Original Article

Expression of *NMU*, *PPBP* and *GNG4* in colon cancer and their influences on prognosis

Danyu Chen[1,2#], Zhen Ye[1,2#], Zhenxian Lew[3], Simin Luo[1,4], Zhong Yu[1,2], Ying Lin[1,2]

[1]Guangdong Provincial Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China; [2]Department of Gastroenterology and Hepatology, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China; [3]Department of Surgery, Guangzhou Concord Cancer Center, Guangzhou, China; [4]Breast Tumor Center, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China

Contributions: (I) Conception and design: D Chen, Y Lin; (II) Administrative support: Z Lew; (III) Provision of study materials or patients: D Chen, Z Ye; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Ying Lin, MD, PhD. Guangdong Provincial Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation Sun Yat-sen Memorial Hospital, Sun Yat-sen University, No. 107 West Yanjiang Road, Guangzhou 510120, China; Department of Gastroenterology and Hepatology, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, No. 107 West Yanjiang Road, Guangzhou 510120, China. Email: linwy@mail.sysu.edu.cn.

Background: This study aims to identify the core genes that influence the prognosis of colon cancer (CC) and analyze their relationships with clinical characteristics.

Methods: The gene expression profiles were downloaded from The Cancer Genome Atlas (TCGA) database. Differentially expressed genes (DEGs) were identified. The top ten core genes were selected by bioinformatics tools and screened through the Oncomine database. The expression of core genes in CC tissues and cells was validated by immunohistochemistry, immunoblotting and quantitative real-time polymerase chain reaction. Spearman correlation was used to analyze the relationship between different parameters. Overall survival was assessed by the Kaplan–Meier method. The area under the curve (AUC) and the receiver operating curve (ROC) were applied to assess the accuracy of genes for predicting prognosis.

Results: There were 1,665 DEGs that were identified from TCGA database. Bioinformatics analysis found that *GNGT1*, *NMU*, *PPBP*, *AGT*, and *GNG4* were differentially expressed in CC tissue. Overexpression of *NMU*, *PPBP*, *AGT*, and *GNG4* in CC was associated with shortened survival time ($P<0.05$). In the validation studies, the high expression levels of *NMU*, *PPBP* and *GNG4* in CC cells and tissues were confirmed compared to the control groups ($P<0.05$) and were adverse prognostic biomarkers ($P<0.01$). The combination prognostic model of the three core genes predicted the 1-, 3-, and 5-year survival of CC with AUCs of 0.868, 0.635 and 0.770, respectively.

Conclusions: High levels of *NMU*, *PPBP*, and *GNG4* were associated with poor prognosis in CC. The combination prognostic model of these three genes could be a new option.

Keywords: Colon cancer (CC); *NMU*; *PPBP*; *GNG4*; prognosis

Submitted May 15, 2022. Accepted for publication Aug 17, 2022.
doi: 10.21037/tcr-22-1377
View this article at: https://dx.doi.org/10.21037/tcr-22-1377

Introduction

In 2021, colon cancer (CC) was the third leading cause of death worldwide (1). The incidence and mortality rates of CC were increasing in the last decade. Although the clinical management of patients with CC has improved after a series of rigorous treatments, due to the lack of effective posterior-line therapy, the 5-year overall survival remains low, which has become a major challenge for global health (2). Therefore, identifying the high risk patients so as to develop precise therapeutic strategy are in urgent need. To accomplish this goal, screening for novel and promising prognostic biomarkers for CC remains the priority.

Bioinformatics is a combination of biological and informatics methods to identify key disease-causing factors to facilitate the exploration of new treatment methods and ultimately solve challenging medical problems, such as cancer. There are a growing number of studies using public databases, such as The Cancer Genome Atlas (TCGA) (3) and Oncomine (4). Statistical tools, such as the R data analysis package and Cytoscape visualization software, provide researchers with more intuitive methods for bioinformatics analysis (5,6). Gene Ontology (GO) (7) and Kyoto Encyclopedia of Genes and Genomes (KEGG, http://www.genetic.jp/) (8) are widely regarded as useful tools for genetic analysis. The Database for Integrated Discovery, Visualization and Annotation (DAVID, http://david.abcc.ncifcrf.gov/) (9) contains comprehensive biological knowledge and a series of analysis tools that can be used to extract genetic biological information. These databases and bioinformatics tools provide much information for tumor research and contribute to precision medicine.

This study was carried out to identify the differentially expressed genes (DEGs) of human CC and analyze their relationships with clinical characteristics. We also evaluated the impact of these genes on clinical prognosis. We present the following article in accordance with the REMARK reporting checklist (available at https://tcr.amegroups.com/article/view/10.21037/tcr-22-1377/rc).

Methods

DEGs screening, signaling pathway enrichment and functional enrichment analysis

The study design is demonstrated in the flow diagram (*Figure 1*). All clinical and mRNA data of CC and tumor-adjacent tissues were obtained from the TCGA database (https://portal.gdc.cancer.gov) for DEG screening. R software was used to convert gene expression into numerical values, average repeated genes, filter low-expressed genes, and draw heat maps and volcano maps. A false discovery rate (FDR) <0.05 and fold change ($|\log_2 FC|$) >2 were set as the criteria to screen statistically significant DEGs. The KEGG pathway enrichment and GO enrichment analyses of the hub genes were executed in the DAVID online tool (https://david.ncifcrf.gov). Statistical significance was defined with a P value beneath 5%.

Protein-protein interaction (PPI) network construction

DEGs were imported into the STRING (10) online tool (https://string-db.org/) to assemble a PPI network. The criteria for protein interaction screening were set (confidence >0.9). The network was reconstructed by the degree algorithm of cytoHubba in Cytoscape software v3.7.1 (11) (http://www.cytoscape.org). This module was applied to obtain the top ten hub genes.

Oncomine database screening

To identify the overlapping key genes across the Oncomine (www.oncomine.org) and TCGA databases, the hub genes obtained from the cytoHubba module were searched in the Oncomine database for cross selection. Therefore, the genes screened out both in TCGA and Oncomine were obtained.

Cell culture and quantitative reverse transcription polymerase chain reaction (qRT-PCR)

Human normal colonic epithelial cells NCM460 (RRID:CVCL_0460) and human colorectal cancer cells RKO (AddexBio Cat# C0009012/374, RRID:CVCL_0504), SW480 (Abcam Cat# ab271146, RRID:CVCL_0546), SW620 (AddexBio Cat# C0009002/68, RRID:CVCL_0547), HCT116 (Abcam Cat# ab255451, RRID:CVCL_0291) and COLO678 (DSMZ Cat# ACC-194, RRID:CVCL_1129) were obtained from the Cell Bank of the National Collection of Authenticated Cell Cultures (Shanghai, China). All cell lines were cultured in high glucose DMEM containing 10% fetal bovine serum and 1% penicillin-streptomycin and maintained in a humidified atmosphere of 95% air and 5% CO_2 at 37 ℃.

An EZ-press RNA purification kit was used to extract cellular RNA. A reverse transcription kit was used to synthesize cDNA, and qRT-PCR assays were performed on

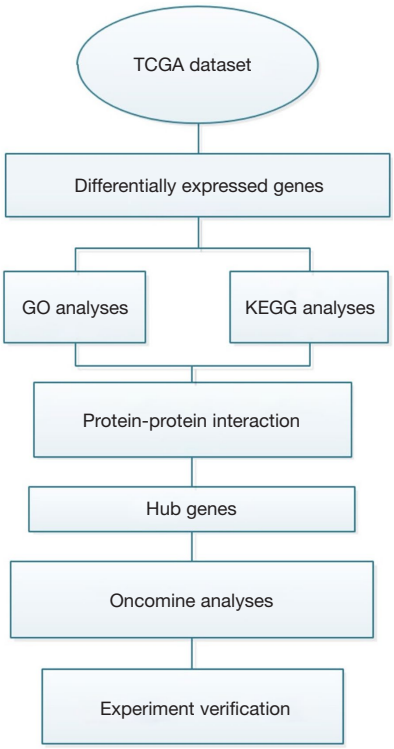

Figure 1 Flow diagram of the study design. TCGA, The Cancer Genome Atlas; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

a LightCycler® 96 thermal cycler. The primer sequences are shown in Table S1. The reagents were purchased from EZBioscience (USA). The work was carried out according to the instructions.

Immunohistochemistry

Tissue microarray sections containing 69 CC samples and 55 adjacent tissue samples were purchased from Weiao Biotechnology (Shanghai, China). The clinical characteristics of the patients are shown in Table S2. Xylene was used for deparaffinization, and graded ethanol dilutions were used to rehydrate the tissue microarray sections. Following incubation at 4 ℃ overnight with rabbit anti-NMU, anti-PPBP, and anti-GNG4 primary antibodies, the peroxide sections were then blocked with 3% hydrogen. After incubation with secondary antibody at room temperature, sections were stained with diaminobenzidine (DAB) and counterstained with hematoxylin. Based on the proportion of positive cells and the integrated staining intensity, the staining scores were evaluated by two pathologists independently. The final score ranged from 0 to 12. The samples with scores of 6–12 were defined as high expression, while the samples with scores of 0–5 were defined as low expression.

Immunoblotting

Total proteins were extracted from the mentioned cells above with RIPA buffer (Beyotime, China) that contained protease and phosphatase inhibitors (Beyotime). Following centrifugation of protein lysates, a bicinchoninic acid assay kit (Pierce, USA) was used to determine the concentration of the supernatants. The final protein lysates were boiled at 100 ℃ for 5 min with 5× loading buffer (Beyotime) and separated on a 15% sodium dodecyl sulfate (SDS) polyacrylamide gel (Fdbio Science, China). The gel was transferred onto a polyvinylidene fluoride (PVDF) membrane (Bio-Rad, USA), which was incubated with methanol for 1 minute. The membrane was blocked in Tris-buffered saline containing 0.1% Tween 20 (TBST), which was added to 5% skimmed milk for 2 h at room temperature and incubated with antibodies against NMU (1:1,000, DF4238, Affinity Biosciences, USA), PPBP (1:1,000, DF6695, Affinity Biosciences, USA), GNG4 (1:1,000, DF9560, Affinity Biosciences, USA), and GAPDH (1:1,000, cat. No. 5174, Cell Signaling Technology, USA) at 4 ℃ overnight. After washing with TBST, the PVDF membrane was incubated with HRP conjugated goat anti-rabbit IgGs (1:10,000, Cell Signaling Technology) for 1 h. Enhanced chemiluminescence (ECL) substrate (Bio-Rad) was used for detection, and GelView 6000Pro (BLT, China) was used for digitizing immunoblots. Band densities were semiquantified by ImageJ 1.52.

Statistical analysis

Spearman correlation was used to analyze the relationship between different parameters. Overall survival was assessed by the Kaplan-Meier method. The area under the curve (AUC) and receiver operating curve (ROC) were applied to assess the accuracy of genes for predicting prognosis. The levels of relative gene expression were calculated using the comparative threshold cycle ($2^{-\Delta\Delta Ct}$) method. Data were analyzed using SPSS 26.0, GraphPad Prism and R software. $P<0.05$ (bilateral) was considered statistically significant.

Ethical statement

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This article does not contain any studies with human participants or animals performed by any of the authors, and thus, there was no need for ethical approval.

Results

Identification of DEGs in CC

This study collected mRNA and clinical information of 480 CCs and 41 adjacent tissues from TCGA database. A total of 18,449 mRNAs were obtained, and 1,665 mRNAs were found to be differentially expressed, of which 911 were upregulated and 754 were downregulated (*Figure 2A*).

We then performed GO and KEGG enrichment analyses on 1,665 genes using the DAVID web tool. The GO analysis showed that DEGs were primarily involved in seven biological functions (*Figure 2B*), including nucleosome assembly, coagulation, digestion, cellular protein metabolism, cell adhesion, cell signal transduction and sodium ion transport. The KEGG analysis indicated that DEGs were involved in many biological pathways closely related to cancer progression, such as drug metabolism-cytochrome P450, neuroactive ligand-receptor interactions, alcoholism, chemical carcinogenesis, and metabolism of cytochrome P450 heterologous organisms (*Figure 2C*). These findings suggested that the DEGs participated extensively in the oncogenesis and progression of CC.

Identification of key genes for prognostic evaluation

We put the proteins encoded by DEGs into the STRING tool to construct a PPI network that was composed of 1,608 nodes and 3,084 edges (Figure S1A). After the PPI network was optimized using Cytoscape, the degree algorithm in cytoHubba was used to screen the top ten hub genes. The correlation between the hub genes (*Figure 2D*) and gene rankings (*Table 1*) was obtained. The top ten hub genes were then imported into the Oncomine database to search for key genes. *GNGT1*, *NMU*, *PPBP*, *AGT*, and *GNG4* were found to be differentially expressed in colorectal cancer in the Oncomine database, while the expression of *GNG13*, *LPAR1*, *NMUR2*, *CASR* and *PENK* was not significant (*Figure 2E*).

The clinical characteristics of the CC cohorts in the TCGA database were extracted, and survival analysis was performed with the five key genes (*Figure 2F*). Among these, the increased expression of *NMU*, *PPBP*, *GNG4*, and *AGT* indicated a poor prognosis (P<0.05), while the increased level of GNGT1 did not affect survival (P>0.05) (Figure S1B).

Expression of the key genes in CC cells and tissues

Quantitative RT-PCR was used to detect the mRNA expression levels of the key genes *GNG4*, *NMU*, *PPBP* and *AGT* in CC cells and normal colon epithelial cells. As shown in *Figure 3A*, the expression levels of *NMU*, *GNG4*, and *PPBP* mRNA in CC cells were significantly higher than those in normal colon epithelial cells (P<0.05), while the expression of AGT mRNA was similar among the different groups (P>0.05). Then, the protein expression levels of *GNG4*, *PPBP* and *NMU* in CC cells and normal epithelial cells were examined (*Figure 3B*). We found that both the mRNA and protein expression levels of *GNG4*, *PPBP* and *NMU* were increased in CC cells compared to normal epithelial cells.

To confirm the expression of the 3 key genes in human CC, we detected the expression of *GNG4*, *PPBP* and *NMU* in a tissue microarray including 69 CC and 55 adjacent normal tissues by immunohistochemistry. The results showed that the positive staining of NMU and GNG4 was mainly located in the cytoplasm, while PPBP was mainly located in the cytoplasm and interstitium of CC cells. The score of each protein was calculated, and a scatter plot involving each case was drawn (*Figure 4A-4C*). The expression scores of NMU, GNG4 and PPBP in CC and adjacent tissues were 8.246±3.863 *vs.* 1.761±2.013 (P<0.0001), 7.667±3.677 *vs.* 2.109±2.034 (P<0.0001), and 5.667±5.011 *vs.* 3.364±3.335 (P<0.0001), respectively. These results suggested that *NMU*, *GNG4* and *PPBP* were overexpressed in CC tissues compared with adjacent normal tissues.

Prognostic biomarkers of CC

We then explored the correlation between the expression of *NMU*, *GNG4*, and *PPBP* and the clinical characteristics of the CC patients in the tissue microarray cohort (*Table 2*). The results showed that positive expression of *PPBP* was associated with lymph node metastasis, distant metastasis and advanced tumor stages (P<0.05).

We used the Kaplan-Meier method to validate the relevance between the expression of NMU, GNG4, and

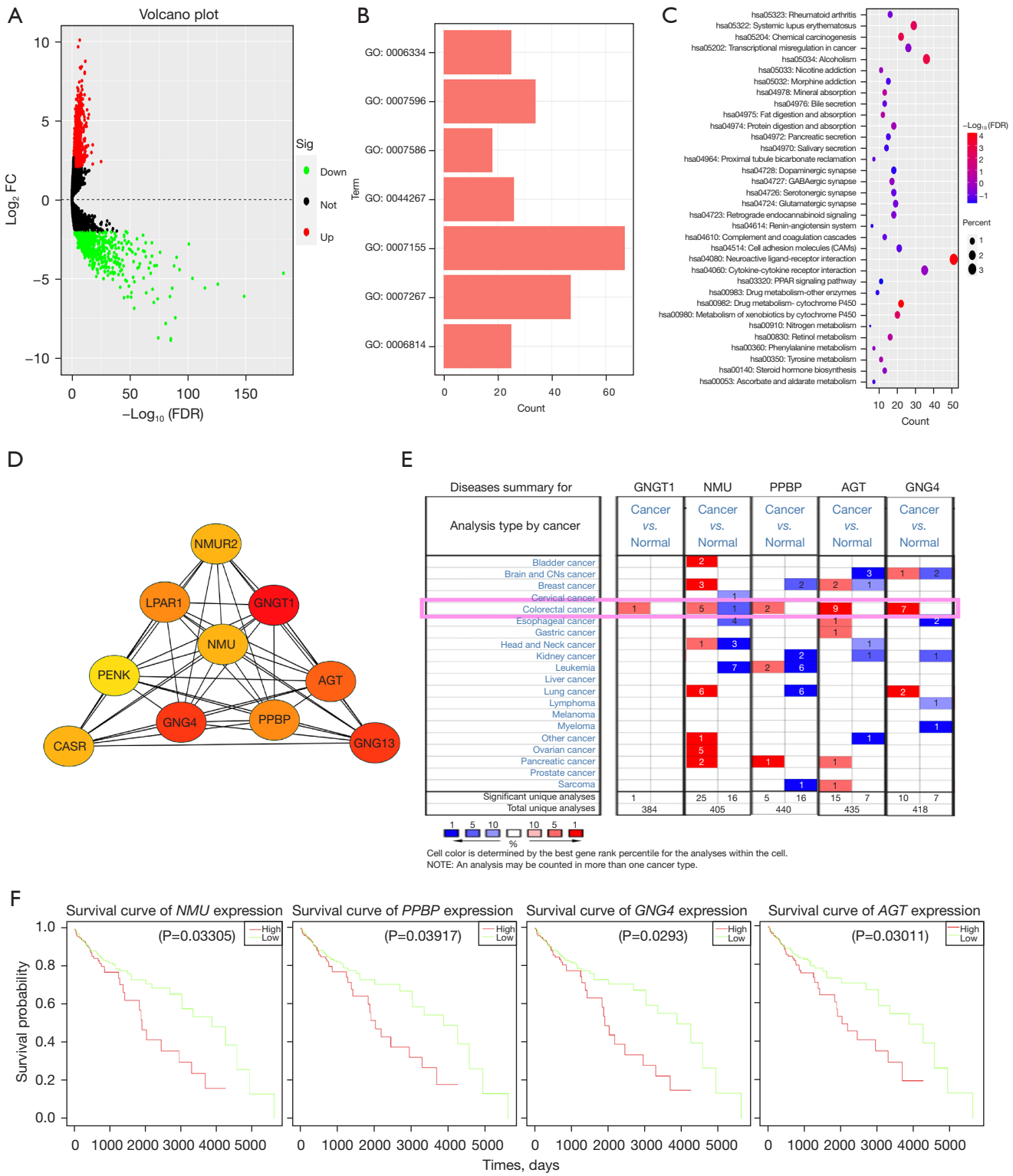

Figure 2 Identification of key genes in colon cancer. (A) Volcano map of DEGs. (B) GO functional enrichment. (C) KEGG pathway enrichment analysis. (D) Interrelationship of the top 10 genes (degree algorithm). (E) The expression of key genes in Oncomine. (F) Overall survival influenced by the key genes (*NMU*, *PPBP*, *GNG4*, and *AGT*). FC, fold change; FDR, false discovery rate; DEGs, differentially expressed genes; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

Table 1 The ranking of the top ten hub genes

Rank	Name	Score
1	*GNGT1*	106
2	*GNG13*	81
2	*GNG4*	81
4	*AGT*	68
5	*LPAR1*	63
5	*PPBP*	63
7	*NMU*	62
7	*NMUR2*	62
7	*CASR*	62
10	*PENK*	59

PPBP and the overall survival of 69 CC cases (*Figure 4D*). At the significance level of 5%, the results of univariate analysis showed that the survival time of patients with high NMU, GNG4, PPBP expression and distant metastasis was shorter than that of the patients with low expression, and without distant metastasis (*Table 3*, *NMU* HR =3.5, P=0.019; *GNG4* HR =3.1, P=0.034; *PPBP* HR =12.0, P<0.0001, distant metastasis HR =3.5, P=0.018). Moreover, the multivariate analysis showed that high *NMU*, *GNG4* expression and low grade of differentiation would shorten the survival time of CC patients (*Table 4*). The results indicated that *NMU*, *GNG4* and *PPBP* were adverse prognostic biomarkers of CC.

Subsequently, we drew the ROC to evaluate the sensitivity and specificity of *NMU*, *GNG4* and *PPBP* in predicting the survival of CC (*Figure 4E*). The combination of these three key genes predicted the overall survival of CC patients with AUCs of 0.868 (1 year), 0.635 (3 years) and 0.770 (5 years) respectively. These results suggested that the combination model surpassed the single biomarker in predicting the prognosis of CC.

Discussion

Although a substantial number of studies have declared that biomarkers are related to CC, only a few markers showed prognostic value (12-16). In this study, we explore the potential molecules that affect the prognosis of patients with CC through a combination of bioinformatics and experimental validation. A total of 1,665 CC-related DEGs were identified. Among these, *NMU*, *PPBP* and *GNG4* were found associated to the prognosis of CC. Moreover, the expression levels of *NMU*, *PPBP* and *GNG4* in CC cells and tissue microarray were validated. The expression levels of *NMU*, *PPBP* and *GNG4* were upregulated in CC cells and tissues, which were consistent with the findings in the TCGA database. These biomarkers may help improve risk stratification, treatment decisions, and prognosis prediction for patients with CC.

NMU is a member of the neuroprotein family that has a highly conserved sequence neuropeptide. NMU is mainly found in the pituitary gland and gastrointestinal tract (17). The "U" in NMU derives from its ability to stimulate strong contractions in the rat uterine smooth muscle (18). It has been revealed that NMU played an important role in regulating immunity, regulating feeding behavior, controlling circadian rhythms and balancing energy metabolism (19). Besides, previous studies have shown that as a neuropeptide, NMU is also related to poor survival in cancers, especially colorectal cancer (13,14,20-24).

PPBP was highly expressed in lymph node and peritoneal metastasis specimens of gastric cancer, which could be associated with the CXCR2 signaling pathway (25). Interference with the CXCR2/PPBP signaling pathway might be a new option for reversing the resistance of CC patients with liver metastases to conversion therapy (26). In addition, Kinouchi *et al.* found that the abnormal expression of PPBP in peripheral blood cells contributed to the diagnosis of renal cell carcinoma (RCC) (27).

Previous studies of GNG4 mainly focused on the nervous system, suggesting that GNG4 was related to cognitive decline and glioblastoma (28,29). However, recent studies found that GNG4 could be an adverse marker of rectal cancer and gallbladder cancer (30-32). Angiotensinogen (AGT) deficiency is related to inflammatory bowel disease and the development of cancer (33). Studies of AGT in CC have mainly focused on the field of liver metastasis (33-36). In the current study, however, we were not able to validate the differential expression of AGT in CC cells.

Although there were some bioinformatics analyses exploring the DEGs in CC using Gene Expression Omnibus (GEO) or TCGA or both databases (13-16,37) for diagnostic and/or prognostic purposes, overlapping

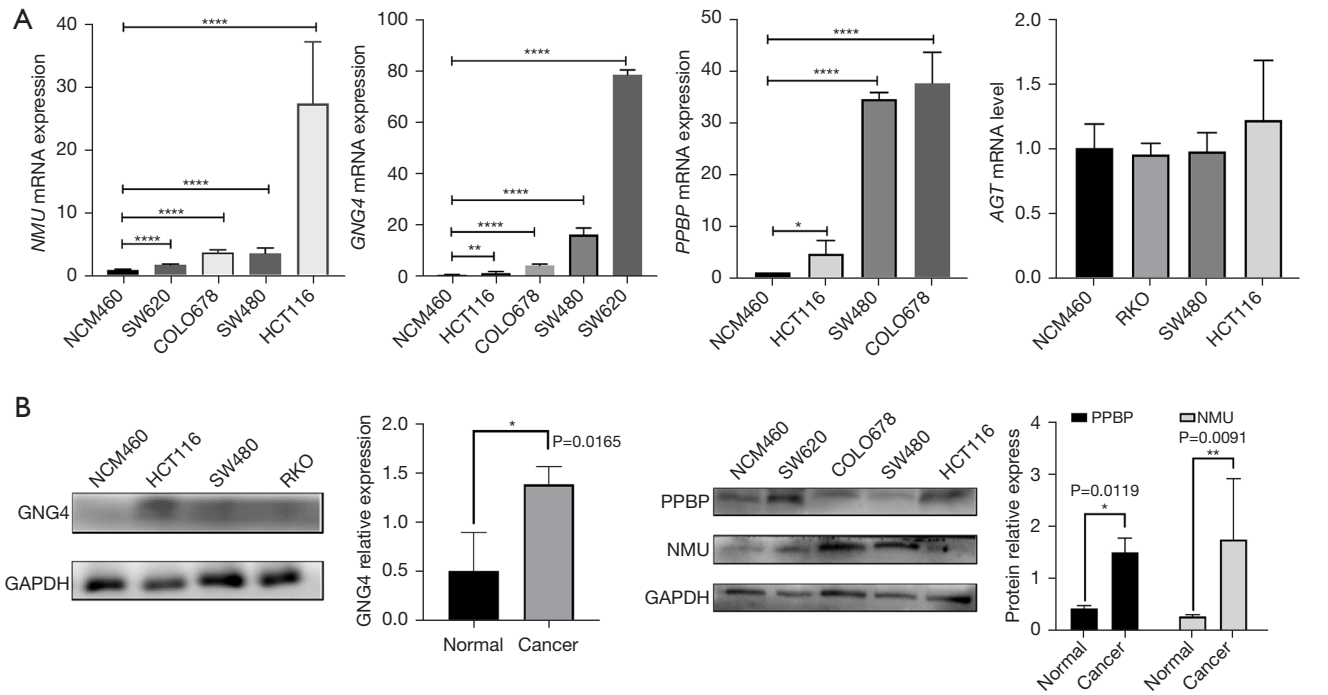

Figure 3 Expression of key genes in colon cancer cells. (A) *NMU*, *GNG4*, *PPBP* and *AGT* mRNA levels were detected by RT-qPCR. (B) The expression of NMU, GNG4 and PPBP was detected by immunoblotting. *, P<0.05; **, P<0.01; ****, P<0.0001. RT-qPCR, reverse transcription-quantitative polymerase chain reaction.

genes were identified consistent with our study, and seldom did they verify their findings with cells and tissue cohorts. Our study made up for this deficiency and focused on the prognostic value of these biomarkers. Nevertheless, the pathophysiological roles and the mechanisms of these 3 key genes participating in the oncogenesis and progression of CC remain to be clarified.

In conclusion, our study found that the combination model of NMU, PPBP and GNG4 expression predicted the overall survival of patients with CC at 1 and 5 years with high accuracy. They may serve as reliable prognostic biomarkers for CC.

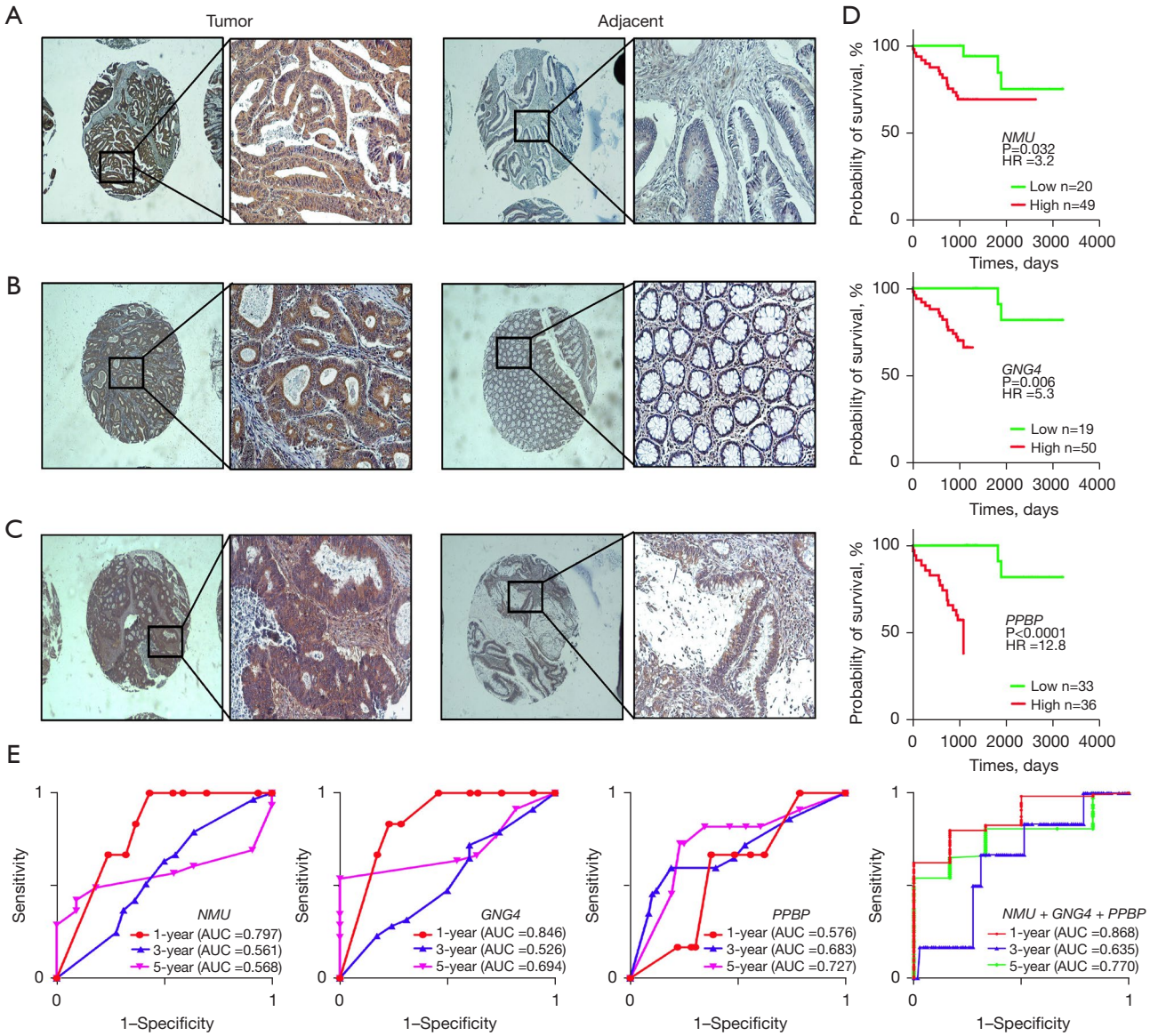

Figure 4 Expression of key genes in colon cancer tissue microarray and their influence on survival. (A) *NMU*, (B) *GNG4* and (C) *PPBP* were highly expressed in CC tissues. (A-C) IHC staining, 40× and 200×. (D) The survival curves of the key genes. (E) The ROC of the key genes and combination model for predicting 1-, 3-, and 5-year survival. HR, hazard ratio; AUC, the area under the ROC; CC, colon cancer; ROC, receiver operating curve; IHC, immunohistochemistry.

Table 2 The relationship between key gene expression and clinical characteristics

Clinical characteristics	No.	*NMU*				*GNG4*				*PPBP*			
		Low N=20	High N=49	r	P	Low N=19	High N=50	r	P	Low N=33	High N=36	r	P
Gender													
F	35	9	26	−0.073	0.550	10	25	0.024	0.848	18	17	0.073	0.550
M	34	11	23			9	25			15	19		
Age													
<65	24	6	18	−0.064	0.600	7	17	0.027	0.828	11	13	−0.029	0.812
≥65	45	14	31			12	33			22	23		
Pathological type													
Poor	66	18	48	−0.177	0.146	17	49	−0.187	0.124	31	35	−0.080	0.511
Well	3	2	1			2	1			2	1		
Node metastasis													
No	46	15	31	0.113	0.356	16	30	0.229	0.058	26	20	0.246	0.041*
Yes	23	5	18			3	20			7	16		
Distant metastasis													
No	61	20	41	0.231	0.056	18	43	0.122	0.318	32	29	0.256	0.034*
Yes	8	0	8			1	7			1	7		
TNM stage													
I/II	44	15	29	0.149	0.221	15	29	0.195	0.109	25	19	0.239	0.048*
III/IV	25	5	20			4	21			8	17		

*, statistically significant (P<0.05). TNM, tumor-node-metastasis.

Table 3 The univariate analysis of key gene expression and clinical characteristics

Variables	Univariate analysis		
	HR	95% CI	P value
GNG4 expression*	3.1	1.2–8.0	0.034*
PPBP expression*	12.0	4.7–30.8	<0.0001*
NMU expression*	3.5	1.4–9.0	0.019*
Sex	1.9	0.7–4.8	0.183
Age (≥65)	1.1	0.4–2.8	0.913
Grade	2.4	0.3–19.5	0.230
Node metastasis	1.9	0.7–5.1	0.181
Distant metastasis	3.5	0.6–19.8	0.018*
TNM stage	1.7	0.6–4.7	0.231

*, statistically significant (P<0.05). HR, hazard ratio; CI, confidence interval; TNM, tumor-node-metastasis.

Table 4 The multivariate analysis of key gene expression and clinical characteristics

Variables	Multivariate analysis								
	HR	95% CI	P value	HR	95% CI	P value	HR	95% CI	P value
GNG4 expression*	10.3	1.1–97.0	0.041*						
NMU expression*				13.3	1.5–120.5	0.021*			
PPBP expression*							–	–	0.570
Sex	2.5	0.9–7.4	0.090	3.0	1.0–8.7	0.048	2.4	0.8–7.1	0.110
Age (≥65)	1.4	0.5–3.8	0.543	1.5	0.5–4.0	0.463	1.4	0.5–3.9	0.483
Grade	14.9	1.8–120.7	0.011*	17.6	2.2–140.7	0.007*	19.6	2.2–168.1	0.007*
Node metastasis	–	–	0.925	–	–	0.925	–	–	0.931
Distant metastasis	3.1	0.8–12.8	0.115	2.6	0.6–10.7	0.186	1.9	0.4–7.8	0.390
TNM stage	–	–	0.925	–	–	0.926	–	–	0.931

HR, hazard ratio; CI, confidence interval; TNM, tumor-node-metastasis.

Acknowledgments

Funding: This work was supported by the Key Laboratory of Malignant Tumor Molecular Mechanism and Translational Medicine of Guangzhou Bureau of Science and Information Technology (No. 2013,163) and the Key Laboratory of Malignant Tumor Gene Regulation and Target Therapy of Guangdong Higher Education Institutes (No. KLB09001).

Footnote

Reporting Checklist: The authors have completed the REMARK reporting checklist. Available at https://tcr.amegroups.com/article/view/10.21037/tcr-22-1377/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tcr.amegroups.com/article/view/10.21037/tcr-22-1377/coif). All authors report that this work was supported by the Key Laboratory of Malignant Tumor Molecular Mechanism and Translational Medicine of Guangzhou Bureau of Science and Information Technology (No. 2013,163) and the Key Laboratory of Malignant Tumor Gene Regulation and Target Therapy of Guangdong Higher Education Institutes (No. KLB09001). The authors have no other 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). This article does not contain any studies with human participants or animals performed by any of the authors, and thus, there was no need for ethical approval.

References

1. Siegel RL, Miller KD, Fuchs HE, et al. Cancer Statistics, 2021. CA Cancer J Clin 2021;71:7-33.
2. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424.
3. Hutter C, Zenklusen JC. The Cancer Genome Atlas: Creating Lasting Value beyond Its Data. Cell

2018;173:283-5.
4. Rhodes DR, Yu J, Shanker K, et al. ONCOMINE: a cancer microarray database and integrated data-mining platform. Neoplasia 2004;6:1-6.
5. Ivanov AA, Revennaugh B, Rusnak L, et al. The OncoPPi Portal: an integrative resource to explore and prioritize protein-protein interactions for cancer target discovery. Bioinformatics 2018;34:1183-91.
6. Phuong T, Nhung N. Predicting gene function using similarity learning. BMC Genomics 2013;14 Suppl 4:S4.
7. Harris MA, Clark J, Ireland A, et al. The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res 2004;32:D258-61.
8. Arakawa K, Kono N, Yamada Y, et al. KEGG-based pathway visualization tool for complex omics data. In Silico Biol 2005;5:419-23.
9. Dennis G Jr, Sherman BT, Hosack DA, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 2003;4:P3.
10. Franceschini A, Szklarczyk D, Frankild S, et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res 2013;41:D808-15.
11. 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.
12. Kadoch C, Hargreaves DC, Hodges C, et al. Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. Nat Genet 2013;45:592-601.
13. Gong B, Kao Y, Zhang C, et al. Identification of Hub Genes Related to Carcinogenesis and Prognosis in Colorectal Cancer Based on Integrated Bioinformatics. Mediators Inflamm 2020;2020:5934821.
14. Chen L, Lu D, Sun K, et al. Identification of biomarkers associated with diagnosis and prognosis of colorectal cancer patients based on integrated bioinformatics analysis. Gene 2019;692:119-25.
15. Liang B, Li C, Zhao J. Identification of key pathways and genes in colorectal cancer using bioinformatics analysis. Med Oncol 2016;33:111.
16. Zhao ZW, Fan XX, Yang LL, et al. The identification of a common different gene expression signature in patients with colorectal cancer. Math Biosci Eng 2019;16:2942-58.
17. Domin J, Polak JM, Bloom SR. The distribution and biological effects of neuromedins B and U. Ann N Y Acad Sci 1988;547:391-403.
18. Raddatz R, Wilson AE, Artymyshyn R, et al. Identification and characterization of two neuromedin U receptors differentially expressed in peripheral tissues and the central nervous system. J Biol Chem 2000;275:32452-9.
19. Teranishi H, Hanada R. Neuromedin U, a Key Molecule in Metabolic Disorders. Int J Mol Sci 2021;22:4238.
20. Wang X, Chen X, Zhou H, et al. The Long Noncoding RNA, LINC01555, Promotes Invasion and Metastasis of Colorectal Cancer by Activating the Neuropeptide, Neuromedin U. Med Sci Monit 2019;25:4014-24.
21. Wang YP, Chen C, Li LX, et al. Neuromedin U expression related to the occurrence of laryngeal carcinoma and the regional lymph node metastasis. Lin Chung Er Bi Yan Hou Tou Jing Wai Ke Za Zhi 2016;30:811-4.
22. Li Q, Han L, Ruan S, et al. The prognostic value of neuromedin U in patients with hepatocellular carcinoma. BMC Cancer 2020;20:95.
23. Przygodzka P, Papiewska-Pajak I, Bogusz H, et al. Neuromedin U is upregulated by Snail at early stages of EMT in HT29 colon cancer cells. Biochim Biophys Acta 2016;1860:2445-53.
24. Zhang JJ, Hong J, Ma YS, et al. Identified GNGT1 and NMU as Combined Diagnosis Biomarker of Non-Small-Cell Lung Cancer Utilizing Bioinformatics and Logistic Regression. Dis Markers 2021;2021:6696198.
25. Yamamoto Y, Kuroda K, Sera T, et al. The Clinicopathological Significance of the CXCR2 Ligands, CXCL1, CXCL2, CXCL3, CXCL5, CXCL6, CXCL7, and CXCL8 in Gastric Cancer. Anticancer Res 2019;39:6645-52.
26. Desurmont T, Skrypek N, Duhamel A, et al. Overexpression of chemokine receptor CXCR2 and ligand CXCL7 in liver metastases from colon cancer is correlated to shorter disease-free and overall survival. Cancer Sci 2015;106:262-9.
27. Kinouchi T, Uemura M, Wang C, et al. Expression level of CXCL7 in peripheral blood cells is a potential biomarker for the diagnosis of renal cell carcinoma. Cancer Sci 2017;108:2495-502.
28. Bonham LW, Evans DS, Liu Y, et al. Neurotransmitter Pathway Genes in Cognitive Decline During Aging: Evidence for GNG4 and KCNQ2 Genes. Am J Alzheimers Dis Other Demen 2018;33:153-65.
29. Pal J, Patil V, Mondal B, et al. Epigenetically silenced GNG4 inhibits SDF1 /CXCR4 signaling in mesenchymal glioblastoma. Genes Cancer 2016;7:136-47.
30. Zhu D, Gu X, Lin Z, et al. High expression of PSMC2 promotes gallbladder cancer through regulation of GNG4

and predicts poor prognosis. Oncogenesis 2021;10:43.
31. Zhao H, Sheng D, Qian Z, et al. Identifying GNG4 might play an important role in colorectal cancer TMB. Cancer Biomark 2021;32:435-50.
32. Lv QY, Zou HZ, Xu YY, et al. Expression levels of chemokine (C-X-C motif) ligands CXCL1 and CXCL3 as prognostic biomarkers in rectal adenocarcinoma: evidence from Gene Expression Omnibus (GEO) analyses. Bioengineered 2021;12:3711-25.
33. Kuniyasu H. Multiple roles of angiotensin in colorectal cancer. World J Clin Oncol 2012;3:150-4.
34. Zhou L, Luo Y, Sato S, et al. Role of two types of angiotensin II receptors in colorectal carcinoma progression. Pathobiology 2014;81:169-75.
35. Shimomoto T, Ohmori H, Luo Y, et al. Diabetes-associated angiotensin activation enhances liver metastasis of colon cancer. Clin Exp Metastasis 2012;29:915-25.
36. Neo JH, Ager EI, Angus PW, et al. Changes in the renin angiotensin system during the development of colorectal cancer liver metastases. BMC Cancer 2010;10:134.
37. Sun G, Li Y, Peng Y, et al. Identification of differentially expressed genes and biological characteristics of colorectal cancer by integrated bioinformatics analysis. J Cell Physiol 2019. [Epub ahead of print]. doi: 10.1002/jcp.28163.

Cite this article as: Chen D, Ye Z, Lew Z, Luo S, Yu Z, Lin Y. Expression of *NMU*, *PPBP* and *GNG4* in colon cancer and their influences on prognosis. Transl Cancer Res 2022;11(10):3572-3583. doi: 10.21037/tcr-22-1377

Supplementary

Table S1 The primer sequences

Gene	Primer	Base sequence (5' to 3')
NMU	F	CCTCAGGCATCCAACGCACTG
NMU	R	CCTGCTGACCTTCTTCCATTCCG
PPBP	F	AGGTGCTGCTGCTTCTGTCATTG
PPBP	R	TGGCTATCACTTCGACTTGGTTGC
GNG4	F	GGCATGTCTAATAACAGCACCACTAG
GNG4	R	CAAAAGAACTTCTTCTCGCGAAAGG
GAPDH(HUMAN)	F	GGAGCGAGATCCCTCCAAAAT
GAPDH(HUMAN)	R	GGCTGTTGTCATACTTCTCATGG

Table S2 The clinical characteristics of the tissue microarray

Location	Gender	Age	Type	T	N	M	Survival time (m)	Survival time (d)	Death
A01	Female	89	Tumor	T3	N0	M0	45	1387	0
A02	Female	89	Adjacent	T3	N0	M0	45	1387	0
A03	Male	52	Tumor	T3	N0	M0	45	1371	0
A04	Male	52	Adjacent	T3	N0	M0	45	1371	0
A05	Female	74	Tumor	T2	N0	M0	44	1366	0
A06	Female	74	Adjacent	T2	N0	M0	44	1366	0
A07	Male	77	Tumor	T3	N0	M0	44	1363	0
A08	Male	77	Adjacent	T3	N0	M0	44	1363	0
A09	Female	60	Tumor	T3	N0	M0	44	1339	0
A10	Female	60	Adjacent	T3	N0	M0	44	1339	0
A11	Male	53	Tumor	T3	N1	M0	43	1329	0
A12	Male	53	Adjacent	T3	N1	M0	43	1329	0
A13	Female	64	Tumor	T3	N1	M0	31	960	1
A14	Female	64	Adjacent	T3	N1	M0	31	960	1
A15	Male	66	Tumor	T3	N1	M0	35	1083	1
A16	Male	66	Adjacent	T3	N1	M0	35	1083	1
B01	Female	77	Tumor	T3	N0	M0	43	1321	0
B02	Female	77	Adjacent	T3	N0	M0	43	1321	0
B03	Female	69	Tumor	T3	N0	M0	42	1290	0
B04	Female	69	Adjacent	T3	N0	M0	42	1290	0
B05	Male	80	Tumor	T3	N0	M0	42	1280	0
B06	Male	80	Adjacent	T3	N0	M0	42	1280	0
B07	Male	64	Tumor	T3	N2	M0	18	553	1
B08	Male	64	Adjacent	T3	N2	M0	18	553	1

Table S2 *(continued)*

https://dx.doi.org/10.21037/tcr-22-1377

Table S2 *(continued)*

Location	Gender	Age	Type	T	N	M	Survival time (m)	Survival time (d)	Death
B09	Female	78	Tumor	T3	N0	M0	20	635	1
B10	Female	78	Adjacent	T3	N0	M0	20	635	1
B11	Male	56	Tumor	T3	N0	M0	23	727	1
B12	Male	56	Adjacent	T3	N0	M0	23	727	1
B13	Female	70	Tumor	T3	N0	M0	40	1237	0
B14	Female	70	Adjacent	T3	N0	M0	40	1237	0
B15	Male	81	Tumor	T3	N1	M0	40	1219	0
B16	Male	81	Adjacent	T3	N1	M0	40	1219	0
C01	Female	78	Tumor	T3	N0	M0	40	1219	0
C02	Female	78	Adjacent	T3	N0	M0	40	1219	0
C03	Female	68	Tumor	T3	N0	M0	39	1209	0
C04	Female	68	Adjacent	T3	N0	M0	39	1209	0
C05	Female	72	Tumor	T3	N1b	M1b	30	929	1
C06	Female	72	Adjacent	T3	N1b	M1b	30	929	1
C07	Female	80	Tumor	T3	N0	M0	39	1196	0
C08	Female	80	Adjacent	T3	N0	M0	39	1196	0
C09	Male	68	Tumor	T3	N0	M0	38	1180	0
C10	Male	68	Adjacent	T3	N0	M0	38	1180	0
C11	Male	71	Tumor	T3	N1	M0	38	1169	0
C12	Male	71	Adjacent	T3	N1	M0	38	1169	0
C13	Female	81	Tumor	T3	N0	M0	28	859	1
C14	Female	81	Adjacent	T3	N0	M0	28	859	1
C15	Female	63	Tumor	T1	N0	M0	38	1162	0
C16	Female	63	Adjacent	T1	N0	M0	38	1162	0
D01	Male	65	Tumor	T1	N0	M0	37	1154	0
D02	Male	65	Adjacent	T1	N0	M0	37	1154	0
D03	Male	26	Tumor	T3	N1	M0	6	184	1
D04	Male	26	Adjacent	T3	N1	M0	6	184	1
D05	Male	63	Tumor	T3	N2a	M1a	24	752	1
D06	Male	63	Adjacent	T3	N2a	M1a	24	752	1
D07	Female	58	Tumor	T3	N1	M0	37	1142	0
D08	Female	58	Adjacent	T3	N1	M0	37	1142	0
D09	Male	79	Tumor	T3	N1	M0	37	1146	0
D10	Male	79	Adjacent	T3	N1	M0	37	1146	0
D11	Male	45	Tumor	T3	N1	M0	36	1120	0

Table S2 *(continued)*

https://dx.doi.org/10.21037/tcr-22-1377

Table S2 *(continued)*

Location	Gender	Age	Type	T	N	M	Survival time (m)	Survival time (d)	Death
D12	Male	45	Adjacent	T3	N1	M0	36	1120	0
D13	Female	70	Tumor	T3	N0	M0	11	361	1
D14	Female	70	Adjacent	T3	N0	M0	11	361	1
D15	Female	75	Tumor	T3	N0	M0	36	1107	0
D16	Female	75	Adjacent	T3	N0	M0	36	1107	0
E01	Male	69	Tumor	T3	N0	M0	36	1097	0
E02	Male	69	Adjacent	T3	N0	M0	36	1097	0
E03	Male	81	Tumor	T3	N0	M0	0	4	1
E04	Male	81	Adjacent	T3	N0	M0	0	4	1
E05	Female	41	Tumor	T3	N0	M0	35	1087	0
E06	Female	41	Adjacent	T3	N0	M0	35	1087	0
E07	Male	64	Tumor	T3	N0	M0	35	1085	0
E08	Male	64	Adjacent	T3	N0	M0	35	1085	0
E09	Female	76	Tumor	T3	N0	M0	35	1081	0
E10	Female	76	Adjacent	T3	N0	M0	35	1081	0
E11	Female	56	Tumor	T3	N0	M0	35	1080	0
E12	Female	56	Adjacent	T3	N0	M0	35	1080	0
E13	Male	72	Tumor	T3	N2b	M1b	1	31	1
E14	Male	72	Adjacent	T3	N2b	M1b	1	31	1
E15	Male	84	Tumor	T3	N0	M0	35	1071	0
E16	Male	84	Adjacent	T3	N0	M0	35	1071	0
F01	Male	67	Tumor	T1	N0	M1a	35	1065	0
F02	Male	67	Adjacent	T1	N0	M1a	35	1065	0
F03	Male	87	Tumor	T3	N0	M0	18	577	1
F04	Male	87	Adjacent	T3	N0	M0	18	577	1
F05	Male	75	Tumor	T3	N1	M0	34	1052	0
F06	Male	75	Adjacent	T3	N1	M0	34	1052	0
F07	Male	84	Tumor	T3	N0	M0	33	1029	0
F08	Male	84	Adjacent	T3	N0	M0	33	1029	0
F09	Female	79	Tumor	T3	N1	M0	34	1035	0
F10	Female	79	Adjacent	T3	N1	M0	34	1035	0
F11	Female	71	Tumor	T3	N1	M0	33	1030	0
F12	Female	71	Adjacent	T3	N1	M0	33	1030	0
F13	Male	61	Tumor	T3	N2a	M1a	33	1028	0
F14	Male	61	Adjacent	T3	N2a	M1a	33	1028	0

Table S2 *(continued)*

https://dx.doi.org/10.21037/tcr-22-1377

Table S2 *(continued)*

Location	Gender	Age	Type	T	N	M	Survival time (m)	Survival time (d)	Death
F15	Female	90	Tumor	T2	N0	M0	33	1017	0
F16	Female	90	Adjacent	T2	N0	M0	33	1017	0
G01	Male	36	Tumor	T3	N0	M0	33	1013	0
G02	Male	36	Adjacent	T3	N0	M0	33	1013	0
G03	Female	83	Tumor	T3	N1	M0	33	1003	0
G04	Female	83	Adjacent	T3	N1	M0	33	1003	0
G05	Male	80	Tumor	T3	N0	M0	32	1001	0
G06	Male	80	Adjacent	T3	N0	M0	32	1001	0
G07	Female	59	Tumor	T3	N1	M0	32	992	0
G08	Female	59	Adjacent	T3	N1	M0	32	992	0
G09	Female	61	Tumor	T3	N2	M0	33	1003	0
G10	Female	61	Adjacent	T3	N2	M0	33	1003	0
G11	Female	82	Tumor	T3	N2a	M1a	32	978	0
G12	Female	82	Adjacent	T3	N2a	M1a	32	978	0
G13	Female	69	Tumor	T3	N0	M0	32	975	0
G14	Female	69	Adjacent	T3	N0	M0	32	975	0
G15	Male	68	Tumor	T4	N1a	M1a	24	731	1
G16	Female	77	Tumor	T3	N0	M0	105	3219	0
H01	Male	83	Tumor	T3	N0	M0	105	3201	0
H02	Female	76	Tumor	T3	N0	M0	60	1826	1
H03	Female	59	Tumor	T3	N0	M0	96	2927	0
H04	Female	60	Tumor	T3	N0	M0	88	2682	0
H05	Male	59	Tumor	T3	N0	M1a	86	2638	0
H06	Male	49	Tumor	T3	N1	M0	84	2569	0
H07	Male	80	Tumor	T3	N0	M0	83	2532	0
H08	Female	82	Tumor	T3	N0	M0	2	68	1
H09	Female	58	Tumor	T3	N0	M0	81	2486	0
H10	Male	59	Tumor	T3	N0	M0	8	266	1
H11	Male	78	Tumor	T3	N0	M0	62	1897	1
H12	Female	83	Tumor	T3	N0	M0	77	2371	0

https://dx.doi.org/10.21037/tcr-22-1377

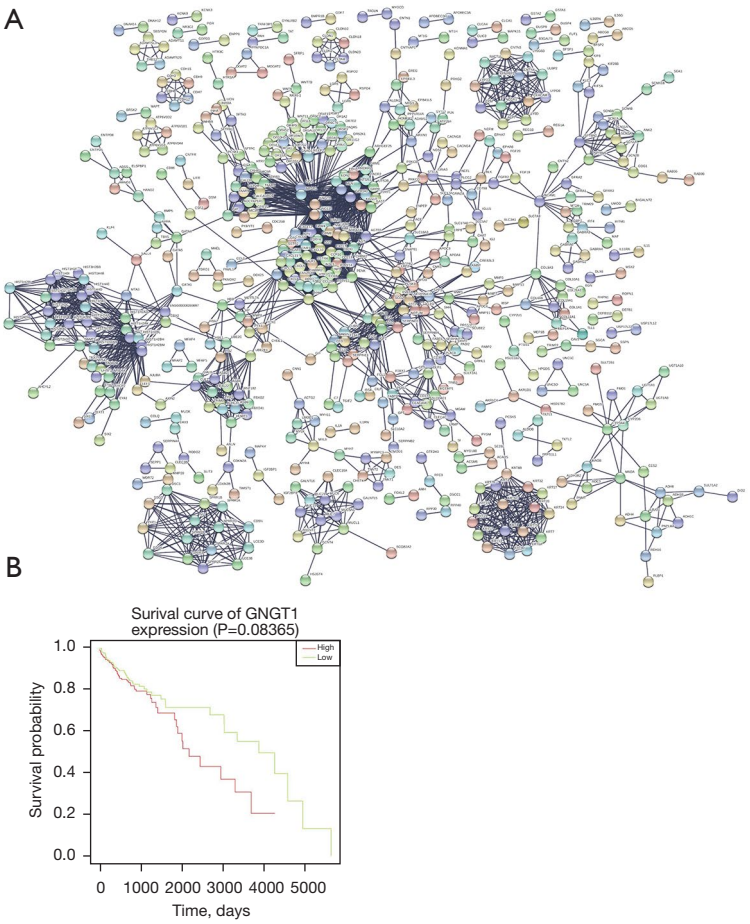

Figure S1 (A) Gene interaction relationship (PPI network). (B) Survival curve of GNGT1. PPI, protein-protein interaction.

https://dx.doi.org/10.21037/tcr-22-1377