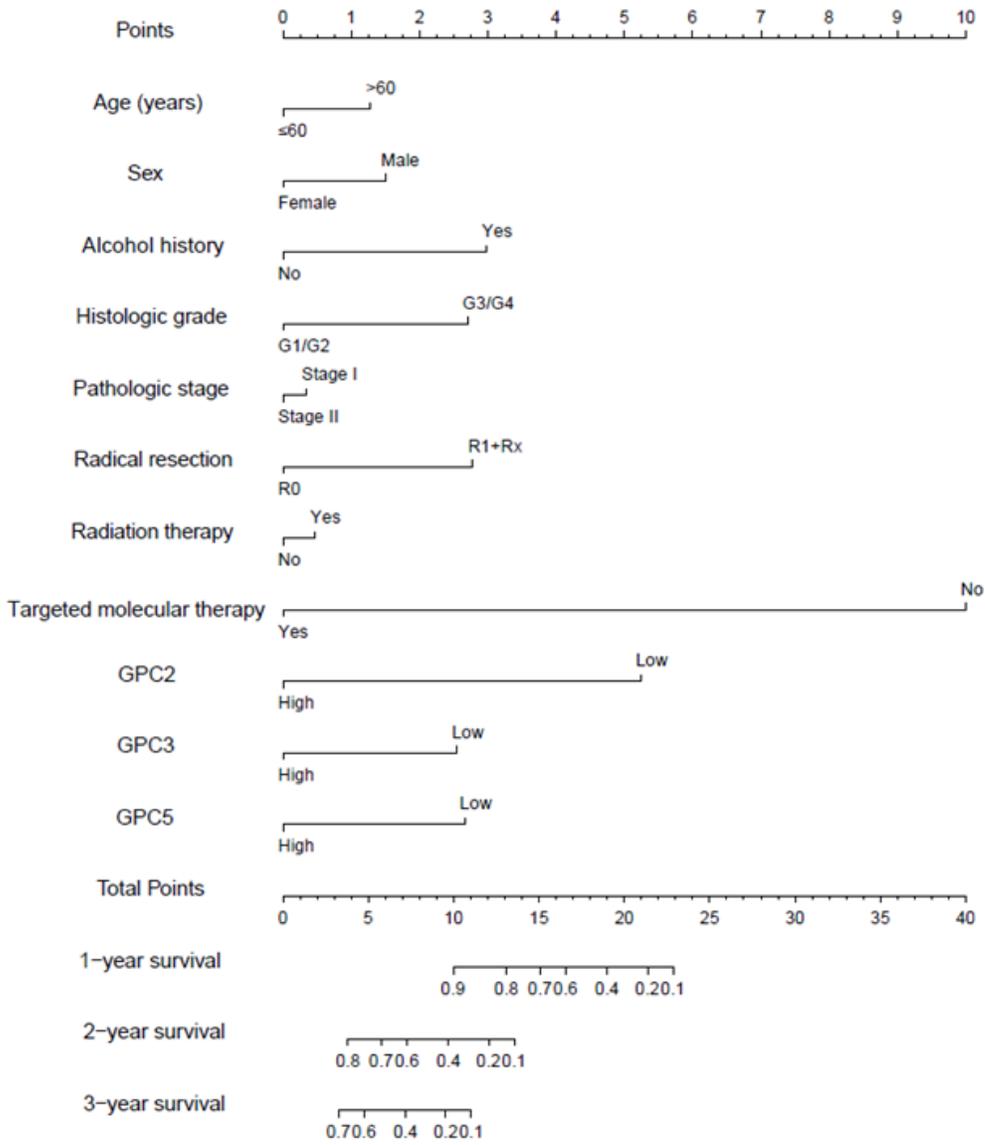
**Figure 6**

6 Prognostic models of Glypican2, Glypican3 and Glypican5 for pancreatic ductal adenocarcinoma in The Cancer Genome Atlas database. From top to bottom are expression values of these genes, survival status of patients and expression heatmaps of these genes at low and high expression levels. (A-C) Prognostic models of Glypican2, Glypican3, and Glypican5, respectively.



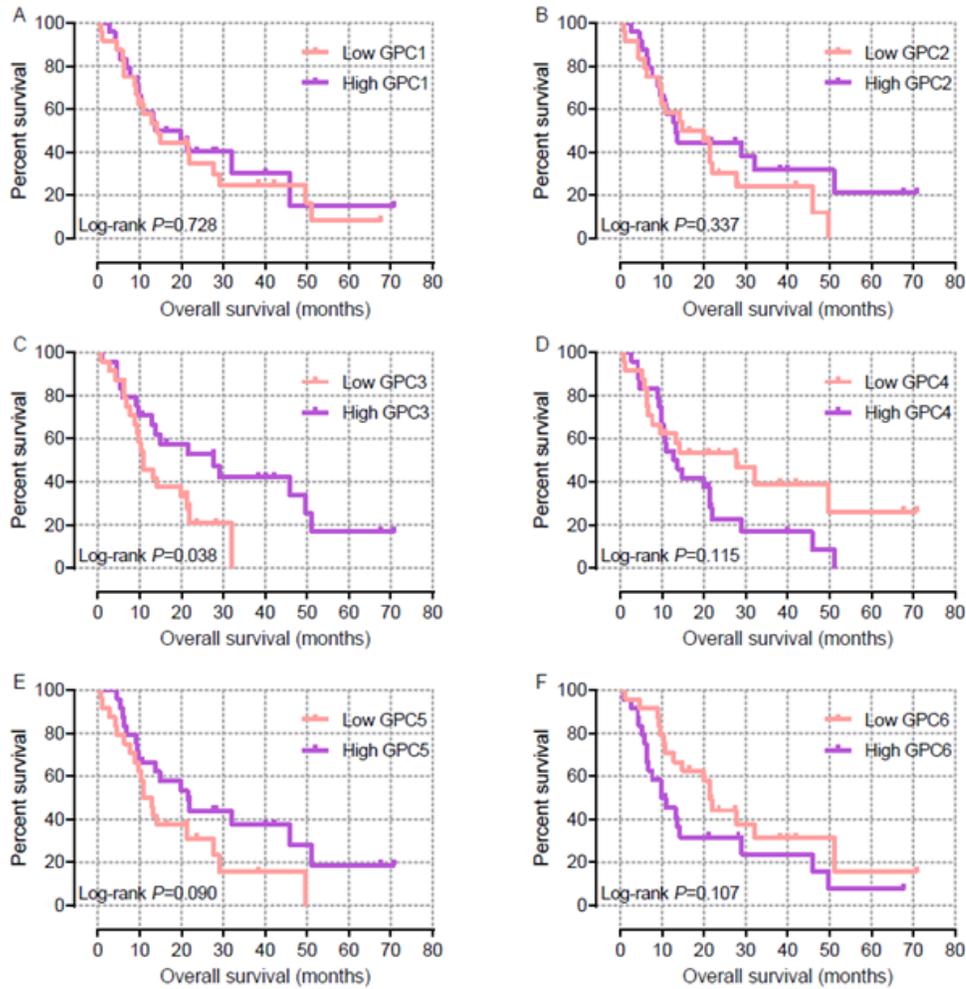
**Figure 7**

Survival curves of joint effects analysis of Glypican2, Glypican3 and Glypican5 in The Cancer Genome Atlas database. (A) Survival curve of Glypican2 and Glypican3; (B) Survival curve of Glypican2 and Glypican5; (C) Survival curve of Glypican3 and Glypican5; (D) Survival curve of Glypican2, Glypican3 and Glypican5.



**Figure 8**

Prognosis nomogram for predicting 1-, 2- and 3- overall survival.



**Figure 9**

Kaplan-Meier survival curves of Glypican family genes for pancreatic ductal adenocarcinoma in Gene Expression Omnibus database. (A-F) Kaplan-Meier survival curves of Glypican1-6, respectively.

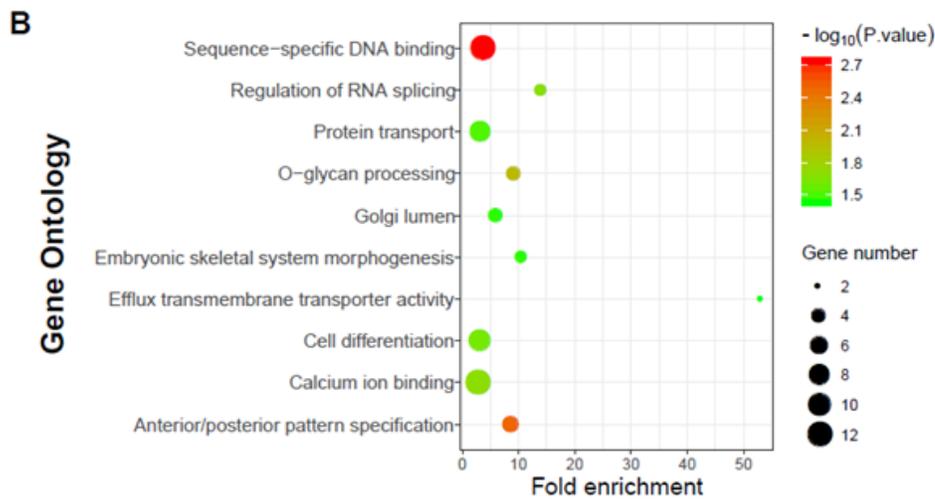
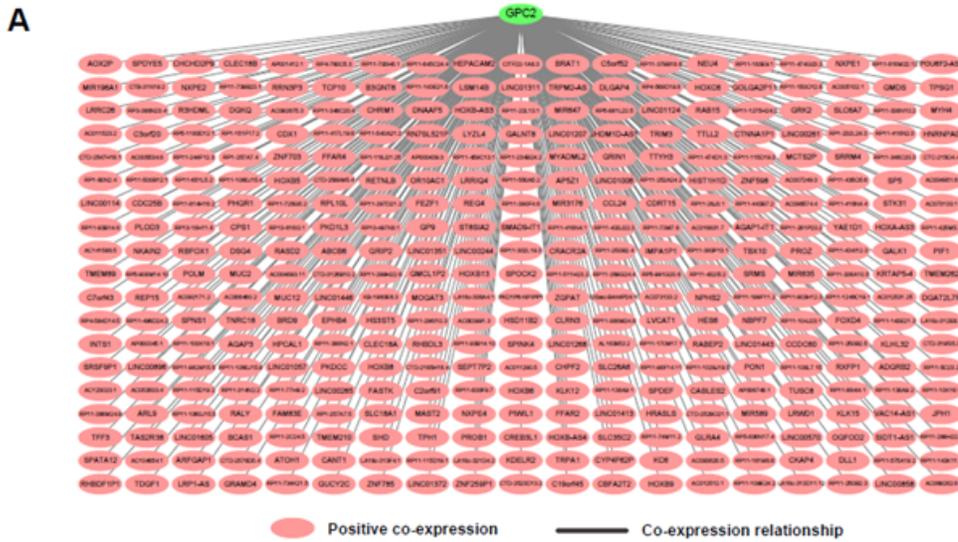
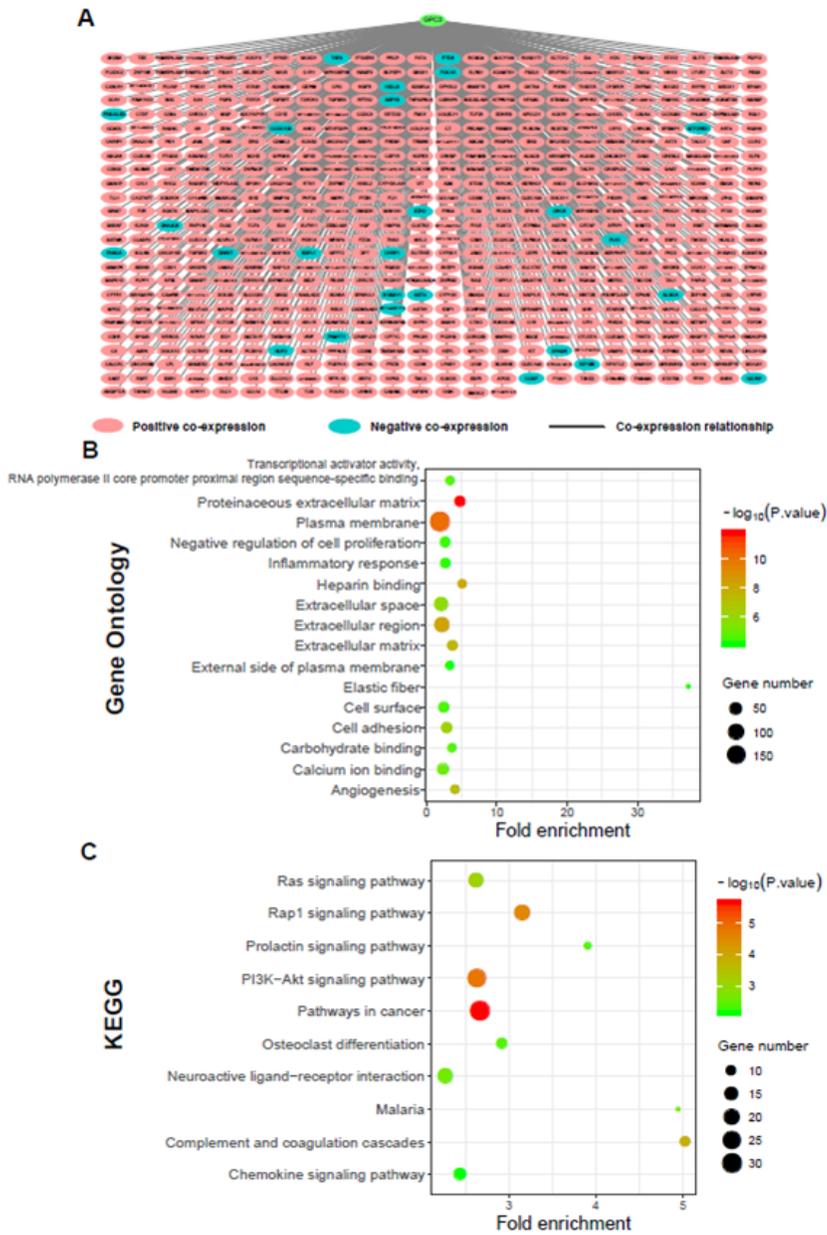


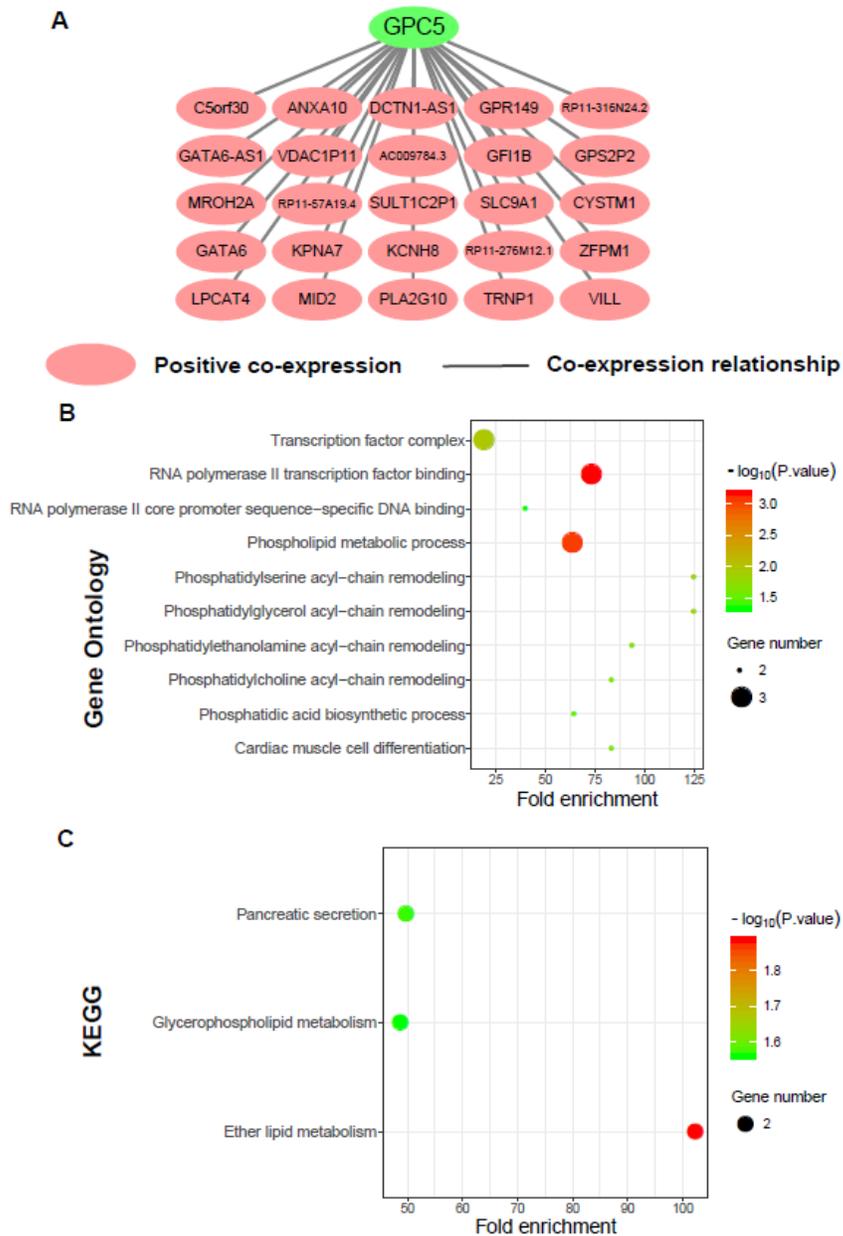
Figure 10

(A) Correlation network for Glypican2 and its co-expression genes in The Cancer Genome Atlas database. The pink nodes are genes correlated positively. (B) Function enrichment analysis of Gene Ontology for Glypican2 and its co-expression genes.



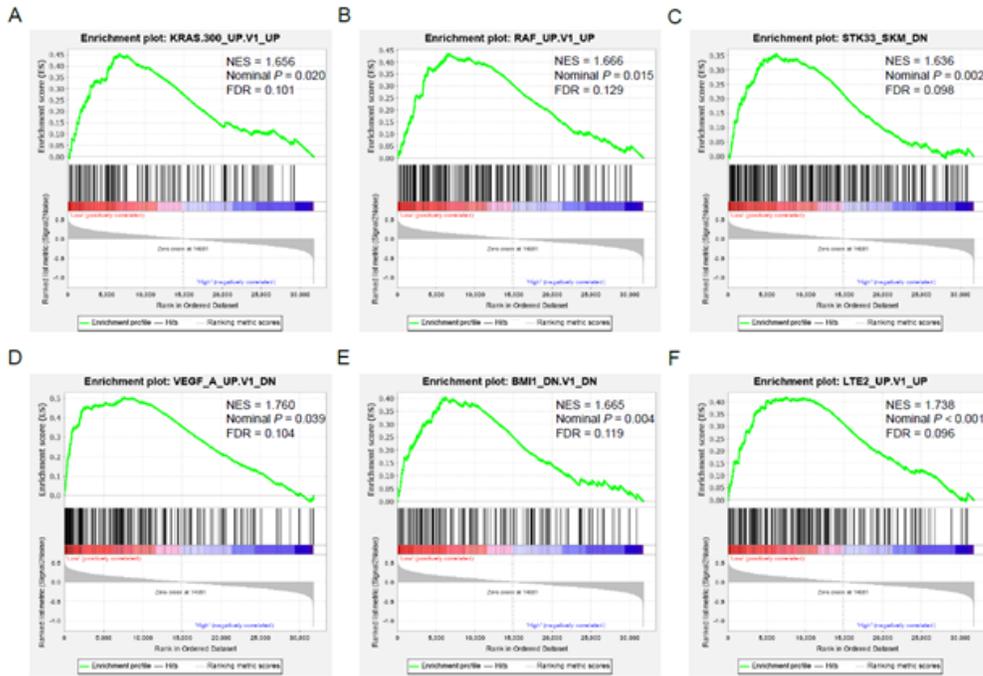
**Figure 11**

(A) Correlation network for Glypican3 and its co-expression genes in The Cancer Genome Atlas database. The pink nodes are genes correlated positively and the blue nodes are genes correlated negatively. (B) Function enrichment analysis of Gene Ontology for Glypican3 and its co-expression genes. (C) Function enrichment analysis of Kyoto Encyclopedia of Genes and Genomes for Glypican3 and its co-expression genes.



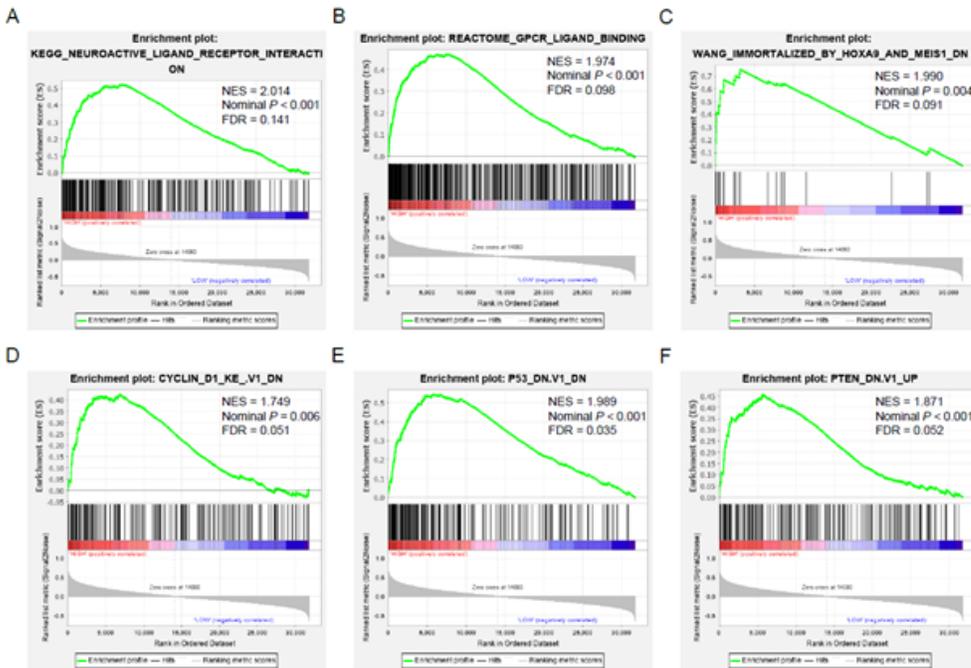
**Figure 12**

(A) Correlation network for Glypican5 and its co-expression genes in The Cancer Genome Atlas database. The pink nodes are genes correlated positively. (B) Function enrichment analysis of Gene Ontology for Glypican5 and its co-expression genes. (C) Function enrichment analysis of Kyoto Encyclopedia of Genes and Genomes for Glypican5 and its co-expression genes



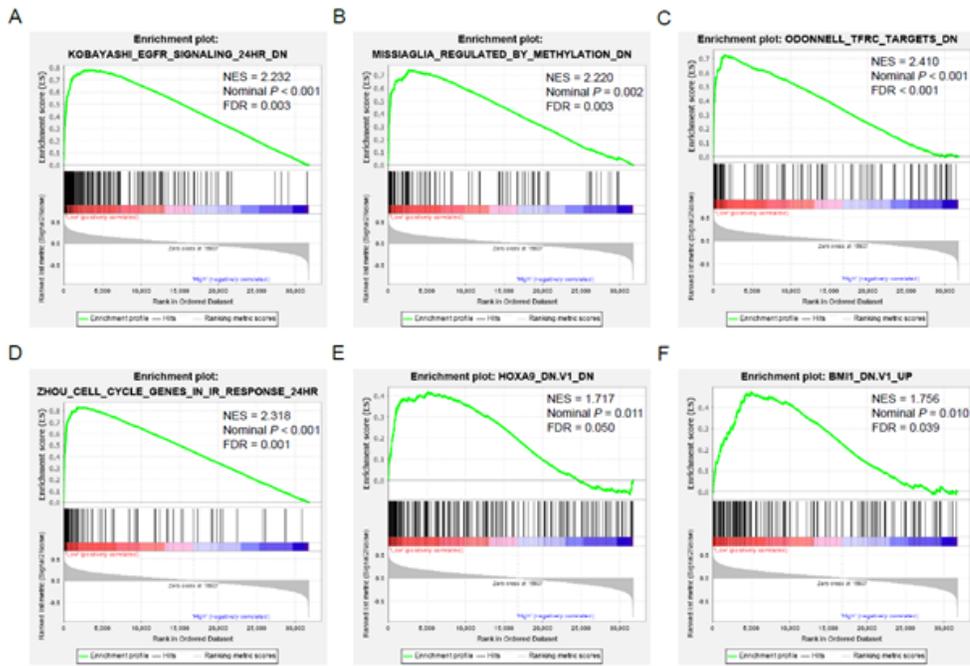
**Figure 13**

Gene Set Enrichment Analysis (GSEA) results of Glypican2 in The Cancer Genome Atlas database. (A-F) GSEA results of c6 reference for the group of low Glypican2 expression. NES, normalized enrichment score; FDR, false discovery rate.



**Figure 14**

Gene Set Enrichment Analysis (GSEA) results of Glypican3 in The Cancer Genome Atlas database. (A-C) GSEA results of c2 reference for the group of high Glypican3 expression; (D-F) GSEA results of c6 reference for the group of high Glypican3 expression. NES, normalized enrichment score; FDR, false discovery rate.



**Figure 15**

Gene Set Enrichment Analysis (GSEA) results of Glypican5 in The Cancer Genome Atlas database. (A-D) GSEA results of c2 reference for the group of low Glypican 5 expression; (E-F) GSEA results of c6 reference for the group of low GPC5 expression. NES, normalized enrichment score; FDR, false discovery rate.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table1Supp.xlsx](#)
- [Table2Supp.docx](#)
- [Table3Supp.docx](#)
- [Table4Supp.xlsx](#)
- [Table5Supp.xlsx](#)
- [Table6Supp.xlsx](#)
- [Table7Supp.xlsx](#)
- [Table8Supp.xlsx](#)
- [Table9Supp.xlsx](#)
- [Table10Supp.xlsx](#)
- [Table11Supp.xlsx](#)
- [Table12Supp.xlsx](#)
- [Table13Supp.xlsx](#)
- [Table14Supp.xlsx](#)
- [Table15Supp.xlsx](#)
- [Table16Supp.xlsx](#)