Original Article

Identification of G-protein signaling modulator 2 as a diagnostic and prognostic biomarker of pancreatic adenocarcinoma: an exploration of its regulatory mechanisms

Xintong Zhou[1], Shengchun Dang[2], Huaji Jiang[2], Min Gu[3]

[1]Department of General Surgery, The Affiliated Zhangjiagang Hospital of Soochow University, Zhangjiagang, China; [2]Department of General Surgery, Affiliated Hospital of Jiangsu University, Zhenjiang, China; [3]Department of Oncology, Zhenjiang Hospital of Traditional Chinese and Western Medicine, Zhenjiang, China

Contributions: (I) Conception and design: M Gu; (II) Administrative support: M Gu; (III) Provision of study materials or patients: S Dang; (IV) Collection and assembly of data: X Zhou, H Jiang; (V) Data analysis and interpretation: X Zhou, H Jiang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Min Gu. Zhenjiang Hospital of Traditional Chinese and Western Medicine, Zhenjiang, China. Email: dangscjda@163.com.

Background: Pancreatic adenocarcinoma (PAAD) has a high rate of mortality. Unfortunately, it is difficult to diagnosis. This study aimed to develop a more in-depth understanding of the disease.

Methods: A total of 177 patients with PAAD were recruited from The Cancer Genome Atlas (TCGA) database. Microarray analysis was performed to identify differentially expressed genes (DEGs) in PAAD. The microarray data were adapted to the ingenuity pathway analysis (IPA) for annotation and visualization, followed by protein-protein interaction (PPI) network analysis. *In vitro* transwell migration assays were conducted to explore the molecular and functional characteristics of pancreatic adenocarcinoma cells (PANC-1) with stable low expression of G-protein signaling modulator 2 (*GPSM2*). Expression of *GPSM2* and the associated hub genes were detected by reverse transcription-quantitative polymerase chain reaction (qPCR).

Results: The overexpression of *GPSM2* was proved in PAAD, as compared with the healthy tissues, as well as its correlation with history of chronic pancreatitis, T stage, TNM stage and tumor grade. We described it as an independent prognostic factor and found that it could influence the infiltration of immune cells in the tumor microenvironment. Silencing of *GPSM2* restrained the and migration of the cells. Microarray analysis identified 1,631 DEGs in PAAD cells. The PPI network analysis identified hub genes including *CD44*, *ITGB1*, *ITGB5*, *ITGA2*, *ITGA5*, *AKT1*, *EGFR*, *NRAS* and *MAP2K1*, and their relationship with *GPSM2* was confirmed by qPCR.

Conclusions: *GPSM2* is a novel prognostic factor and therapeutic target for PAAD. *GPSM2* promoted the migration of pancreatic adenocarcinoma cells .Targeting *GPSM2* and its downstream genes may prolong the survival time of patients with PAAD.

Keywords: G-protein signaling modulator 2 (GPSM2); pancreatic carcinoma; bioinformatics analysis; PANC-1 cells; migration

Submitted Jan 28, 2021. Accepted for publication May 26, 2021.
doi: 10.21037/jgo-21-224
View this article at: https://dx.doi.org/10.21037/jgo-21-224

Introduction

Worldwide, the number of newly diagnosed cases of pancreatic cancer and the associated mortality continue to escalate every year (1). While pancreatic carcinoma is relatively rare, it is one of the most common causes of cancer mortality in Western countries (2). As invasion or distant metastasis was not rare in early stage patients, only 10–20% of patients had surgical resection of the pancreas,

© Journal of Gastrointestinal Oncology. All rights reserved.

while the 5-year survival rate of patients undergoing surgical resection is only 23.4%, and the number drops to 6% for patients of all stages (3,4). A recent study suggested that by 2018, pancreatic cancer had be the third leading cause of cancer death in USA (5). Therefore, reducing the invasion and metastasis of pancreatic cancer has great significance for the treatment and prognosis of these patients.

G-protein signaling modulator 2 (*GPSM2*) belongs to a family of proteins involved in G-protein signal conditioning, and has received considerable attention in recent years (6,7). Mutations in the *GPSM2* gene have been associated with poor development of the auditory system, and can lead to autosomal recessive nonsyndromic deafness (8). *GPSM2* was originally found to play a vital role in ensuring the correct orientation of cell spindles and symmetric cell division (9,10). One study discovered that overexpression of *GPSM2* suppressed cell growth and influenced chromosomal segregation in breast cancer (7). Pishas *et al.* found that inhibition of *GPSM2* by XI-006, a 4-nitrobenzofuroxan derivative, had a positive therapeutic effect on Ewing sarcoma (11). Our previous studies demonstrated that *GPSM2* is overexpressed in pancreatic cancer, and its expression is closely related to the T stage, TNM stage, tumor grade, and prognosis of pancreatic cancer patients (12). However, the role of GSPM2 in the growth and migration of pancreatic cancer, as well as its specific regulatory mechanism, has not been clarified.

High throughput sequencing methods are growing in popularity due to its success in early diagnosis and the prognostic evaluation of cancers (13). Since the sample size of our previous study was not large (54 samples), the results from the latter study were validated using a public database. Univariate and multivariate cox regression analyses were used to explore the factors associated with survival time in pancreatic carcinoma patients. *GPSM2* was identified and shown to influence the microenvironment of pancreatic carcinomas. Since there's no research on using microarray technology to explore the regulatory mechanism of GPSM2 at cellular whole level, stable low *GPSM2* expressing pancreatic cancer cells were constructed and microarray analysis was performed to identify differentially expressed genes (DEGs). The DEGs were then examined using ingenuity pathway analysis (IPA) and protein-protein interaction (PPI) network analysis for in-depth exploration of the associated regulatory mechanisms.

We present the following article in accordance with the REMARK reporting checklist (available at https://dx.doi.org/10.21037/jgo-21-224).

Methods

Online analysis of gene expression and prognosis

GEPIA (Gene Expression Profiling Interactive Analysis; http://gepia.cancer-pku.cn/index.html) is a novel interactive web server developed for processing the RNA sequencing expression data of 9,736 cancers and 8,587 normal tissues obtained from the TCGA and GTEx projects by applying a standard processing pipeline (14). It has shown great promise in data analysis and presentation. The differential expression of *GPSM2* between cancer tissues and normal tissues has previously been shown in a boxplot or an anthropotomical heatmap constructed using the gganatogram package (15). Preliminary survival analysis of hub genes was performed and displayed on a Kaplan-Meier curve.

TCGA data of pancreatic adenocarcinomas (PAADs)

UCSC is a data visualization and analysis website for exploring various sequencing data and related clinical phenotypic annotations from TCGA, PCAWG, Pan-Cancer Atlas, ICGC, GDC, and GTEx (16). All available TCGA data on the PAAD project were obtained from the UCSC Cancer Genomics Browser (https://genome-cancer.ucsc.edu/). In September 2017, there were RNA sequencing data on 182 PAAD samples, which included 173 single tumor samples, 4 pairs of PAAD and adjacent non-tumor pancreatic tissues, and clinical data including survival time and survival status records of 185 patients. Prior to analysis, raw sequencing data were normalized and annotated by gene mapping. After screening, a total of 177 pancreatic cancer samples were picked up, including *GPSM2* expression levels and clinical data. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Evaluation of the tumor microenvironment and immune infiltration

The amount of immunological and stromal infiltration in TCGA pancreatic carcinoma was evaluated using ESTIMATE in R software (17). The population of tumor-infiltrating immune cells in heterogeneous tissues from the transcriptomic data was quantified using the analytical platform TIMER at https://cistrome.shinyapps.io/timer (18). TIMER is a comprehensive resource for systematic analysis of immune infiltration across diverse

© Journal of Gastrointestinal Oncology. All rights reserved.

cancer types. The infiltration abundances of six types of immune cells (B cells, CD4+ T cells, CD8+ T cells, neutrophils, macrophages, and dendritic cells) were estimated by using a bioinformatic algorithm on the basis of rigorous statistics.

Cell culture and transfection

To induce stable low expression of *GPSM2* in the PANC-1 pancreatic adenocarcinoma cell line (Institute of Cell Biology, Chinese Academy of Sciences, Shanghai, China), short hairpin (sh)RNA for *GPSM2* was cloned into the GV248 vector (Genechem, Shanghai) and verified by sequencing. Three different target sequences for *GPSM2* (shGPSM2#1, shGPSM2#2, shGPSM2#3), as well as the scrambled sequence as a negative control (shNC), were designed and are shown in Table S1. The PANC-1 cells were cultivated with RPMI 1640 medium (Hyclone, Logan, UT, USA) supplemented with 10% fetal bovine serum (FBS) and maintained in a fully humidified atmosphere at 37 ℃ and 5% CO_2. The culture medium was refreshed daily. Log-phase PANC-1 cells were harvested and transferred into six-well plates with serum-free medium, at a density of 5×10^5 cells per well. When confluency reached 70%, the cells were transfected with 4 μL of shGPSM2 or shNC sequence by applying polybrene (Genechem, Shanghai) according to the manufacturer's protocol. The cells were transferred to complete culture medium after incubation for 16 hours. The silencing efficiency of shGPSM2 was evaluated by qPCR. Mixed populations of GPSM2-silenced cells were then used for microarray analysis and construction of an orthotopic model for future use.

Transwell migration assay

For the invasion assay, cells were cultured with serum-free medium for 24 hours after transfection. Untransfected cells (Blank group), cells transfected with the scrambled sequence (negative control, shNC group) and shGPSM2#2 sequence (shGPSM2#2 group) were then harvested and resuspended at a density of 2×10^5/mL in 100 μL of serum-free medium and transferred to the upper transwell chamber (Corning Costar). Next, 700 μL of complete medium containing 10% FBS was added to the lower chamber, and cells were incubated for 24 hours at 37 ℃. After migration for 24 hours, the microporous membrane and the migrated cells were fixed with 4% paraformaldehyde solution for 0.5 hours and stained with 2% crystal violet (Google biotechnology, Wuhan, China) for 20 minutes. Finally, a microscope (Leica) was used for counting cells and acquiring images. Each group was repeated 3 times.

Microarray and IPA

Total RNA Isolation Reagent (SuperfecTRI) was used to extract RNA from PANC-1 cells according to the manufacturer's protocol. The quality of RNA was verified by Agilent 2100 (Agilent), using an RNA 6000 Nano Kit (Agilent). After quality inspection, the RNA was reverse transcribed to cDNA, which was then used as a template for *in vitro* transcription using the GeneChip 3' IVT Express Kit. Biotin-labeled RNA was purified and fragmented before the next step of hybridizing to Affymetrix PrimeView Human Gene Expression Array (Affymetrix) for 16 hours. The microarray was washed and stained using GeneChip Fluidics Station 450 (Affymetrix). Fluorescent images were obtained using a GeneChip Scanner 3000 (Affymetrix).

Before analysis, raw data were adjusted by mean normalization and the lowest 20% range of signal intensity of the probe was filtered out as background noise, as well as those with a coefficient of variation greater than 25% (19-21). Sequencing data were then applied to the limma package for DEGs identification based on the linear model of empirical bayesian distribution. The Benjamini-Hochberg method was applied for the correction of significant differences [false discovery rate (FDR)] (22). Genes with large variations of expression (|Fold Change| >1.5, FDR <0.05) were selected as DEGs and applied to the Ingenuity Pathways Knowledge Base (IPA Ingenuity Systems; Qiagen, Redwood City, CA) for biological information mining. According to the changes in the expression values and the correlation between them, the obtained DEGs were mapped to the predicted functional modules and shown in graph form.

Reverse transcription-quantitative PCR

The steps of RNA extraction and purification are described above. Total RNA was reverse transcribed using the RT-PCR Quick Master Mix Kit (TOYOBO, Osaka, Japan) according to the manufacturer's protocol. After extraction, the RNA was reverse transcribed to cDNA with an oligo dT primer (Sangon Biotech, Shanghai) and M-MLV-RTase (Promega). The gene expression

© Journal of Gastrointestinal Oncology. All rights reserved.

levels were analyzed by real-time qPCR using SYBR Green assays (TaKaRa, Otsu, Japan) on a real-time PCR System (Roche). Glyceralde-3-hydephosphate dehydrogenase (*GAPDH*) was used as an internal control, and the equation $2^{-\Delta\Delta Ct}$ was used to describe the relative expression of the gene. The specific forward primer for *GPSM2* was CCAGAGAGCTTAATGACAAGG and the reverse primer was GGACCAGGGCAACCAAAACT. The forward primer for *GAPDH* was TGACTTCAACAGCGACACCCA and the reverse primer was CACCCTGTTGCTGTAGCCAAA.

The PPI network and hub gene analysis

The PPI networks can analyze the specific molecular mechanisms of intracellular activity. DEGs were mapped to a PPI network by using the Search Tool for the Retrieval of Interacting Genes (STRING), which is an online database designed for visualizing PPI information (23). A confidence score of 0.4 was set as the cut-off criterion. We used the CentiScaPe plugin in Cytoscape software to calculate the node degrees, and the top 30 DEGs of node degrees were selected as hub genes (24,25). Visualization of the PPI network was also performed using Cytoscape.

Statistical analysis

IBM SPSS22.0 statistics was applied for statistical calculations. Single comparisons of measurement data between two groups were performed using the Wilcoxon Rank-Sum test, while the Kruskal-Wallis test was adopted for more groups. Survival analysis using overall survival (OS) and disease-free survival (DFS) curves was conducted according to the Kaplan-Meier method and log-rank tests were used to assess the statistical significance of differences. The prognostic significance of clinicopathologic parameters and *GPSM2* expression was determined using univariate and multivariate Cox-regression analysis via a backward stepwise procedure. Differences were considered statistically significant when the P values were less than 0.05 (*P<0.05). The relative risk (RR), 95% confidence intervals (CIs), and log rank P values were computed. OS and DFS data were acquired from the clinical data of pancreatic cancer samples on the TCGA database. Graphs and curves were constructed by GraphPad Prism 7 (GraphPad Software, San Diego, CA, USA).

Results

GPSM2 is upregulated in pancreatic adenocarcinoma (PAAD) and is correlated with clinicopathological features of PAAD

By using the GEPIA online tool (accessed on December 20, 2017), the differential expression of *GPSM2* was compared between primary pancreatic carcinomas and normal samples. There was a remarkably higher expression of *GPSM2* in PAAD compared to the adjacent healthy tissues (*Figure 1A*, P<0.001). Our previous studies confirmed this at the protein level (12). In addition, *GPSM2* was overexpressed in colon adenocarcinoma, rectum adenocarcinoma, and stomach adenocarcinoma (*Figure 1B*, Fold Change >1), while the expression was low in thyroid carcinoma (Fold Change <1).

Furthermore, to determine whether *GPSM2* mRNA expression levels were associated with pancreatic carcinoma progression, the relationship between *GPSM2* and the clinicopathological characteristics of pancreatic carcinoma were explored by applying data from the TCGA database. The *GPSM2* expression profiles of 185 samples, as well as the clinical information, were obtained from the UCSC website (https://xena.ucsc.edu/). A total of 177 PAAD samples were analyzed through rigorous selection, and the mean age of the sample cohort was 65 years (35–88 years). *GPSM2* expression was markedly associated with a history of chronic pancreatitis (P=0.004), T stage (P=0.033), TNM stage (P=0.045), and tumor grade (P<0.001) (*Table 1*), but poorly associated with age, gender, family history of cancer, history of diabetes, tumor size, N stage, and M stage (P>0.05). These results were basically consistent with our earlier findings (12).

Association of GPSM2 expression with clinical prognosis and the tumor microenvironment

Analysis of the 177 pancreatic cancer samples showed a median OS time of 18.9 months (range, 0 to 91.4 months). A total of 92 patients died from the disease and 85 patients survived. The median DFS time was 16.4 months (range, 0 to 91.4 months). A total of 80 patients suffered a recurrence or progression of disease. Kaplan-Meier survival analysis and log-rank tests showed that patients with high *GPSM2* expression suffered poorer OS and DFS compared to patients with low *GPSM2* expression (*Figure 1C*, P=0.002; *Figure 1D*, P=0.008). The univariate

© Journal of Gastrointestinal Oncology. All rights reserved.

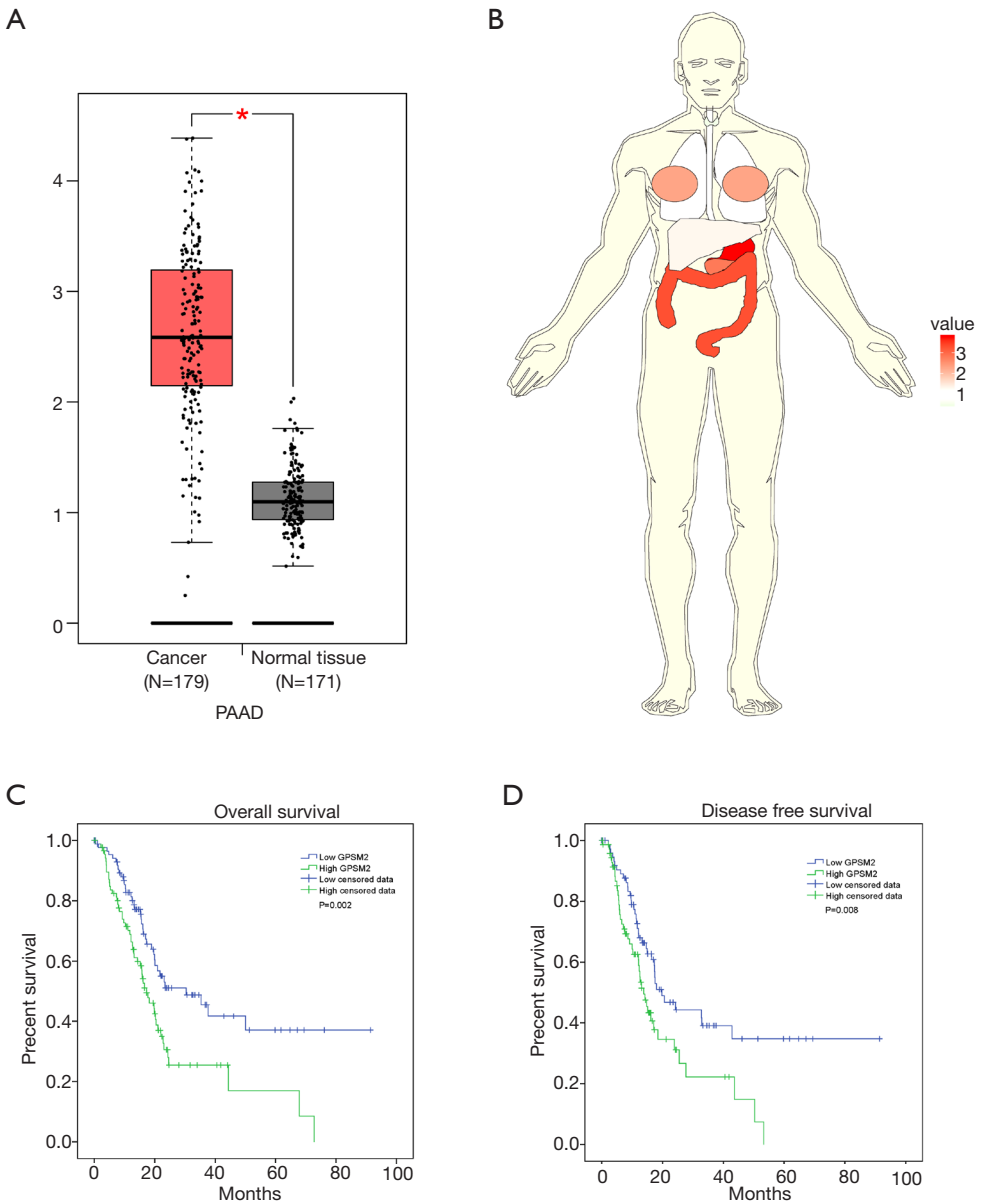

Figure 1 Expression of GPSM2 in pancreatic carcinoma. (A,B) Abundance and fold change of GPSM2 expression in primary cancer and adjacent normal tissues. Green: fold change <1; Red: fold change >1; *, P<0.05. (C,D) Survival curves showing the overall survival (C) and disease-free survival (D) of patients with high or low GPSM2 expression. GPSM2, G-protein signaling modulator 2; PAAD, pancreatic adenocarcinoma.

analysis of OS in correlation with the clinicopathologic parameters and *GPSM2* expression is shown in *Table 2*. In addition to *GPSM2* protein expression (P<0.001), OS was also associated with age, T stage, N stage, and tumor grade, (P<0.05), but not with other clinicopathologic factors (P≥0.05). Multivariate analysis revealed that *GPSM2* expression was an independent prognostic factor for OS in PAAD patients, in addition to age and N stage (*Table 2*).

The sequencing data was further analyzed using the ESTIMATE package. Samples with higher *GPSM2* expression were characterized by less dispersed infiltration of populations of immune cells (P=0.018) and stroma cells (P>0.05) (*Table 3*). Consistent with the lower stromal or non-tumor content, samples with higher *GPSM2* expression

© Journal of Gastrointestinal Oncology. All rights reserved.

Table 1 The expression of GPSM2 was correlated with the clinicopathologic features of pancreatic carcinoma

Characteristics	n	GPSM2 expression (mean ± standard deviation)	P value
Gender			0.915
Male	97	14.87±0.68	
Female	80	14.87±0.70	
Age (y)			0.949
≤60	58	14.88±0.62	
>60	119	14.87±0.72	
Family history of cancer			0.699
No	47	14.85±0.81	
Yes	63	14.81±0.74	
History of chronic pancreatitis			0.004
No	128	14.79±0.70	
Yes	13	15.42±0.68	
History of diabetes			0.195
No	108	14.87±0.70	
Yes	38	14.79±0.78	
Tumor size (cm)			0.343
≤2	10	14.64±0.84	
2–4	100	14.84±0.64	
>4	54	14.96±0.77	
T stage			0.044
T1	7	14.45±0.90	
T2	24	14.62±0.76	
T3, 4	144	14.95±0.65	
N stage			0.427
N0	49	14.82±0.74	
N1	123	14.93±0.64	
M stage			0.617
M0	79	14.87±0.63	
M1	4	15.03±0.49	
TNM stage			0.046
I	21	14.46±0.81	
II	146	14.95±0.65	
III, IV	7	14.76±0.53	
Grade			<0.001
G1	31	14.34±0.74	
G2	94	15.01±0.61	
G3, 4	50	14.94±0.62	

GPSM2, G-protein signaling modulator 2.

© Journal of Gastrointestinal Oncology. All rights reserved.

Table 2 Univariate and multivariate cox regression analyses on GPSM2 level and clinicopathologic features for prognosis of pancreatic carcinoma

Characteristics	Univariate analysis			Multivariate analysis		
	RR	95% CI	P value	RR	95% CI	P value
GPSM2	1.667	1.258–2.208	<0.001	1.412	1.030–1.936	0.032
Gender	0.820	0.544–1.235	0.342			
Age	1.029	1.008–1.051	0.007	1.024	1.002–1.047	0.029
Family history of cancer	1.118	0.650–1.922	0.686			
History of chronic pancreatitis	1.178	0.563–2.466	0.664			
History of diabetes	0.927	0.532–1.614	0.788			
Tumor size	1.006	0.899–1.127	0.916			
T stage	1.561	1.008–2.419	0.046			
N stage	2.105	1.253–3.537	0.005	2.067	1.213–3.521	0.008
M stage	1.028	0.247–4.298	0.970			
Stage	1.312	0.898–1.920	0.161			
Grade	1.448	1.089–1.926	0.011			

CI, confidence interval; GPSM2, G-protein signaling modulator 2, RR: relative risk.

Table 3 Immunocyte score correlated with GPSM2 expression in the tumor microenvironment of pancreatic carcinoma

	GPSM2		P value
	High (n=88)	Low (n=89)	
Stromal score	581.25	720.40	0.129
Immune score	1,040.53	1,253.61	0.018
Estimate score	1,621.78	1,974.01	0.035

GPSM2, G-protein signaling modulator 2.

had a higher tumor content, which can be inferred by the estimate score (*Table 3*, P=0.034). Immunological infiltration estimated with TIMER confirmed that tumors with lower expression of *GPSM2* were characterized by a higher infiltration of macrophage cells (P<0.001), while infiltration of CD8+ T cells (P=0.01) and dendritic cells were lower (*Table 4*, P=0.011).

Silencing of GPSM2 expression decreased the migration ability of pancreatic cancer cells

To further examine the association between *GPSM2* and overall patient survival, *GPSM2* was silenced in PANC-1 cells using lentiviral-mediated RNA interference (RNAi). As presented in *Figure 2A*, the mRNA levels of *GPSM2* were decreased by 88.5% in PANC-1 cells after transfection with the shGPSM2#2 sequence (P=0.0063). To examine the biological functions of *GPSM2* in PAAD cell migration, transwell assays were performed after transfecting cells with the lentiviral vector. Transwell assays demonstrated that silencing *GPSM2* markedly decreased the numbers of migrating cells compared to cells transfected with the scrambled sequence (shNC) and untransfected cells (Blank group) (*Figure 2B*,*C*,*D*,*E*, P<0.001). These results suggested that *GPSM2* facilitated tumor migration in pancreatic cancer cells.

Microarray analysis and identification of DEGs

Microarray analysis for the mixed population of *GPSM2*-

© Journal of Gastrointestinal Oncology. All rights reserved.

Table 4 Immunological infiltrating status of six types of cells and their association with GPSM2

	GPSM2		P value
	High (n=88)	Low (n=89)	
B cells	0.133	0.123	0.056
CD4 T cells	0.100	0.129	0.094
CD8 T cells	0.227	0.147	0.010
Neutrophils	0.098	0.110	0.304
Macrophage	0.039	0.169	0.000
Dendritic cells	0.446	0.358	0.011

GPSM2, G-protein signaling modulator 2.

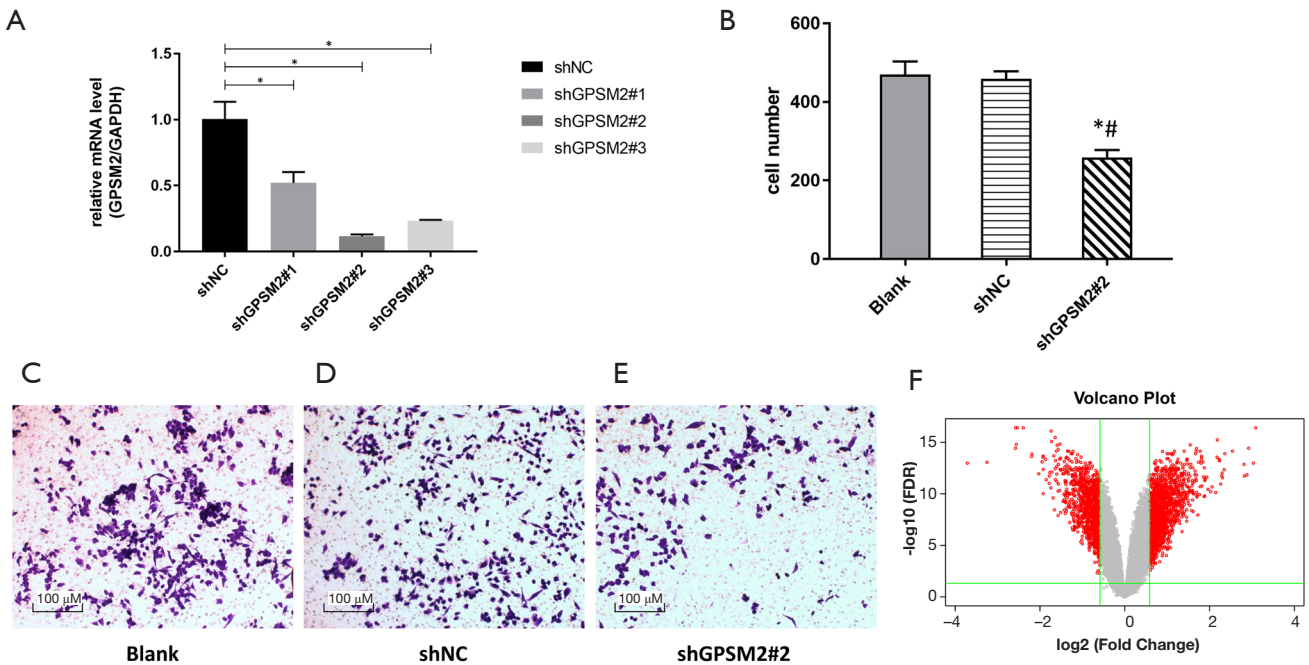

Figure 2 Silencing of GPSM2 expression in PANC-1 cells. (A) Relative expression of GPSM2 mRNA in cells transfected with shGPSM2 (#1, #2, and #3) and cells transfected with the scrambled sequence (negative control, shNC) (P<0.05). (B,C,D,E) Transwell migration assays of untransfected cells (Blank group), cells transfected with the scrambled sequence (negative control, shNC group), and cells transfected with shGPSM2#2 (shGPSM2#2 group). The graph (B) shows the average colonies per group (*, P<0.05 *vs.* shNC; #, P<0.05 *vs.* Blank). Staining method: 2% crystal violet for 20 minutes. (F) Volcano plot constructed from genes detected and filtered by microarray analysis. Red plot, |Fold Change| >1.5 with FDR <0.05; grey plot: |Fold Change| <1.5. GPSM2, G-protein signaling modulator 2; GAPDH, Glyceraldehyde 3-phosphate dehydrogenase; sh, short hairpin; NC, negative control; FDR, false discovery rate.

silenced PANC-1 cells was performed. The gene expression profiles between cells transfected with the scrambled sequence (shNC) and shGPSM2#2 sequence were compared (Figure S1). By comparing the RNAseq read counts of the various genes and subsequently applying the cut-off criteria, 1,631 genes were identified as DEGs, including 1,039 upregulated and 592 downregulated genes. Genes with large variations of expression [Fold Change >1.5, false discovery rate (FDR) <0.05] were included in the volcano plot, while low expression genes were excluded

© Journal of Gastrointestinal Oncology. All rights reserved.

(*Figure 2F*). A heatmap was painted and the DEGs in shGPSM2#2 and shNC groups were organized in obviously separate clusters (Figure S2). Comparing the mRNA expression profiles revealed that mesenchymal promoter *CD44* was downregulated at the transcriptional level in the GPSM2-silenced cells, while the epithelial marker *CDH1* (E-cadherin) was notably upregulated (Figure S2). These results suggested that *GPSM2* promoted epithelial-to-mesenchymal transition (EMT) in PAAD cells.

IPA of the genes

To gain further insight into the function of the identified DEGs for PAAD, enrichment analysis was performed using IPA. Several DEGs were significantly enriched in various IPA domains. In the canonical pathway, the genes were mainly enriched in neuregulin signaling, thrombin signaling, acute myeloid leukemia signaling, *IL-8* signaling, and glioblastoma multiforme signaling (*Figure 3A*). Analysis of diseases and functions showed a focus on cancer, organismal injury and abnormalities, cellular movement, and gastrointestinal disease (*Figure 3B*). A heat map was constructed to examine the relationship between the expression levels of the DEGs and the activation or inhibition of the disease and function. Significantly activated diseases or functions included viral infections (Z-score =3.521), liver lesions (2.514), and others as shown in *Figure 3C*. Marked inhibition of disease or function involved congenital anomaly of the cardiovascular system (–3.215), congenital heart disease (–3.194), and others as shown in *Figure 3C*.

Analysis of upstream regulators

Upstream regulatory factors were analyzed based on the IPA dataset. Upstream regulators that were significantly activated included microRNA *(miR)-124-3p*, *ESR1*, *PD98059*, and *MYC* (*Table 5*). Upstream regulators that were significantly suppressed included *TNF*, *PDGF BB*, and *ERK* (*Table 6*). The top ten upstream regulators, as well as their target genes, are shown in Figure S3.

The functional implications of these upstream regulators in the development of downstream diseases and functions were investigated and several regulator effects networks were constructed (Table S2). The network with the highest consistency score (*Figure 3D*) suggested that upstream genes such as *ALB*, *ATF4*, *C3*, *CD40LG*, *CSF1*, *CYR61*, *DICER1*, *EDN1*, *EGR1*, *ERK1/2*, *EZH2*, *F2*, *F2R*, *F3*, *FGF7*, *IL17A*, *IL36B*, and integrin beta 1 (*ITGB1*) may inhibit accumulation of neutrophils, cell movement of connective tissue cells, homing of tumor cell lines, increased levels of AST, migration of endothelial cell lines, and proliferation of hepatic stellate cells through the downstream genes *AHR*, *AKAP12*, *AKT1*, *ANGPTL4*, *ATF3*, *ATP2A2*, *AXL*, *CAPN2*, *CCL2*, *CD44*, *CD69*, *CXCL8*, *EDIL3*, and *EGFR*.

The PPI network and the validation of hub genes

A PPI network consisting of 1,501 nodes and 8,903 edges was established by retrieving 1,631 DEGs in the STRING database (version 10.5) and visualizing through the use of Cytoscape software (*Figure 4*). Based on the STRING database, the DEGs with the highest PPI scores identified by the number of degrees are shown in Table S3. Among these genes, AKT1 scored maximum node degrees. The top 30 genes, including *AKT1*, *EGFR*, *VEGFA*, *CDH1*, *HSP90AA1*, *CXCL8*, *KRAS*, *CD44*, *MMP2*, *MDM2*, and *ITGB1*, were selected as hub genes for further analysis. The prognostic information of the 30 hub genes was freely available at http://gepia.cancer-pku.cn/index.html. The findings demonstrated that *CD44* expression (*Figure 5A*, P=0.0051) was closely related to OS in PAAD, as was *EGFR* (*Figure 5B*, P=0.031), *ITGB1* (*Figure 5C*, P=0.031), *KRAS* (*Figure 5D*, P=0.014), *NRAS* (*Figure 5E*, P=0.034), *RAB5A* (*Figure 5F*, P=0.029), *STAT5B* (*Figure 5G*, P=0.0016), and *TFRC* (*Figure 5H*, P=0.036).

The expressions of these genes, as well as several members of the integrin family, were assessed (*Figure 5I*). In addition to decreased expression of *CD44* (P=0.008), levels of *AKT1* (P=0.023), *EGFR* (P=0.048), MAP2K1(P=0.033), *NRAS* (P=0.005), *ITGB1* (P=0.021), *ITGB5* (P=1.4E-04) , *ITGA2* (P=9.5E-03), *ITGA5* (P=5.9E-04) were all significantly lower in GPSM2-silenced PNAC-1 cells compared to control cells.

Discussion

PAAD is one of the most common fatal tumors (1). This study identified *GPSM2* as a new oncogene in PAAD, and revealed the correlation between *GPSM2* and T stage, TNM stage, and tumor grade, which are crucial prognostic factors in PAAD. Similarly, high *GPSM2* expression was also observed in several other gastrointestinal tumors, such as colon adenocarcinoma, rectum adenocarcinoma, and stomach adenocarcinoma (Figure S4). The expression of *GPSM2* was low in thyroid carcinoma, which may

© Journal of Gastrointestinal Oncology. All rights reserved.

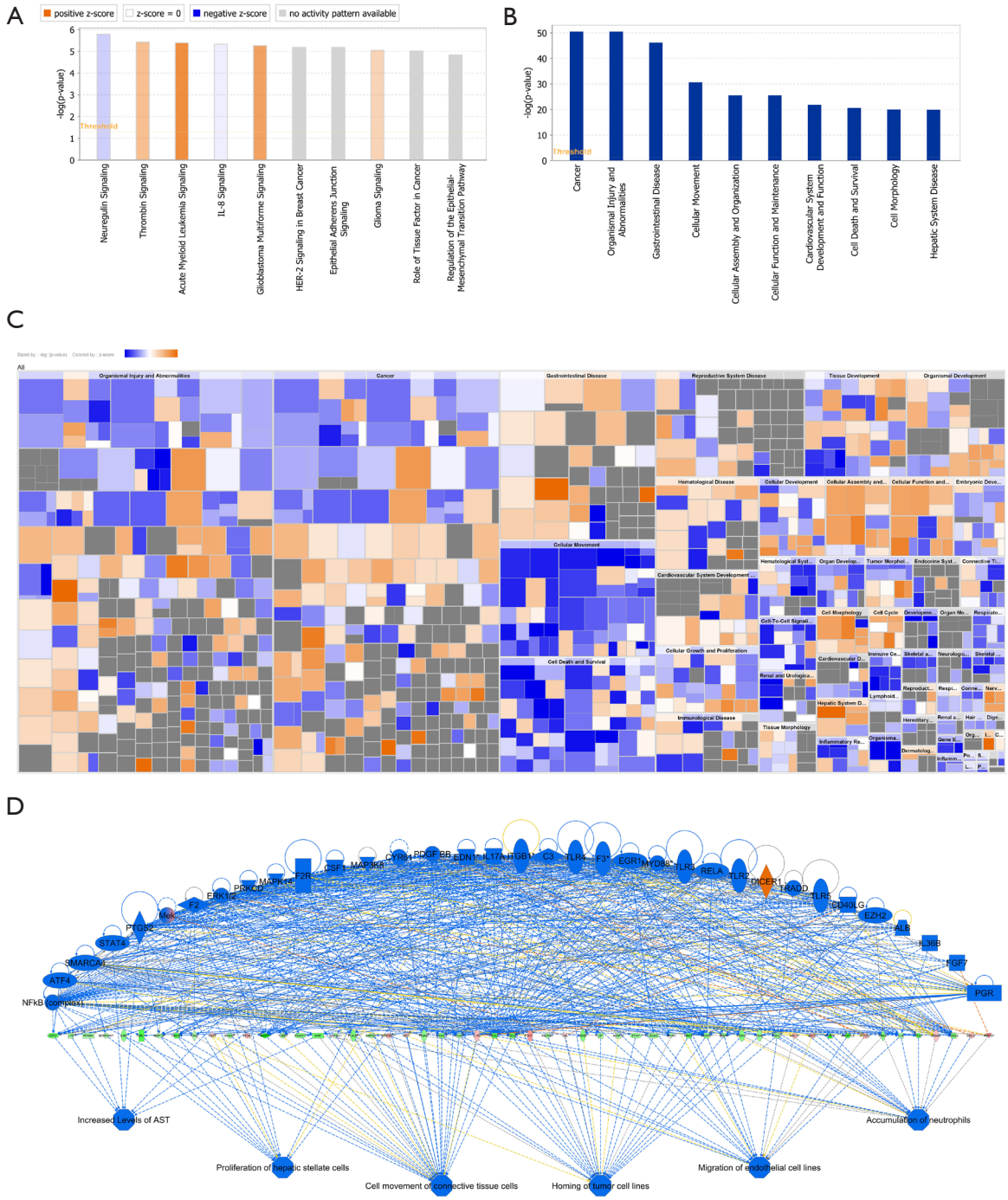

Figure 3 Ingenuity pathway analysis (IPA) of the differentially expressed genes. (A) Canonical Pathway analysis of differentially expressed genes in the microarray data. The length of the bar denotes –log(P value). (B) Enrichment analysis of differentially expressed genes in disease and function categories. X-axis: name of disease and function categories. Y-axis: –log(P value). (C) Activation or inhibition state of disease and function. Each block represents a specific function. Gray indicates that the disease or functional state is not determined. (D) Regulator effect network with the highest consistency score. Genes at the top are the upstream regulators; genes in the middle are the downstream genes; and the predicted functions are at the bottom. Orange, activated upstream regulators and functions (z-score >0); blue, inhibited upstream regulators and functions (z-score <0); red, activated downstream genes (Fold Change >1.5); green, inhibited downstream genes (Fold Change <–1.5).

© Journal of Gastrointestinal Oncology. All rights reserved.

Table 5 Top ten upstream regulators that were activated by differentially expressed genes

Activated upstream regulator	Description	z-score	P value
ESR1	estrogen receptor 1	4.332	2.81E-26
miR-124-3p	miRNA	5.112	2.21E-17
PD98059	extracellular signal-regulated kinase (ERK)1/2 inhibitor	3.5	3.78E-13
MYC	MYC Proto-Oncogene	2.681	1.05E-08
SFTPA1	Surfactant Protein A1	3.491	0.0000311
DICER1	Dicer 1, Ribonuclease II	2.505	0.000187
SP600125	c-Jun N-terminal kinases (JNK) inhibitor	3.193	0.00019
Sb202190	p38 MAPK inhibitor	2.868	0.0056
H-7	protein kinase C (PKC) inhibitor	2.619	0.0073
BDNF	Brain Derived Neurotrophic Factor	2.688	0.0134

z-score was used to predict the activation or inhibition of upstream regulators: z-score >0, activated; z-score <0, inhibited; P value <0.05.

Table 6 Top ten upstream regulators that were inhibited by differentially expressed genes

Inhibited upstream regulator	Description	z-score	P value
TNF	Tumor necrosis factor	–3.591	3.44E-25
PDGF BB	Platelet derived growth factor BB	–3.474	3.75E-14
ERK	Extracellular signal-regulated kinase	–3.309	1.15E-11
lipopolysaccharide	glycolipid	–5.272	7.9E-11
NUPR1	Nuclear Protein 1	–4.588	1.23E-10
Mek	Mitogen-activated protein kinase kinase	–3.272	2.97E-10
NFkB (complex)	Nuclear Factor kappa B	–3.441	4.35E-09
F2	Peptidase	–3.603	8.36E-09
IFNG	Interferon Gamma	–5.568	2.25E-08
IL1B	Interleukin 1 Beta	–3.596	3.31E-08

z-score was used to predict the activation or inhibition of upstream regulators: z-score >0, activated; z-score <0, inhibited; P value <0.0

partially explain the good prognosis in patients with thyroid carcinoma.

These findings improve our understanding of *GPSM2* biology and may help clinicians predict tumor progression and prognosis, especially in patients with a high level of *GPSM2* expression in the primary tumor. Furthermore, patients with tumors that exhibited a high level of *GPSM2* were associated with a poor prognosis, suggesting that *GPSM2* may act as a potential diagnostic and prognostic biomarker in PAAD.

With the increasing clinical success of cancer immunotherapies, there is a growing need to comprehensive understanding specific tumor immune interactions (26,27). This study demonstrated that *GPSM2* could promote the infiltration of immune cells in the tumor microenvironment and reduce the content of tumor cells. Dendritic cells have been reported to internalize antigens from the tumor and process them to present them to $CD8^+$ T cells (28). Also, macrophages play large roles in the promotion of pancreatic tumor growth and development of an immunosuppressive microenvironment (29). Our study found that GPSM2 expression is associated with the infiltration of macrophage cells, CD8+ T cells and dendritic cells, suggesting that the function of GPSM2 in pancreatic cancer may be related to

© Journal of Gastrointestinal Oncology. All rights reserved.

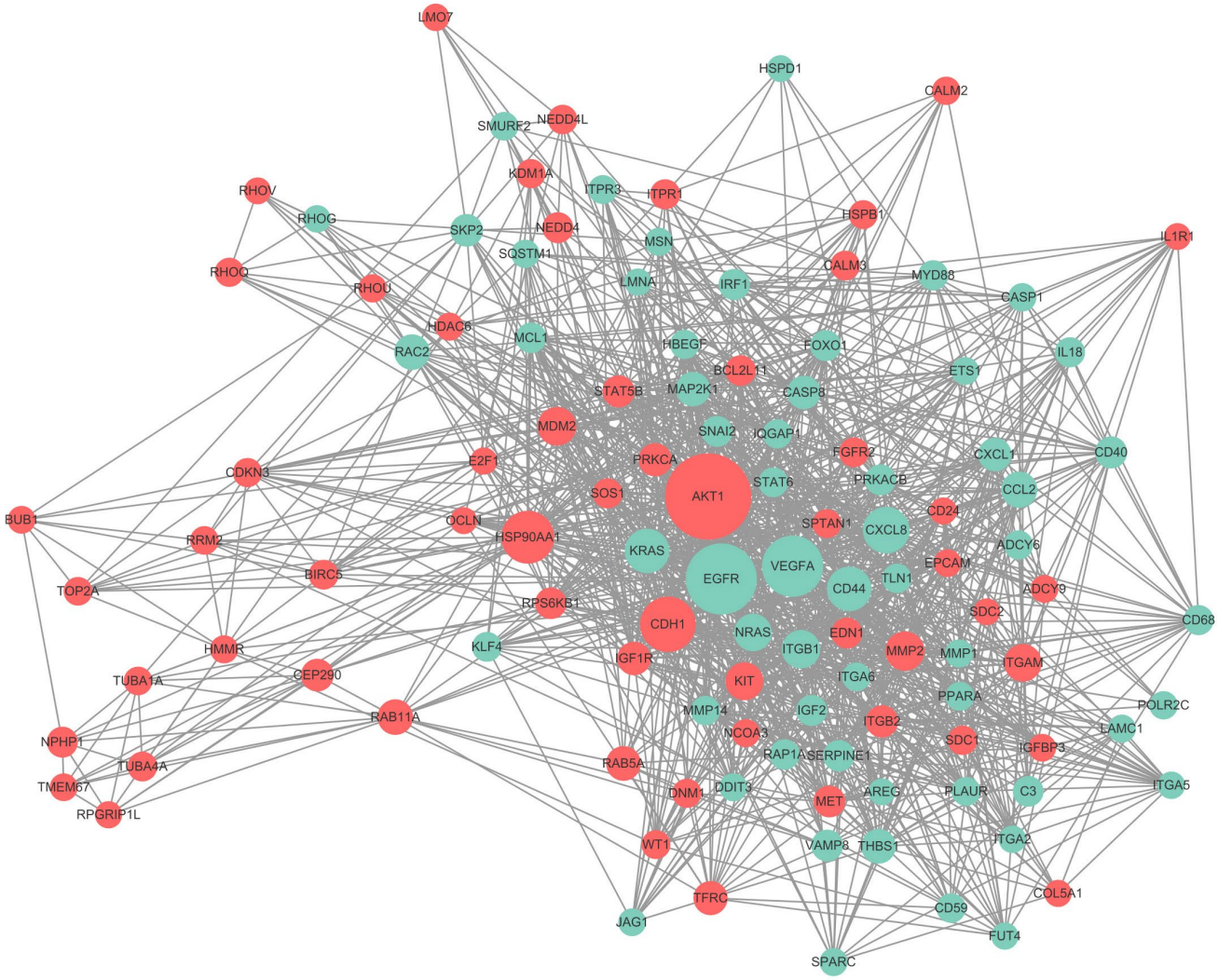

Figure 4 The protein-protein interaction (PPI) network for products of the 1,631 differentially expressed genes. The nodes with higher scores of degrees are shown in larger shapes and brighter colors. Red, up-regulated genes; green, down-regulated genes. Degree, the number of interactions between target gene and other differentially expressed genes. Genes with fewer degrees (<30) are not shown.

these components.

It has been reported that G-protein coupled receptors (*GPCRs*) play a critical role in the enhancement of breast cancer cell growth (7). However, it remained to be determined whether *GPSM2* is also involved in the carcinogenesis of the pancreas. The transwell assays in this investigation revealed that *GPSM2* facilitated cell migration in PANC-1 cells.

In this analysis, 1,631 DEGs were identified, consisting of 1,039 upregulated genes and 592 downregulated genes. The canonical pathways analysis showed that these DEGs were involved in epithelial adherens junction signaling and regulation of the EMT pathway. There is increasing evidence that EMT represents the key cellular reprogramming required for metastasis and chemoresistance in PAAD (30). Several growth factors have been documented to trigger EMT (31,32). A recent study reported that epithelial cadherin (*CDH1*) is involved in mechanisms regulating cell-cell adhesions, mobility, and proliferation of epithelial cells (33).

This investigation demonstrated that the expressions of the EMT-related proteins *CD44* and *SNAII2* were decreased, while E-cadherin was increased in the shRNA cells (*Figure 5I*), suggesting that *GPSM2* is a novel EMT

© Journal of Gastrointestinal Oncology. All rights reserved.

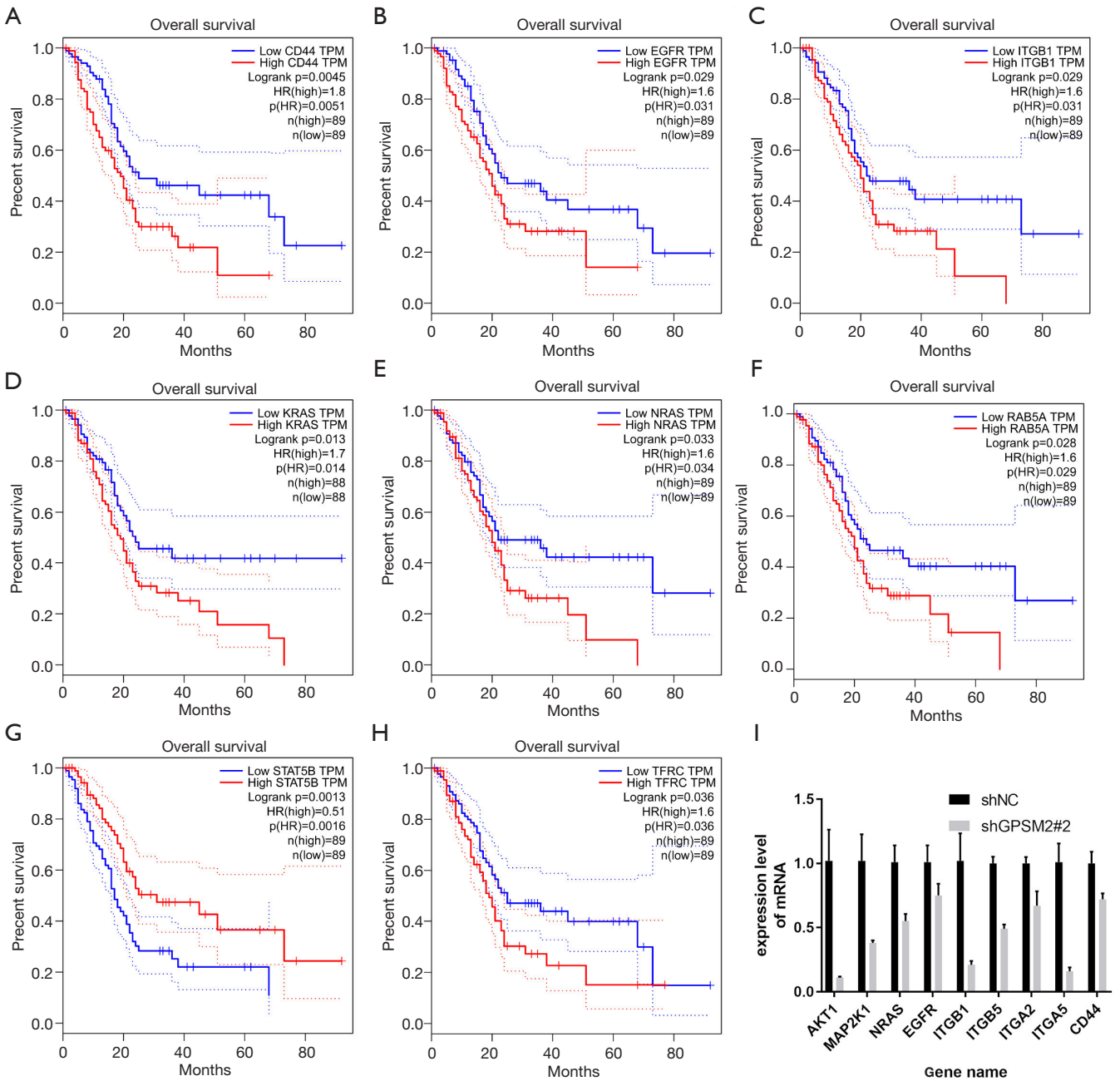

Figure 5 Detection of downstream hub genes. (A,B,C,D,E,F,G,H) Survival analysis of the hub genes in PAAD. Survival curves showing 8 genes (A–H represent CD44, EGFR, ITGB1, KRAS, NRAS, RAB5A, STAT5B, and TFRC, sequentially) that were related to overall patient survival rate (all P value <0.05). (I) mRNA expression in cells transfected with shGPSM2#2 and cells transfected with the scrambled sequence (shNC) (all P value <0.05).

inducer in PAAD.

Some reports have suggested that the *PI3K/AKT* pathway that transmits signals downstream of *GPCRs* is dysregulated in breast cancer (34). Another study suggested that *GPSM2* may be a critical molecule for carcinoma progression and may act as a potential biomarker and therapeutic target for the prevention and treatment of PAAD (6). The present study found that *GPSM2* could regulate *AKT1*, and this may

© Journal of Gastrointestinal Oncology. All rights reserved.

be the mechanism by which *GPSM2* promotes migration of pancreatic carcinoma cells. Integrins are cell-adhesion molecules, and reports have shown that its upregulation is associated with tumor progression, migration, invasion, and metastasis (7,35).

Studies have demonstrated that *ITGB1* can promote the growth and migration period of tumors in many types of human cancers, including pancreatic cancer (36-39), and thus *ITGB1* may be a therapeutic target for the treatment of these cancers. In fact, a monoclonal antibody against *ITGB1* has been shown to block hepatocellular cancer cell invasion (40) and knockdown of *ITGB1* by lentivirus-based RNAi inhibited PC cell growth and migration *in vitro* and *in vivo* (38). Indeed, the results of this current study suggested that *ITGB1* may be a downstream gene of *GPSM2* involved in oncogenesis.

In conclusion, this investigation identified *GPSM2* as a potential diagnostic and prognostic biomarker and its specific regulation mechanisms were preliminarily explored. Future studies should include gain-of-function experiments and *in vivo* experiments using xenografts in nude mice to further examine the effects of *GPSM2* in PAAD. In addition, the detailed mechanisms of *GPSM2* and its downstream pathways warrant further investigation. Finally, this study suggested that *GPSM2* may be a critical molecule for carcinoma progression and may be a potential biomarker and therapeutic target for the prevention and treatment of PAAD.

Acknowledgments

We thank LetPub (www.letpub.com) for linguistic assistance during the preparation of this manuscript.
Funding: This work was supported by the Key Research and Development Project of Zhenjiang (Social Development Project; SH2019061).

Footnote

Reporting Checklist: The authors have completed the REMARK reporting checklist. Available at https://dx.doi.org/10.21037/jgo-21-224

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/jgo-21-224). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin 2016;66:7-30.
2. Raimondi S, Maisonneuve P, Lowenfels AB. Epidemiology of pancreatic cancer: an overview. Nat Rev Gastroenterol Hepatol 2009;6:699-708.
3. Sener SF, Fremgen A, Menck HR, et al. Pancreatic cancer: a report of treatment and survival trends for 100,313 patients diagnosed from 1985-1995, using the National Cancer Database. J Am Coll Surg 1999;189:1-7.
4. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. CA Cancer J Clin 2013;63:11-30.
5. Feng RM, Zong YN, Cao SM, et al. Current cancer situation in China: good or bad news from the 2018 Global Cancer Statistics? Cancer Commun (Lond) 2019;39:22.
6. He XQ, Zhang YF, Yu JJ, et al. High expression of G-protein signaling modulator 2 in hepatocellular carcinoma facilitates tumor growth and metastasis by activating the PI3K/AKT signaling pathway. Tumour Biol 2017;39:1010428317695971.
7. Fukukawa C, Ueda K, Nishidate T, et al. Critical roles of LGN/GPSM2 phosphorylation by PBK/TOPK in cell division of breast cancer cells. Genes Chromosomes Cancer 2010;49:861-72.
8. Doherty D, Chudley AE, Coghlan G, et al. GPSM2 mutations cause the brain malformations and hearing loss in Chudley-McCullough syndrome. Am J Hum Genet 2012;90:1088-93.

© Journal of Gastrointestinal Oncology. All rights reserved.

9. Williams SE, Beronja S, Pasolli HA, et al. Asymmetric cell divisions promote Notch-dependent epidermal differentiation. Nature 2011;470:353-8.
10. Du Q, Macara IG. Mammalian pins is a conformational switch that links NuMA to heterotrimeric G proteins. Cell 2004;119:503-16.
11. Pishas KI, Adwal A, Neuhaus SJ, et al. XI-006 induces potent p53-independent apoptosis in Ewing sarcoma. Sci Rep 2015;5:11465.
12. Peng Y, Cui L, Shi J, et al. Expression of G-protein signaling modulator 2 in pancreatic cancer tissues. Jiangsu Daxue Xuebao Yixueban 2016;26:231-4.
13. Kulasingam V, Diamandis EP. Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. Nat Clin Pract Oncol 2008;5:588-99.
14. Tang Z, Li C, Kang B, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 2017;45:W98-W102.
15. Maag JLV. gganatogram: an R package for modular visualisation of anatograms and tissues based on ggplot2. F1000Res 2018;7:1576.
16. Goldman M, Craft B, Hastie M, et al. The UCSC Xena platform for public and private cancer genomics data visualization and interpretation. 2019 [cited 2019 July 24]. Available online: http://dx.doi.org/10.1101/326470
17. Yoshihara K, Shahmoradgoli M, Martínez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun 2013;4:2612.
18. Li B, Severson E, Pignon JC, et al. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. Genome Biol 2016;17:174.
19. Bolstad BM, Irizarry RA, Astrand M, et al. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 2003;19:185-93.
20. Irizarry RA, Bolstad BM, Collin F, et al. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 2003;31:e15.
21. Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003;4:249-64.
22. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.
23. Szklarczyk D, Franceschini A, Wyder S, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 2015;43:D447-52.
24. Scardoni G, Tosadori G, Faizan M, et al. Biological network analysis with CentiScaPe: centralities and experimental dataset integration. F1000Res 2014;3:139.
25. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498-504.
26. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer 2012;12:252-64.
27. Sharma P, Wagner K, Wolchok JD, et al. Novel cancer immunotherapy agents with survival benefit: recent successes and next steps. Nat Rev Cancer 2011;11:805-12.
28. Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. Immunity 2013;39:1-10.
29. Cui R, Yue W, Lattime EC, et al. Targeting tumor-associated macrophages to combat pancreatic cancer. Oncotarget 2016;7:50735-54.
30. Maier HJ, Wirth T, Beug H. Epithelial-mesenchymal transition in pancreatic carcinoma. Cancers (Basel) 2010;2:2058-83.
31. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. Nat Rev Mol Cell Biol 2014;15:178-96.
32. Biamonti G, Bonomi S, Gallo S, et al. Making alternative splicing decisions during epithelial-to-mesenchymal transition (EMT). Cell Mol Life Sci 2012;69:2515-26.
33. Meigs TE, Fedor-Chaiken M, Kaplan DD, et al. Gα12 and Gα13 negatively regulate the adhesive functions of cadherin. J Biol Chem 2002;277:24594-600.
34. Ye Y, Tang X, Sun Z, et al. Upregulated WDR26 serves as a scaffold to coordinate PI3K/AKT pathway-driven breast cancer cell growth, migration, and invasion. Oncotarget 2016;7:17854-69.
35. Gotzmann J, Mikula M, Eger A, et al. Molecular aspects of epithelial cell plasticity: implications for local tumor invasion and metastasis. Mutat Res 2004;566:9-20.
36. Caccavari F, Valdembri D, Sandri C, et al. Integrin signaling and lung cancer. Cell Adh Migr 2010;4:124-9.
37. Felding-Habermann B. Integrin adhesion receptors in tumor metastasis. Clin Exp Metastasis 2003;20:203-13.
38. Grzesiak JJ, Cao HST, Burton DW, et al. Knockdown of the β1 integrin subunit reduces primary tumor growth and inhibits pancreatic cancer metastasis. Int J Cancer 2011;129:2905-15.

© Journal of Gastrointestinal Oncology. All rights reserved.

39. Song J, Zhang J, Wang J, et al. β1 integrin mediates colorectal cancer cell proliferation and migration through regulation of the Hedgehog pathway. Tumour Biol 2015;36:2013-21.
40. Masumoto A, Arao S, Otsuki M. Role of β1 integrins in adhesion and invasion of hepatocellular carcinoma cells. Hepatology 1999;29:68-74.

(English Language Editor: J. Teoh)

Cite this article as: Zhou X, Dang S, Jiang H, Gu M. Identification of G-protein signaling modulator 2 as a diagnostic and prognostic biomarker of pancreatic adenocarcinoma: an exploration of its regulatory mechanisms. J Gastrointest Oncol 2021;12(3):1164-1179. doi: 10.21037/jgo-21-224

© Journal of Gastrointestinal Oncology. All rights reserved.

Supplementary

Table S1 The three different target sequences for GPSM2 (siGPSM2#1, siGPSM2#2, and siGPSM2#3)

No.	Target Seq
ShGPSM2#1	AGATACTATTGGAGATGAA
ShGPSM2#2	ACTTTACAATCTTGGGAAT
ShGPSM2#3	ATGATTATGCCAAAGCATT
Scrambled sequence	TTCTCCGAACGTGTCACGT

GPSM2, G-protein signaling modulator 2.

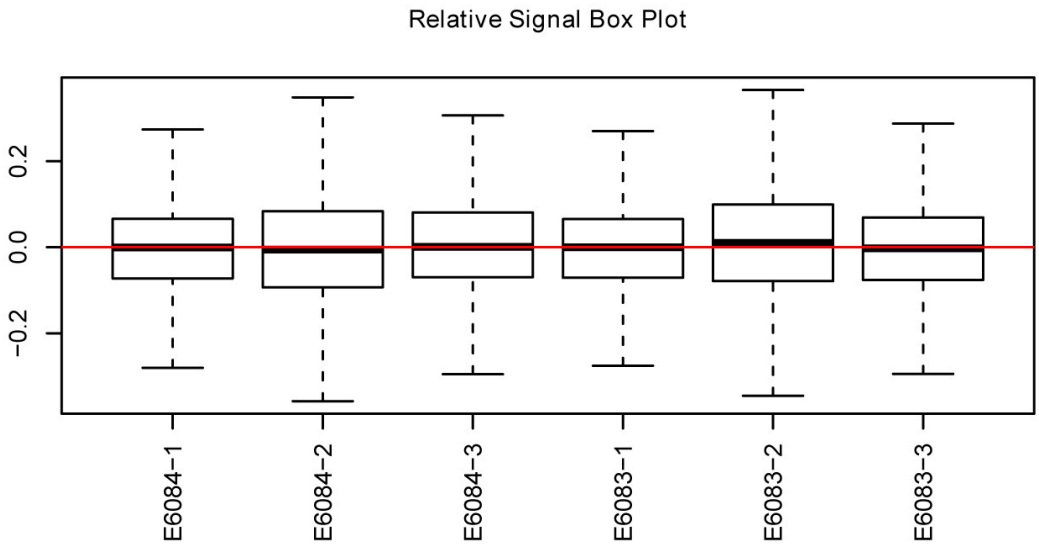

Figure S1 The relative signal boxplot of six microarrays after normalization.

© Journal of Gastrointestinal Oncology. All rights reserved.
https://dx.doi.org/10.21037/jgo-21-224

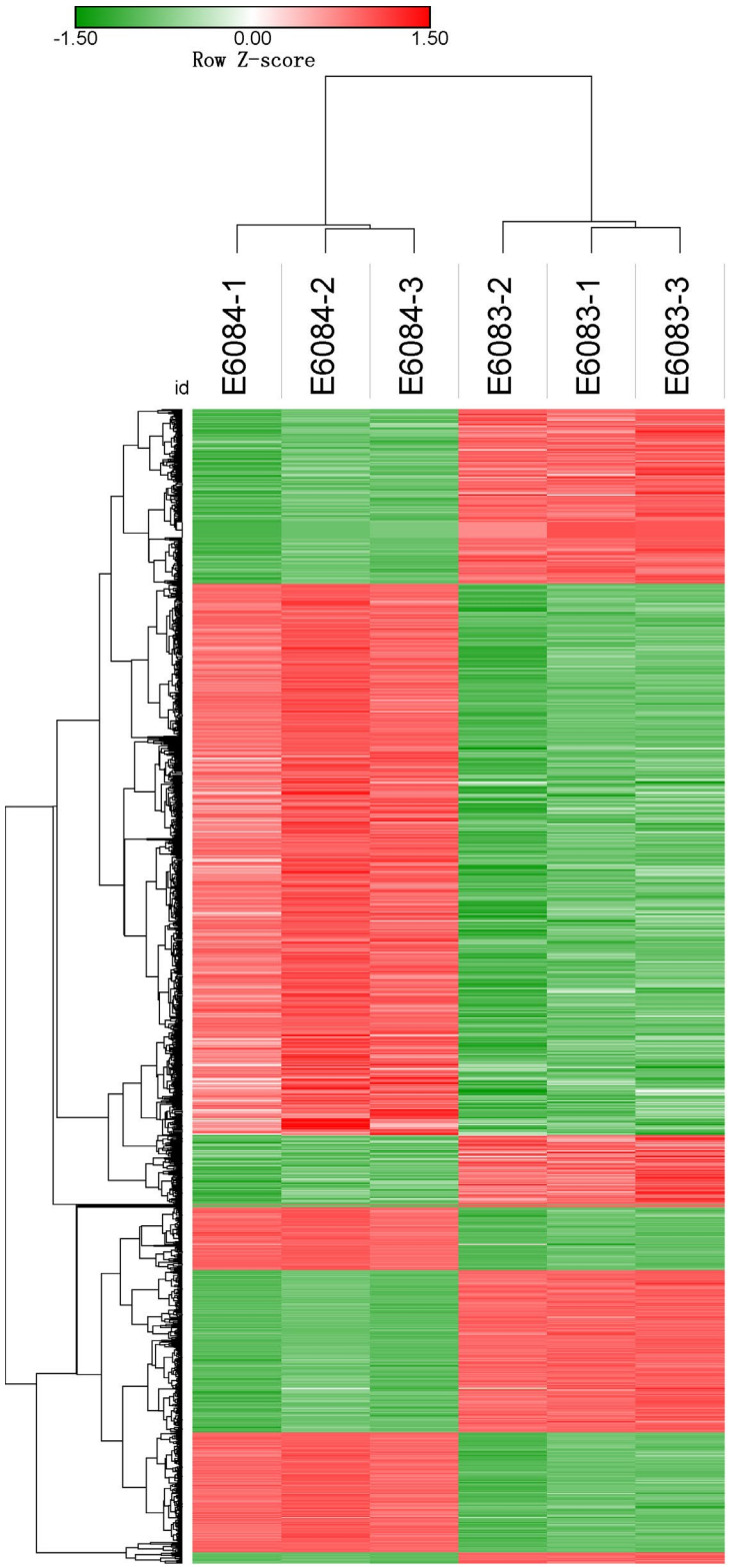

Figure S2 Heatmap of the differentially expressed genes (DEGs) in cells transfected with shGPSM2#2 (E6083-1, E6083-2 and E6083-3) and cells transfected with the scrambled sequence (shNC) (E6084-1, E6084-2 and E6084-3). Red patch, upregulation; green patch, downregulation. FDR <0.05 and |Fold Change| >1.5 were set as the cut-off criteria. FDR, false discovery rate.

© Journal of Gastrointestinal Oncology. All rights reserved.
https://dx.doi.org/10.21037/jgo-21-224

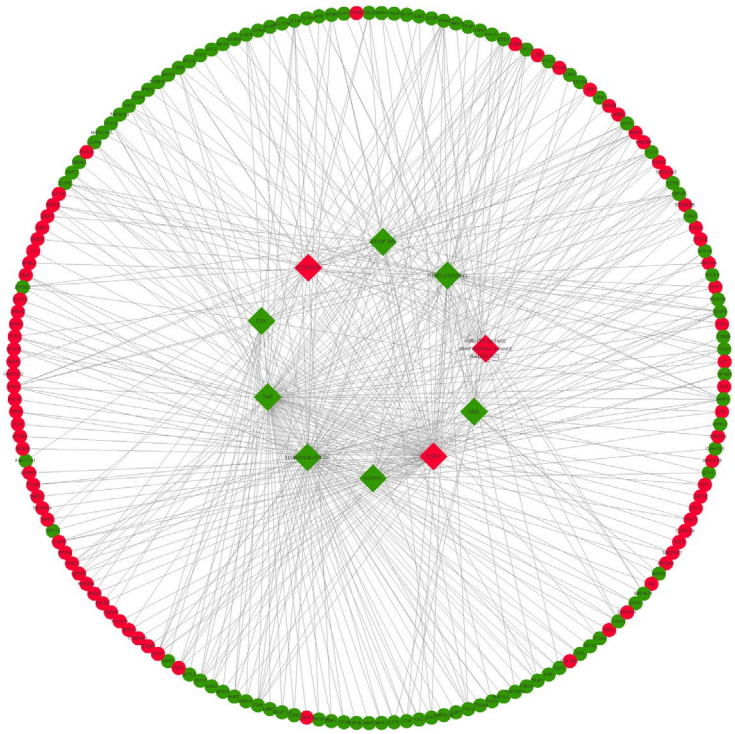

Figure S3 Network of the top ten upstream regulators and their downstream genes. Diamond, upstream regulators; circle, downstream genes. Red, upregulation; green, downregulation.

© Journal of Gastrointestinal Oncology. All rights reserved.
https://dx.doi.org/10.21037/jgo-21-224

Table S2 Top ten regulator effects analysis of upstream regulatory factors on downstream functions

ID	Consistency score	Node total	Regulator total	Regulators	Target total	Target molecules in dataset	Disease & function total	Diseases & functions	Known regulator-disease/function relationship	Counts of antibody list
1	31.843	95	35	ALB, ATF4, C3, CD40LG, CSF1, CYR61, DICER1, EDN1, EGR1, ERK1/2, EZH2, F2, F2R, F3, FGF7, IL17A, IL36B, ITGB1, MAP3K8, MAPK14, Mek, MYD88, NFkB (complex), PDGF BB, PGR, PRKCD, PTGS2, RELA, SMARCA4, STAT4, TLR2, TLR3, TLR4, TLR5, TRADD	54	AHR, AKAP12, AKT1, ANGPTL4, ATF3, ATP2A2, AXL, CAPN2, CCL2, CD44, CD69, CX3CL1, CXCL1, CXCL2, CXCL8, EDIL3, EGFR, ETV4, F2RL1, FGFR2, FOXO1, HBEGF, IL18, IL1R1, IL6R, ITGA2, ITGA5, ITGA6, ITGAM, KIT, KRAS, LIF, LTB, MAP2K1, MELTF, MET, MMP1, MMP14, NRG1, PLAUR, PLPP3, PPARA, PTX3, RAC2, RICTOR, SDC2, SERPINF1, SERPINH1, SNAI2, SPARC, SPHK1, STON1, THBS1, VEGFA	6	Accumulation of neutrophils,Cell movement of connective tissue cells,Homing of tumor cell lines,Increased Levels of AST,Migration of endothelial cell lines,Proliferation of hepatic stellate cells	20% (41/210)	14
2	24.4	37	7	ALB, C3, EDN1, F2, IL36B, ITGB1, TRADD	25	ANGPTL4, ATP2A2, CASP1, CCL2, CD44, CX3CL1, CXCL1, CXCL2, CXCL8, DDIT3, EGFR, F3, HBEGF, IL18, IRF1, ITGA2, ITGA5, ITGAM, MAP2K1, MMP1, NRG1, PLAUR, RAC2, THBS1, VEGFA	5	Accumulation of neutrophils,Apoptosis of kidney cells,Cell movement of connective tissue cells,Homing of tumor cell lines,Migration of endothelial cell lines	29% (10/35)	10
3	24.218	48	13	CREB1, ERK1/2, F2, FGF7, IL17A, IL36B, NFkB (complex), SMAD3, SPP1, TGFB3, TNFSF11, TP73, TRADD	32	AHR, AKT1, ANGPTL4, ATP2A2, C3, CAPN2, CCL2, CD44, CD69, CX3CL1, CXCL1, CXCL2, CXCL8, EDIL3, EDN1, EGFR, F3, FGFR2, IL18, ITGA2, ITGA5, ITGAM, ITGB1, KIT, LIF, MAP2K1, PLAUR, PTX3, SERPINF1, SNAI2, SPHK1, VEGFA	3	Accumulation of neutrophils,Homing of tumor cell lines,Migration of endothelial cell lines	10% (4/39)	11
4	21.372	53	10	ALB, BRAF, CCL5, ECSIT, EGR1, F2, GLIS2, IL36B, MAPK3, mir-515	37	AHR, ANGPTL4, ATF3, ATP2A2, C3, CAPN2, CCL2, CD44, CD69, CDH1, CX3CL1, CXCL1, CXCL2, CXCL8, DUSP6, EDN1, EGFR, ETS1, F3, HBEGF, IGF1R, IL18, ITGAM, ITGB1, LTBP2, MMP1, MMP14, NDRG1, PLAUR, RAC2, RFFL, SERPINE1, SNAI2, SPARC, TFPI, THBS1, VEGFA	6	Accumulation of neutrophils,Cell movement of connective tissue cells,Cell movement of melanoma cell lines,Homing of tumor cell lines,Migration of breast cancer cell lines,Migration of endothelial cell lines	13% (8/60)	12
5	19.941	80	9	CSF1, CYR61, EDN1, ERK1/2, F2, F2R, IL17A, PDGF BB, TLR4	66	AKAP12, ANGPTL4, ATP2A2, BCL10, BCL2L11, BIRC5, C3, CCL2, CD40, CD44, CD68, CDH1, CX3CL1, CXCL1, CXCL2, CXCL8, DCBLD2, DOCK10, DST, EDIL3, EGFR, ENPP1, ETS1, F3, FGFR2, FOXO1, FZD7, HBEGF, HMGA1, HSP90AA1, ID1, IDH2, IGF1R, IL18, IRF1, ITGA2, ITGA5, ITGA6, ITGAM, ITGB1, ITGB5, ITPR1, KCNH2, KRAS, LIF, LMNA, MAP2K1, MCL1, MET, MMP1, MMP14, MMP2, MYD88, NRG1, OCLN, PLAUR, PTX3, RAC2, RUNX1, SKP2, SNAI2, SPHK1, TFPI, THBS1, TUBB3, VEGFA	5	Accumulation of neutrophils,Cell movement of connective tissue cells,Homing of tumor cell lines,Leukemia,Migration of endothelial cell lines	49% (22/45)	24
6	19.936	46	10	ALB, BRAF, C3, EGR1, F2, GLIS2, IL17A, IL36B, ITGB1, TRADD	31	ANGPTL4, ATP2A2, CCL2, CD44, CDH1, CX3CL1, CXCL1, CXCL2, CXCL8, EDIL3, EDN1, EGFR, F3, HBEGF, IL18, ITGA5, ITGAM, ITPR1, LIF, LTBP2, MMP1, MMP14, MMP2, PLAUR, RAC2, RFFL, SERPINE1, SNAI2, SPARC, THBS1, VEGFA	5	Accumulation of neutrophils,Cell movement of connective tissue cells,Cell movement of melanoma cell lines,Homing of tumor cell lines,Migration of endothelial cell lines	24% (12/50)	9
7	19.894	58	21	CD40LG, CHUK, EDN1, F2, F3, IKBKG, IL18, IL1B, IL2, MAPK14, MYD88, NFkB (complex), PDGF BB, PTGS2, TLR2, TLR3, TLR4, TLR5, TLR7, TNFRSF1A, TYROBP	34	AHR, ANGPTL4, ATP2A2, CCL2, CD44, CX3CL1, CXCL1, CXCL2, CXCL8, EGFR, ENAH, FGFR2, FOXO1, HBEGF, IL1R1, ITGA2, ITGA5, ITGA6, ITGAM, ITGB1, KRAS, L1CAM, MAP2K1, MET, MMP1, MMP14, NRG1, PLAUR, PLPP3, PTX3, RAC2, SPARC, THBS1, VEGFA	3	Accumulation of neutrophils,Cell movement of connective tissue cells,Migration of endothelial cell lines	25% (16/63)	10
8	16.384	59	16	ALB, C3, CCL5, ECSIT, ETV5, F2, F7, IL17A, IL36B, ITGB1, MAPK3, miR-200b-3p (and other miRNAs w/seed AAUACUG), mir-515, NOD2, RHOA, TRADD	38	AHR, ANGPTL4, CAPN2, CCL2, CD44, CD69, CDH1, CX3CL1, CXCL1, CXCL2, CXCL8, DUSP6, EDIL3, EDN1, ETS1, F3, FHOD1, HBEGF, IGF1R, IL18, ITGA5, ITGAM, ITGB5, ITGBL1, LIF, MMP1, MMP14, MMP2, PLAUR, PPM1F, RAC2, SDC2, SERPINE1, SNAI2, TFPI, THBS1, TWIST2, VEGFA	5	Accumulation of neutrophils,Cell movement of connective tissue cells,Homing of tumor cell lines,Migration of breast cancer cell lines,Migration of endothelial cell lines	20% (16/80)	14
9	15.629	59	8	CYR61, EDN1, F2, F2R, F3, IL17A, TNFSF11, TRADD	46	ANGPTL4, ATP2A2, C3, CAPN2, CCL2, CD44, CEBPD, CLDN3, CX3CL1, CXCL1, CXCL2, CXCL8, DUSP6, EDIL3, EFEMP1, EGFR, EHD1, FOXO1, FZD7, HBEGF, HSP90AA1, IGF1R, IGFBP7, IL15RA, IL18, ITGA2, ITGA5, ITGA6, ITGAM, ITGB1, JAG1, KRAS, LIF, MAP2K1, MMP1, MMP2, NRG1, PLAUR, PLPP3, RAC2, SDC1, SERPINE1, STMN1, TFPI, THBS1, VEGFA	5	Accumulation of neutrophils,Cell movement of connective tissue cells,Homing of tumor cell lines,Migration of endothelial cell lines,Secondary neoplasm of digestive system	30% (12/40)	15
10	12.4	35	7	ERK1/2, F2R, IL17A, MYD88, TICAM1, TNFSF11, TRADD	25	C3, CAPN2, CCL2, CD44, CX3CL1, CXCL1, CXCL2, CXCL8, EDIL3, EDN1, F3, FGFR2, IGF1R, IL15RA, IL18, ITGA2, ITGA5, ITGA6, ITGAM, LIF, MAP2K1, PLAUR, SERPINE1, SNAI2, VEGFA	3	Accumulation of neutrophils,Homing of tumor cell lines,Liver metastasis	10% (2/21)	11

© Journal of Gastrointestinal Oncology. All rights reserved.
https://dx.doi.org/10.21037/jgo-21-224

Table S3 The top 15 differentially expressed genes with higher scores, identified by the number of degrees

Name	Description	Degree
AKT1	Protein kinase B	235
EGFR	Epidermal Growth Factor Receptor	183
VEGFA	Vascular Endothelial Growth Factor A	149
CDH1	Cadherin 1	129
HSP90AA1	Heat Shock Protein 90 Alpha Family Class A Member 1	120
CXCL8	C-X-C Motif Chemokine Ligand 8	98
KRAS	Kirsten rat sarcoma viral oncogene homolog	91
CD44	CD44 Molecule	91
MMP2	Matrix Metallopeptidase 2	73
MDM2	Mouse double minute 2	71
ITGB1	Integrin Subunit Beta 1	70
KIT	KIT proto-oncogene	68
ITGAM	integrin subunit alpha M	67
NRAS	Neuroblastoma RAS Viral Oncogene Homolog	67
CCL2	C-C Motif Chemokine Ligand 2	62

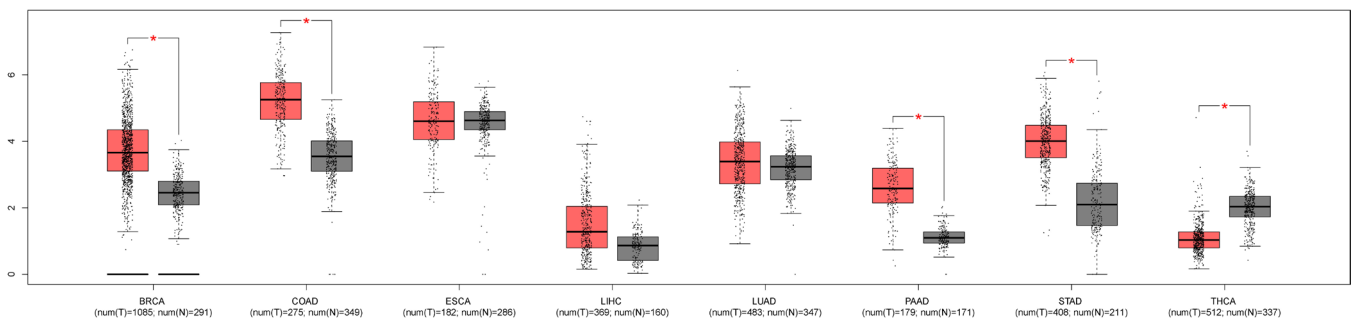

Figure S4 Expression levels of GPSM2 in primary cancer and adjacent normal tissues. *, P<0.05. BRCA, breast invasive carcinoma; COAD, colon adenocarcinoma; ESCA, esophageal carcinoma; LUAD, lung adenocarcinoma; PAAD, pancreatic adenocarcinoma; LIHC, liver hepatocellular carcinoma; STAD, stomach adenocarcinoma; THCA, thyroid carcinoma.

© Journal of Gastrointestinal Oncology. All rights reserved.
https://dx.doi.org/10.21037/jgo-21-224