Identification of Hub Genes and Biological Characteristics of Gastric Cancer by Multiple-Microarray Analysis

Xiangyuan Ding
Jingzhou No1 People's Hospital and First Affiliated Hospital of Yangtze University
Junjie Liu
Jingzhou No1 People's Hospital and First Affiliated Hospital of Yangtze University
Siqi Yan
Jingzhou No1 People's Hospital and First Affiliated Hospital of Yangtze University
Wei Yan (16634026543@163.com)
Jingzhou No1 People's Hospital and First Affiliated Hospital of Yangtze University

Research

Keywords: Gene Expression, Bioinformatics, Gastric cancer

Posted Date: May 4th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-415887/v1

Abstract

Background

Gastric cancer (GC) is not only listed as the third leading cause of cancer death, but is also ranked as the fifth most common cancer. The aims of the present study were to find the potential hub genes and uncover the possible molecular mechanisms of gastric cancer.

Methods

The gene expression profiles (GSE29272, GSE54129, GSE13911, GSE79973, GSE19826) were gained from the Gene Expression Omnibus (GEO) database. Then, we filtered out the differentially expressed genes (DEGs) online by GEO2R. According to Database for Annotation, Visualization and Integrated Discovery program (DAVID), functional and pathway enrichment analyses were fulfilled. Based on the Search Tool for the Retrieval of Interacting Genes/Proteins database (STRING) database and Cytoscape software, protein protein interaction (PPI) network were conducted and visualized. Following, the hub genes were selected by the CytoHubba plugin and analyzed using The Cancer Genome Atlas (TCGA) databases.

Results

A total of 105 co-regulated DEGs were extracted, containing 57 downregulated genes and 48 upregulated genes. The enrichment analysis indicated that they were remarkably enriched in various cancer-related functions and pathways. Eight genes (BGN, SPARC, COL5A2, COL5A1, COL1A2, COL4A1, COL6A3 and COL11A1) were recognized as hub genes in Cytoscape. These hub genes were the same as those in GC in the TCGA database. The survival analysis of COL5A2 and COL4A1 implied that worse survival status is caused by high expression levels in GC patients.

Conclusion

Our results showed that BGN, SPARC, COL5A2, COL5A1, COL1A2, COL4A1, COL6A3 and COL11A1 genes might be served as potential targets to improve diagnosis and immunotherapy biomarkers for GC.

Background

Of the worldwide population of patients with cancer, gastric cancer (GC) is not only listed as the third leading cause of cancer death, but is also ranked as the fifth most common cancer [1]. The most common tumor in the upper digestive tract is gastric cancer [1]. Although progress has been made in biopsy diagnosis and surgical treatment, most cases of gastric cancer are diagnosed as advanced and have a poor prognosis [2]. Moreover, the long-term survival rate of GC is still very dismal [3, 4]. Therefore, illumination of the molecular mechanisms is urgently needed to ameliorate the value of effective diagnosis and prognostic assessment for GC.

During the last decades, gene analysis using the high-throughput platforms has been extensively used to mine some potentially critical biomarkers for cancer diagnosis, prognostic assessment and new therapeutic target [5]. However, false-positive rates in a single expression profile analysis make it difficult to gain persuasive results.

Hence, in the present study, firstly, we downloaded and analyzed the gene expression profiles (GSE29272, GSE54129, GSE13911, GSE79973, GSE19826) between human GC and non-cancer tissues and obtained differentially expressed genes (DEGs) from the Gene Expression Omnibus (GEO) database. Secondly, functional and pathway enrichment analyses were further discussed. Screening potential hub genes through Cytoscape. Lastly, the expression of mRNA level of BGN, SPARC, COL5A2, COL5A1, COL1A2, COL4A1, COL6A3 and COL11A1 was obtained from the GEPIA databases, the prognostic roles and genetic alterations of hub genes were performed. The results of the present study may aid in providing useful insights into the novel prognostic biomarkers and the promising therapeutic targets for GC patients.

Results

Identification of differentially expressed genes

The gene expression profiles (GSE29272, GSE54129, GSE13911, GSE79973 and GSE19826) were identified in this analysis (Table 1). DEGs (2086 downregulated genes and 1858 upregulated genes in GSE54129, 2305 downregulated genes and 1001 upregulated genes in GSE13911, 174 downregulated genes and 176 upregulated genes in GSE29272, 754 downregulated genes and 408 upregulated genes in GSE19826, 919 downregulated genes and 487 upregulated genes in GSE79973) were screened, adjust P-value < 0.01 and |logFC|≥1 were considered as statistically significant. When the DEGs were obtained, the Draw Venn Diagram was used to draw a Venn diagram. 105 genes were preliminarily validated as co-differentially expressed genes, including 57 downregulated genes and 48 upregulated genes (Fig. 1).

Biological annotation terms

KEGG pathway analysis and GO enrichment analysis were conducted using DAVID. The outcomes of the GO analysis revealed that the upregulated genes were mainly involved in biological processes (BP), the major enriched terms were “extracellular matrix organization” and “cell adhesion”; in molecular function (MF), the major enriched terms were “calcium ion binding” and “extracellular matrix structural constituent”; in cell component(CC), the major enriched terms were “extracellular exosome”, “extracellular space” and “extracellular region”. However, the downregulated genes were particularly involved in biological processes (BP), the major enriched terms were “oxidation-reduction process” and “xenobiotic metabolic process”; in molecular function (MF), the major enriched terms were “oxidoreductase activity”; in cell component(CC), the main enriched terms were “extracellular exosome” and “organelle membrane”.

Moreover, KEGG analysis reflected that the upregulated DEGs were particularly involved in "focal adhesion", "PI3K-Akt signaling pathway" and "ECM-receptor interaction"; while the downregulated DEGs were particularly involved in "Drug metabolism - cytochrome P450" and "Gastric acid secretion" (Table 2).

Table 1
A summary of gastric cancer microarray datasets from different gene expression omnibus datasets

	Series	Platform	Affymetrix GeneChip	Normal Samples	Cancer Samples
1	GSE29272	GPL96	Affymetrix Human Genome U133A Array	134	134
2	GSE54129	GPL570	Affymetrix Human Genome U133 Plus 2.0 Array	21	111
3	GSE13911	GPL570	Affymetrix Human Genome U133 Plus 2.0 Array	31	38
4	GSE79973	GPL570	Affymetrix Human Genome U133 Plus 2.0 Array	10	10
5	GSE19826	GPL570	Affymetrix Human Genome U133 Plus 2.0 Array	15	12

Table 2
GO and KEGG analysis of differentially expressed genes in gastric cancer.

Category	Term	FDR	P-Value	Count	%
upregulated genes					
BP	GO:0030198 ~ extracellular matrix organization	3.56e-22	2.42e-25	20	42.55
BP	GO:0030574 ~ collagen catabolic process	1.37e-10	9.32e-14	10	21.28
BP	GO:0007155 ~ cell adhesion	1.70e-08	1.15e-11	15	31.91
BP	GO:0030199 ~ collagen fibril organization	1.36e-06	9.18e-10	7	14.90
BP	GO:0001501 ~ skeletal system development	1.25e-04	8.47e-08	8	17.02
BP	GO:0071230 ~ cellular response to amino acid stimulus	1.19e-02	8.06e-06	5	10.64
BP	GO:0016525 ~ negative regulation of angiogenesis	3.61e-02	2.45e-05	5	10.64
CC	GO:0031012 ~ extracellular matrix	1.72e-19	1.54e-22	20	42.55
CC	GO:0005576 ~ extracellular region	9.06e-17	8.13e-20	30	63.83
CC	GO:0005578 ~ proteinaceous extracellular matrix	1.22e-13	7.28e-17	16	34.04
CC	GO:0005788 ~ endoplasmic reticulum lumen	1.43e-09	1.29e-12	12	25.53
CC	GO:0005581 ~ collagen trimer	1.50e-09	1.35e-12	10	21.28
CC	GO:0005615 ~ extracellular space	2.00e-08	1.80e-11	21	44.68
CC	GO:0005604 ~ basement membrane	4.97e-05	4.46e-08	7	14.90
CC	GO:0070062 ~ extracellular exosome	1.27e-03	1.14e-06	22	46.81
MF	GO:0005201 ~ extracellular matrix structural constituent	1.30e-12	1.06e-15	11	23.40
MF	GO:0050840 ~ extracellular matrix binding	7.14e-06	6.01e-09	6	12.77

Category	Term	FDR	P-Value	Count	%
MF	GO:0005178 ~ integrin binding	1.09e-05	9.33e-09	8	17.02
MF	GO:0048407 ~ platelet-derived growth factor binding	1.53e-05	1.31e-08	5	10.64
MF	GO:0008201 ~ heparin binding	4.06e-03	3.46e-06	7	14.90
MF	GO:0005509 ~ calcium ion binding	1.41e-02	1.20e-05	11	23.40
MF	GO:0005518 ~ collagen binding	2.06e-02	1.76e-05	5	10.64
KEGG	hsa04512:ECM-receptor interaction	3.30e-14	3.44e-17	13	27.66
KEGG	hsa04510:Focal adhesion	3.89e-11	4.05e-14	14	29.79
KEGG	hsa04974:Protein digestion and absorption	1.00e-08	1.05e-11	10	21.28
KEGG	hsa04151:PI3K-Akt signaling pathway	5.80e-07	6.04e-10	13	27.66
KEGG	hsa05146:Amoebiasis	2.03e-06	2.12e-09	9	19.15
Hyper-methylated down-regulated genes					
BP	GO:0055114 ~ oxidation-reduction process	2.57e-04	1.87e-07	13	22.81
BP	GO:0006805 ~ xenobiotic metabolic process	7.74e-04	5.71e-03	6	10.53
BP	GO:0006081 ~ cellular aldehyde metabolic process	1.14e-03	6.23e-03	4	7.02
BP	GO:0008202 ~ steroid metabolic process	1.26e-03	1.27e-02	5	8.77
CC	GO:0070062 ~ extracellular exosome	8.84e-04	8.32e-07	25	43.86
CC	GO:0031090 ~ organelle membrane	7.40e-03	6.96e-06	6	10.53
MF	GO:0016491 ~ oxidoreductase activity	3.57e-02	2.96e-05	7	12.28
KEGG	hsa04971:Gastric acid secretion	7.10e-04	6.92e-07	7	12.28

Category	Term	FDR	P-Value	Count	%
KEGG	hsa00982:Drug metabolism - cytochrome P450	1.16e-02	1.13e-05	6	10.53

PPI network generation and hub genes selection

To gain the protein interactions information for each DEG, the STRING database was applied to generate a PPI network, which was visualized by Cytoscape. Also, 338 edges and 85 nodes were displayed in the Cytoscape (Fig. 2). In addition, eight hub genes in the four methods were chosen using CytoHubba plugin and sequentially ordered as follows: BGN, SPARC, COL5A2, COL5A1, COL1A2, COL4A1, COL6A3 and COL11A1.

Hub genes analysis and Clinic stage analyses

To determine the clinical significances of BGN, SPARC, COL5A2, COL5A1, COL1A2, COL4A1, COL6A3 and COL11A1 in patients with GC, we performed data mining and analyzed BGN, SPARC, COL5A2, COL5A1, COL1A2, COL4A1, COL6A3 and COL11A1 expressions from the publicly available GEPIA database. The expression of these eight hub genes was found higher in GC specimens than that in normal gastric specimens (Fig. 3A). Meanwhile, the expression of eight hub genes in GC patients were performed based on the TNM stage. The expression of seven genes (BGN, SPARC, COL5A2, COL5A1, COL1A2, COL6A3 and COL11A1) in patients with stage II, III and IV and was higher than that in patients with stage I, which suggests that these up-regulated hub genes might be positively correlated with tumor progression. The statuses of the expression of seven genes were the same as in our study, and their p-values < 0.05, except COL4A1 (p-value = 0.486) (Fig. 3B). Furthermore, the target analysis of TFs in TRRUST database was used for identification of TFs for hub genes. Three targeted transcription factors were identified. The outcomes obtained were displayed in Table 3. Among 3 TFs, two TFs, RELA and NFKB1, regulates the functions of BGN and COL1A2.

Table 3
Result of transcription factors prediction of hub genes obtained from TRRUST database.

	transcription factor	targets	Mode of Regulation	P value	FDR
1	CEBPZ	COL11A1	Repression	2.36e-06	7.07e-06
		COL1A2	Unknown	2.36e-06	7.07e-06
2	RELA	BGN	Activation	6.66e-03	6.75e-03
		COL1A2	Repression	6.66e-03	6.75e-03
3	NFKB1	BGN	Activation	6.75e-03	6.75e-03
		COL1A2	Repression	6.75e-03	6.75e-03

Mining genomic alterations of hub genes by cBioPortal

To accurately assess the genomic alterations of hub targets genes in gastric cancer, eight hub genes (BGN, SPARC, COL5A2, COL5A1, COL1A2, COL4A1, COL6A3 and COL11A1) was performed through TCGA Provisional. The outcomes displayed that 202 cases (55%) have alterations in eight genes (Fig. 4A); the frequency of alteration of eight hub genes in gastric cancer is displayed in Fig. 4B. For COL11A1 (16%), most alterations were missense mutation and mRNA upregulation, while a few were amplification, truncating mutation. For COL6A3 (16%), the majority of alterations were missense mutation and mRNA upregulation, with a small part of amplification and truncating mutation. For COL4A1 (15%), the majority of alterations were amplification and mRNA upregulation, with a small part of deep deletion and truncating mutation. For COL1A2 (18%), the majority of alterations were missense mutation, amplification and mRNA upregulation, with a small part of deep deletion. For COL5A1 (12%), the majority of alterations were missense mutation, mRNA upregulation, and truncating mutation, with a small part of deep deletion. For COL5A2 (10%), the majority of alterations were amplification, mRNA upregulation and missense mutation, with a small part of deep deletion and truncating mutation. Gene changes related to BGN (9%) were amplification, deep deletion, mRNA upregulation, and changes related to SPARC (5%) were missense mutation, mRNA upregulation and amplification.

The eight hub genes (BGN, SPARC, COL5A2, COL5A1, COL1A2, COL4A1, COL6A3 and COL11A1) had 28 gene pairs showing concurrent alterations, in which 21 gene pairs were significant ($P < 0.05$; data not shown).

Survival analysis in gastric cancer

To evaluate the prognostic value of these eight hub genes, the overall survival (OS) analyses of hub genes were displayed by GEPIA database. Survival curves according to GEPIA database showed that two hub genes (COL4A1 and COL5A2) were found to be associated with OS. The survival trend of BGN, SPARC, COL5A1, COL1A2, COL6A3 and COL11A1 was consistent with the prognosis, but there was no significant difference (Fig. 5). The analysis of these two genes implied that worse survival status is caused by high expression levels in GC patients. The investigation carried out by us has revealed that COL5A2 and COL4A1 were involved in the progression of GC and might be used as a prognostic marker for GC.

CMap analysis

To find small molecular compounds, CMap analysis was predicted. The results obtained were presented in Table 8. The most significant small molecular compounds were picotamide (enrichment score = -0.899) and TTNPB (enrichment score = -0.858). The prediction of small molecule drugs aims to develop potential drugs for GC and make existing drugs fully utilized. However, further researches were needed to verify the above results.

Table 4
The list of 7 most significant small molecular compounds provided by CMAP analysis

CAap name	Mean	Enrichment	P	Percent non-null
picotamide	-0.719	-0.899	0.00004	100
TTNPB	-0.441	-0.858	0.04032	100
spiradoline	-0.518	-0.832	0.00145	75
carmustine	-0.405	-0.825	0.01076	66
corticosterone	-0.458	-0.77	0.00571	75
emetine	-0.423	-0.701	0.01671	75
minocycline	-0.453	-0.615	0.02313	60

Discussion

Although progress has been made in biopsy diagnosis and surgical treatment, most cases of gastric cancer are diagnosed as advanced and have a poor prognosis [2]. Identification of favorable prognostic biomarkers related to GC is a pivotal step for seeking a contributing treatment.

To search for biomarkers for GC prognosis, we selected DEGs in GC tissues based on data available from public databases. To further understand these DEGs, these DEGs were analyzed for KEGG pathway and GO function. The outcomes of the GO biological processes displayed that the upregulated genes were particularly involved in extracellular matrix organization and cell adhesion. Meanwhile, KEGG pathway indicated that the upregulated DEGs were particularly involved in ECM-receptor interaction, focal adhesion and PI3K-Akt signaling pathway. Previous studies have reported that extracellular matrix organization and cell adhesion act a pivotal part in the development or progression of tumours [6–8]. Besides, dysregulation of the ECM-receptor interaction and focal adhesion and PI3K-Akt signaling pathway also act a pivotal part in cancer development and progression [9–11]. All in all, all theories are consistent with our findings.

The research we have done suggests that the expression of BGN, SPARC, COL5A2, COL5A1, COL1A2, COL4A1, COL6A3 and COL11A1 were drastically increased in gastric tumour tissues. Many members of the collagen family have been reported to be involved in tumorigenesis [12]. COL5A2, COL5A1, COL1A2, COL4A1, COL6A3 and COL11A1 belong to the collagen family. Collagen is the major structural protein of the extracellular matrix and overexpressed in many cancers [13]. Type V collagen (COL5) is a regulatory fibrocollagen composed of three different chains of COL5A1, COL5A2, and COL5A3 [14]. COL4A1 encodes type IV collagen alpha protein, which is an important component of the basal membrane found in most connective and embryonic tissues [15, 16]. Evidence indicates that COL4A1 plays a crucial role in tumor invasion through the induction of tumor budding [17]. COL11A1 has been paid more attention to in

the process of tumor transformation, containing metastasis, invasion and epithelial mesenchymal transformation. COL11A1 was remarkably overexpressed in colorectal cancer, ovarian cancer, and esophageal squamous cell carcinoma [18–20]. The expression of COL1A2 was positively correlated with the size and depth of invasion of gastric cancer [21]. At the same time, studies have shown that COL1A2 may play an important role in the invasion, metastasis and proliferation of pancreatic cancer [22]. COL6A3 is a ubiquitous extracellular matrix protein. Type VI collagen plays an important role in affecting metabolic levels, inhibiting apoptosis and oxidative damage, promoting cell growth. It has been reported that COL6A3 was remarkably overexpressed in colorectal cancer [23]. SPARC could bind to fibrous collagens I, III, V and IV [24]. BGN is a proteoglycan of the extracellular matrix. Evidence suggested that BGN is highly expressed in gastrointestinal tumors [25]. There are also reported that in some cancers, BGN up-regulation was associated with poor prognosis, advanced and metastatic cancer [26, 27].

Next, we used cBioportal to further explore the genomic alterations of hub genes in gastric cancer. In gastric cancer, most of the alterations of BGN, SPARC, COL5A2, COL5A1, COL1A2, COL4A1, COL6A3 and COL11A1 are amplification and mRNA upregulation, which may result in the upregulation of their expression correlated with the occurrence in gastric cancer.

In this study, we identified several small molecular drugs, which might improve GC. Picotamide is an antiplatelet drug that inhibits both TxA2 synthase and thromboxane A2 (TxA2) receptors. It may be potentially useful in various clinical settings characterized by atherosclerotic disease [28]. TTNPB is a retinoic acid receptor agonist, and is a very potent agent in carcinoprevention and carcinotherapy [29]. However, it has not been reported that small molecular compounds have the function of reversing gastric cancer. These small molecular compounds could be studied as new targets for the treatment of gastric cancer.

In this study, DEGs in GC were comprehensively analyzed by combining bioinformatics analysis. However, the present study presented limitations. First, in the present study, the experimental validation of these genes in GC is lacking. Second, a larger sample size is required to increase the credibility of the findings. Future studies should require many experiments and samples to verify our results.

Conclusion

In summary, this study was applied to identify 195 DEGs and select eight hub genes in GC, these genes may be related to the tumorigenesis or progression of GC. We hypothesized that our results can support potential targets and novel drugs for the treatment of GC. However, further large-scale experimental researches are needed to verify our findings.

Materials And Methods

Data source

The gene expression profiles (GSE29272, GSE54129, GSE13911, GSE79973 and GSE19826) were gained from the GEO (https://www.ncbi.nlm.nih.gov/geo/). The GSE54129 contained 132 samples, providing 111 GC sample tissues and 21 normal gastric sample tissues. The GSE29272 contained 268 samples, including 134 GC sample tissues and 134 normal gastric sample tissues. The GSE13911 provided 69 samples, including 38 GC sample tissues and 31 normal gastric sample tissues. The GSE79973 contained 20 samples, including 10 GC sample tissues and 10 normal gastric sample tissues. The GSE19826 included 27 samples, containing 12 GC sample tissues and 15 normal gastric sample tissues.

Data preprocessing and analyzing

We analysed the DEGs online using GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r). The threshold for the DEGs was defined as $|logFC| \geq 1$ and adjust P-value < 0.05. Venn diagram was performed to determine the intersection part using Draw Venn Diagram (https://bioinformatiocs.psb.ugent.be/webtools/Venn/).

Biological annotation terms

GO analysis was conducted to uncover the biological functions of the DEGs. Based on the KEGG database, biological pathway enrichment analysis was performed. In addition, Database for Annotation, Visualization and Integrated Discovery (DAVID) (https://david.ncifcrf.gov/) was applied to perform KEGG pathway enrichment analysis and GO analysis [30]. A false discovery rate (FDR) of < 0.05 was defined significant.

Protein-protein interaction network generation and hub genes selection

DEGs were imported into STRING online database (https://string-db.org/) [31]. An interaction with a combined score > 0.4 was chosen as the threshold.. PPI network of DEGs was analyzed via STRING and then the results were downloaded in table TSV format data. The results were visualized by Cytoscape v3.7.1 software, which was used to construct a protein-interaction relationship. Based on four ranked methods in cytoHubba, providing MNC, DMNC, MCC and EPC, hub genes were determined through overlapping the top 15 genes.

Hub genes analysis

To investigate the expression status of hub genes in GC, the Gene Expression Profiling interactive analysis (GEPIA, http://gepia.cancer-pku.cn) database was utilized for analysis [32]. The Transcriptional Regulatory Relationships Unraveled (TRRUST, http://www.grnpedia.org/trrust/) database is based on the existing literature to display the relationships between genes and TFs [33]. According to TRRUST, the TFs in hub genes were predicted. p < 0.005 were chosen as the threshold.

Genetic alterations of hub genes and survival analysis

The cBioPortal (www.cbioportal.org) is an integrated data mining database that provides visualization, analysis, and download of large-scale cancer genomics data sets [34]. The relevant information gained is related to the clinical outcomes, which helps to uncover the new connection. To further explore the

validity of this link, cBioportal was used to explore the genetic alteration of hub genes in gastric cancer. In order to assess the prognostic value of hub genes, the overall survival (OS) was used to analyze the hub genes by GEPIA (http://gepia.cancer-pku.cn) database.

CMap analysis

The Connectivity Map (CMap; https://portals.broadinstitute.org/cmap) is a comprehensive and publicly available resource that links drugs, genes and disease through opposite or similar genes. CMAP database was utilized to predict the small molecular compounds that may reverse altered expression of co-regulated DEGs. $P < 0.05$ and Mean < − 0.4 were considered as the threshold criteria. An up-regulated genes and a down-regulated genes were uploaded to CMap database, we gained small molecules with enrichment scores ranging from − 1 to + 1. Positive connectivity score (closer to + 1) suggested that a small molecule may induce GC gene expression, whereas negative connectivity score (closer to -1) suggested the similarity between small molecules and genes, which would reverse the status of GC. The outcomes were ranked by enrichment value.

Declarations

Acknowledgements

The authors gratefully acknowledge the Gene Expression Omnibus (GEO) database, which made the data available.

Authors' contributions

This research was conducted in collaboration with all authors. XYD and JJL performed the data curation and analysis. SQY and WY analyzed and interpreted the results. XYD, SQY and WY drafted and reviewed the manuscript. All authors read and approved the final manuscript.

Funding

NONE

Availability of data and materials

The raw data of this study are derived from the GEO data portal (https://www.ncbi.nlm.nih.gov/geo/), which are publicly available databases.

Ethics approval and consent to participate

Not necessary.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Contributor Information

Xiangyuan Ding, Email: 437754785@qq.com.

Wei Yan, Email: 16634026543@163.com.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.
2. Van Cutsem E, Sagaert X, Topal B, Haustermans K, Prenen H. Gastric cancer. The Lancet. 2016;388(10060):2654–64.
3. Orditura M, Galizia G, Sforza V, Gambardella V, Fabozzi A, Laterza MM, et al. Treatment of gastric cancer. World J Gastroenterol. 2014;20(7):1635–49.
4. Dicken BJ, Bigam DL, Cass C, Mackey JR, Joy AA, Hamilton SM. Gastric adenocarcinoma: review and considerations for future directions. Ann Surg. 2005;241(1):27–39.
5. Cancer Genome Atlas Research N. Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 2014;513(7517):202–9.
6. Jones CE, Hammer AM, Cho Y, Sizemore GM, Cukierman E, Yee LD, et al. Stromal PTEN Regulates Extracellular Matrix Organization in the Mammary Gland. Neoplasia. 2019;21(1):132–45.
7. Kobayashi H, Boelte KC, Lin PC.. Endothelial Cell Adhesion Molecules and Cancer Progression. Current Medicinal Chemistry. 2007; 14(4).
8. Makrilia N, Kollias A, Manolopoulos L, Syrigos K. Cell adhesion molecules: role and clinical significance in cancer. Cancer Invest. 2009;27(10):1023–37.
9. Buhrmann C, Shayan P, Goel A, Shakibaei M. Resveratrol Regulates Colorectal Cancer Cell Invasion by Modulation of Focal Adhesion Molecules. Nutrients. 2017;9(10).
10. Osaki M, Oshimura M, Ito H. PI3K-Akt pathway: Its functions and alterations in human cancer. Apoptosis. 2004;9(6):667–76.
11. Yeh MH, Tzeng YJ, Fu TY, You JJ, Chang HT, Ger LP, et al. Extracellular Matrix-receptor Interaction Signaling Genes Associated with Inferior Breast Cancer Survival. Anticancer Res. 2018;38(8):4593–605.
12. Ricard-Blum S. The collagen family. Cold Spring Harb Perspect Biol. 2011;3(1):a004978.
13. Fang M, Yuan J, Peng C, Li Y. Collagen as a double-edged sword in tumor progression. Tumour Biol. 2014;35(4):2871–82.

14. Mak KM, Png CY, Lee DJ. Type V Collagen in Health, Disease, and Fibrosis. Anat Rec (Hoboken). 2016;299(5):613–29.
15. Vitale G, Pichiecchio A, Ormitti F, Tonduti D, Asaro A, Farina L, et al. Cortical malformations and COL4A1 mutation: Three new cases. Eur J Paediatr Neurol. 2019;23(3):410–7.
16. Gould DB, Phalan FC, Breedveld GJ, van Mil SE, Smith RS, Schimenti JC, et al. Mutations in Col4a1 cause perinatal cerebral hemorrhage and porencephaly. Science. 2005;308(5725):1167–71.
17. Miyake M, Hori S, Morizawa Y, Tatsumi Y, Fujimoto K. Collagen type IV alpha 1 (COL4A1) and collagen type XIII alpha 1 (COL13A1) produced in cancer cells promote tumor budding at the invasion front in human urothelial carcinoma of the bladder. Oncotarget. 2017; 8(22).
18. Fischer H, Stenling R, Rubio C, Lindblom A. Colorectal carcinogenesis is associated with stromal expression of COL11A1 and COL5A2. Carcinogenesis. 2001;22(6):875.
19. Wu YH, Chang TH, Huang YF, Huang HD, Chou CY. COL11A1 promotes tumor progression and predicts poor clinical outcome in ovarian cancer. Oncogene. 2014;33(26):3432–40.
20. Zhang B, Zhang C, Yang X, Chen Y, Zhang H, Liu J, et al. Cytoplasmic collagen XIalphaI as a prognostic biomarker in esophageal squamous cell carcinoma. Cancer Biol Ther. 2018;19(5):364–72.
21. Li J, Ding Y, Li A. Identification of COL1A1 and COL1A2 as candidate prognostic factors in gastric cancer. World J Surg Oncol. 2016;14(1):297.
22. Wu J, Liu J, Wei X, Yu Q, Niu X, Tang S, et al. A feature-based analysis identifies COL1A2 as a regulator in pancreatic cancer. J Enzyme Inhib Med Chem. 2019;34(1):420–8.
23. Liu W, Li L, Ye H, Tao H, He H. Role of COL6A3 in colorectal cancer. Oncol Rep. 2018;39(6):2527–36.
24. Arnold SA, Rivera LB, Miller AF, Carbon JG, Dineen SP, Xie Y, et al. Lack of host SPARC enhances vascular function and tumor spread in an orthotopic murine model of pancreatic carcinoma. Disease Models Mechanisms. 2015;3(1–2):57–72.
25. Aprile G, Avellini C, Reni M, Mazzer M, Foltran L, Rossi D, et al. Biglycan expression and clinical outcome in patients with pancreatic adenocarcinoma. Tumour Biol. 2013;34(1):131–7.
26. Niedworok C, Rock K, Kretschmer I, Freudenberger T, Nagy N, Szarvas T, et al. Inhibitory role of the small leucine-rich proteoglycan biglycan in bladder cancer. PLoS One. 2013;8(11):e80084.
27. Maishi N, Ohba Y, Akiyama K, Ohga N, Hamada JI, Nagao-Kitamoto H, et al. Tumour endothelial cells in high metastatic tumours promote metastasis via epigenetic dysregulation of biglycan. Sci Rep. 2016;6(1):28039.
28. Celestini A, Violi F. A review of picotamide in the reduction of cardiovascular events in diabetic patients. Vasc Health Risk Manag. 2007;3(1):93–8.
29. Bonchev D, Mountain CF, Seitz WA. The conformational flexibility of aromatic retinoids. Biophys Chem. 1994;51(1):59–69.
30. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44–57.

31. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607-D13.
32. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res. 2017;45(W1):W98–102.
33. Han H, Cho JW, Lee S, Yun A, Kim H, Bae D, et al. TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. Nucleic Acids Res. 2018;46(D1):D380-D6.
34. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2(5):401–4.

Figures

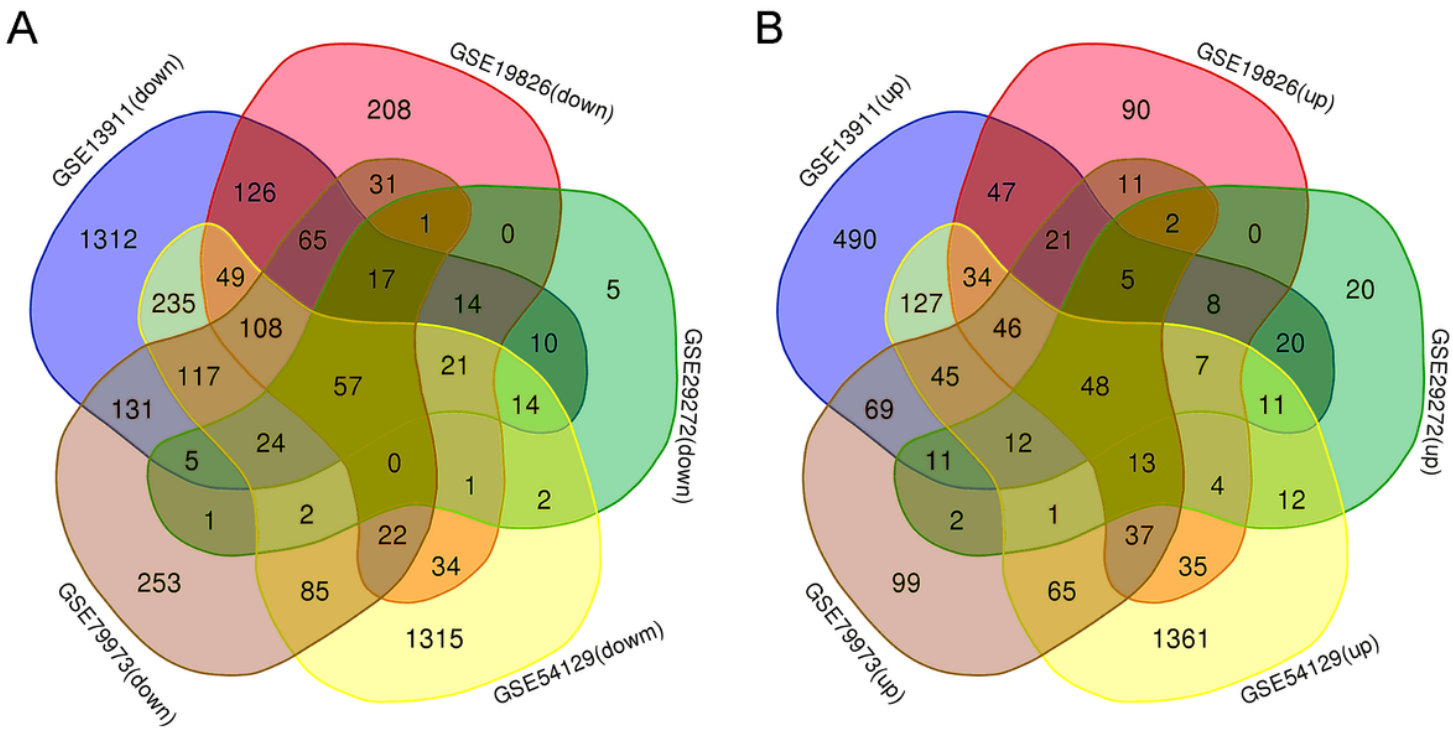

Figure 1

Selection of overlapping differentially expressed genes (DEGs). Statistically significant DEGs were acknowledged with adjust p-value<0.05 and |logFC|≥1 as the cut-off standard for each data set. The Venn diagram displayed an overlap of 287 genes. (A) A total of 57 shared downregulated genes were displayed. (B) A total of shared 48 upregulated genes were displayed.

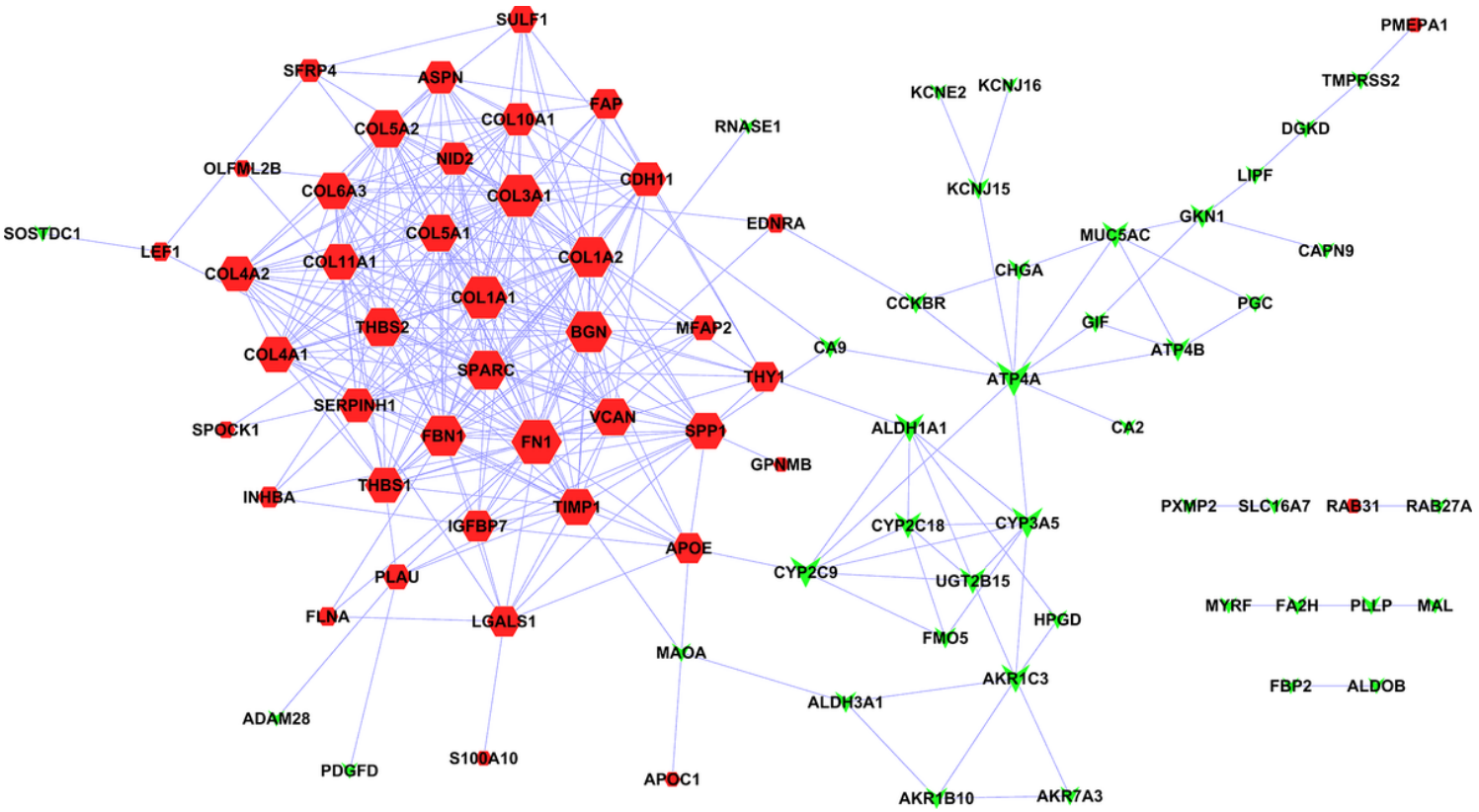

Figure 2

The outcomes of PPI network were visualized by Cytoscape software. Red color represents upregulated genes; green color represents downregulated genes.

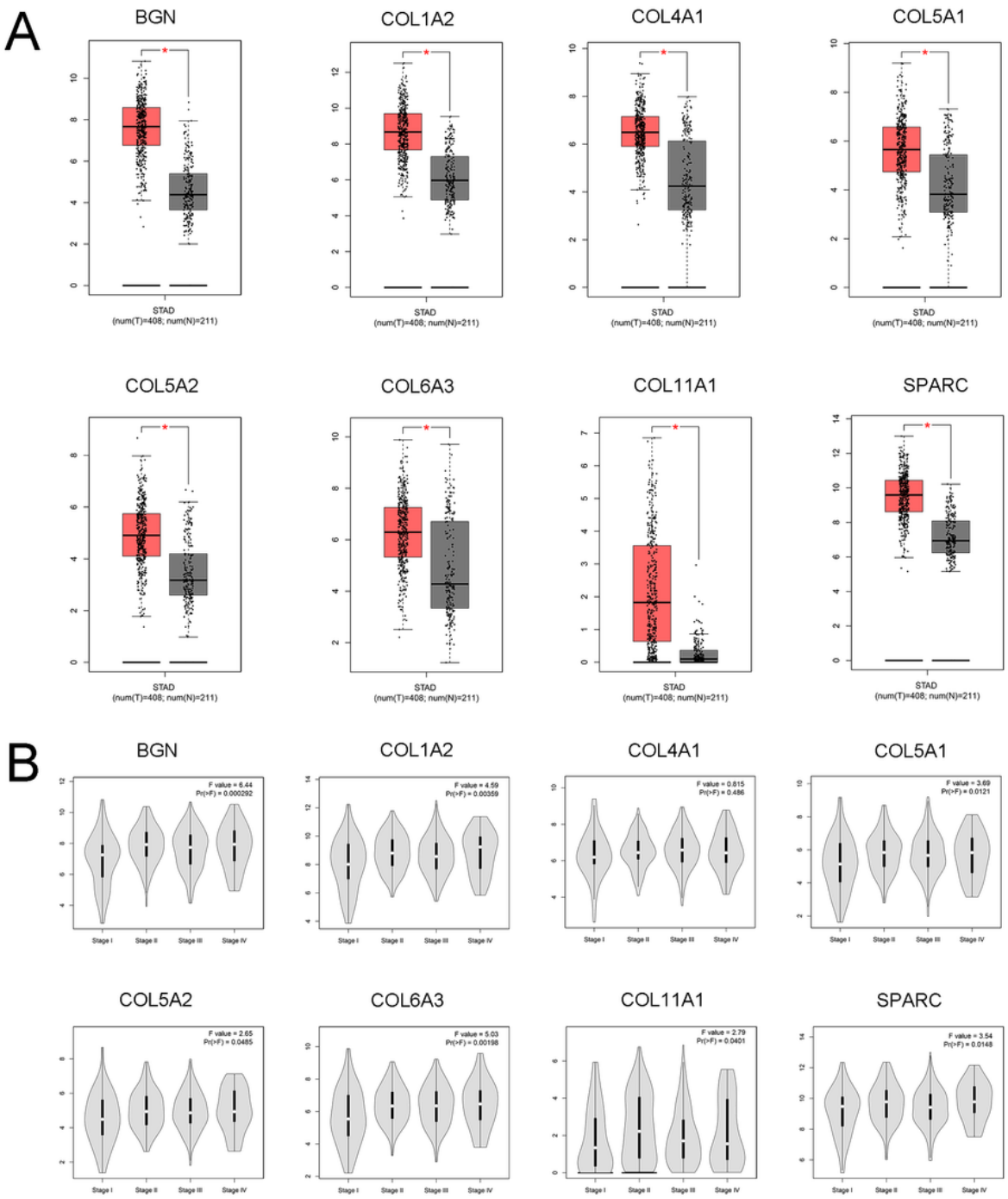

Figure 3

The expression of hub genes in GEPIA database. (A) BGN, SPARC, COL5A2, COL5A1, COL1A2, COL4A1, COL6A3 and COL11A1 were overexpressed in gastric cancer tissues compared with normal tissues. (B) The expression of BGN, SPARC, COL5A2, COL5A1, COL1A2, COL6A3 and COL11A1 in patients with stage II, III and IV and was higher than that in patients with stage I. COL4A1 were not statistically significant.

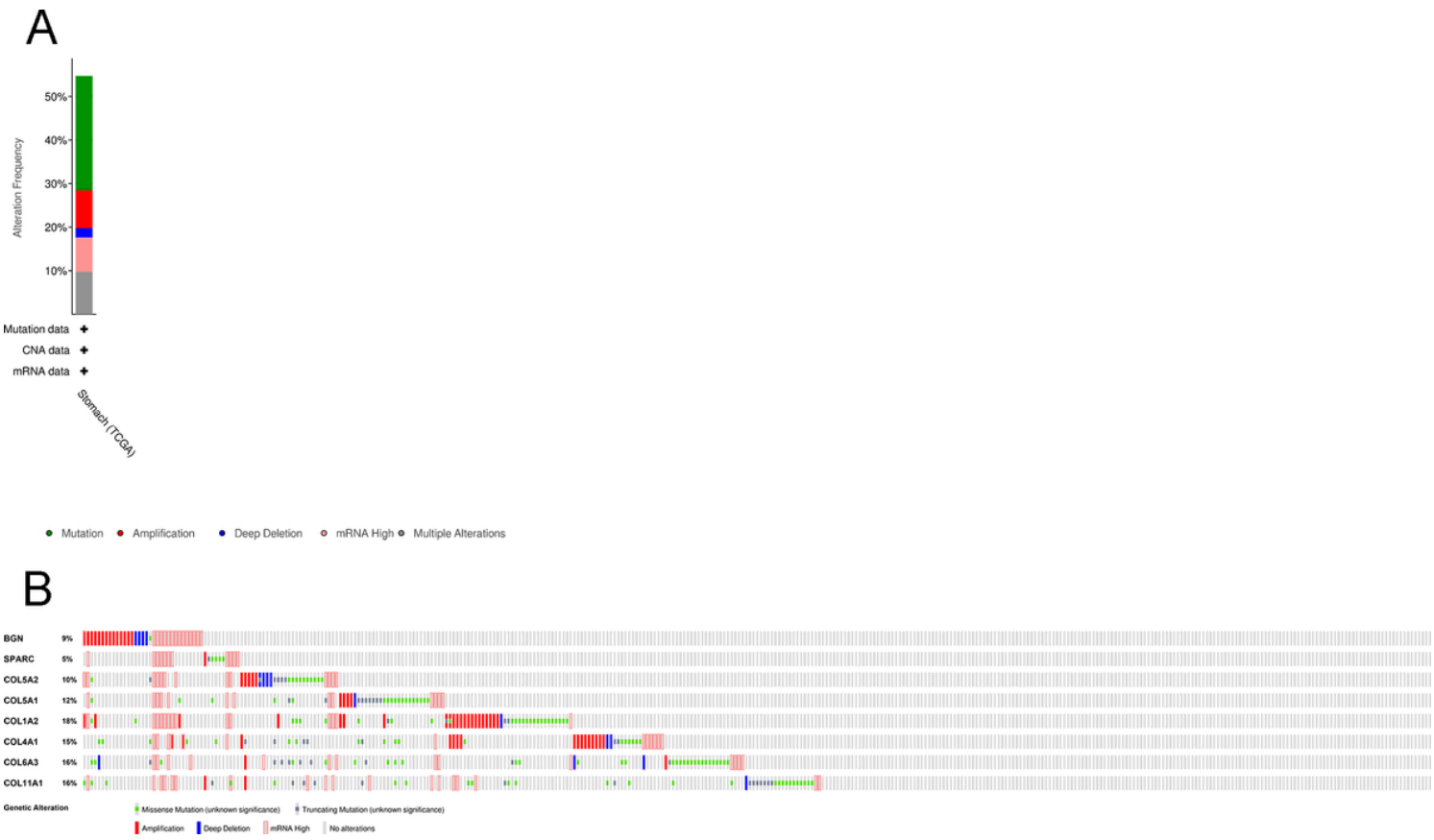

Figure 4

Mining genomic alterations of hub genes in gastric cancer using cBioPortal. (A) Overview of changes in BGN, SPARC, COL5A2, COL5A1, COL1A2, COL4A1, COL6A3 and COL11A1 genes in genomic database across a set of gastric cancer samples. (B) Alteration frequencies of seven genes (BGN, SPARC, COL5A2, COL5A1, COL1A2, COL4A1, COL6A3 and COL11A1) in cBio cancer genomics portal.

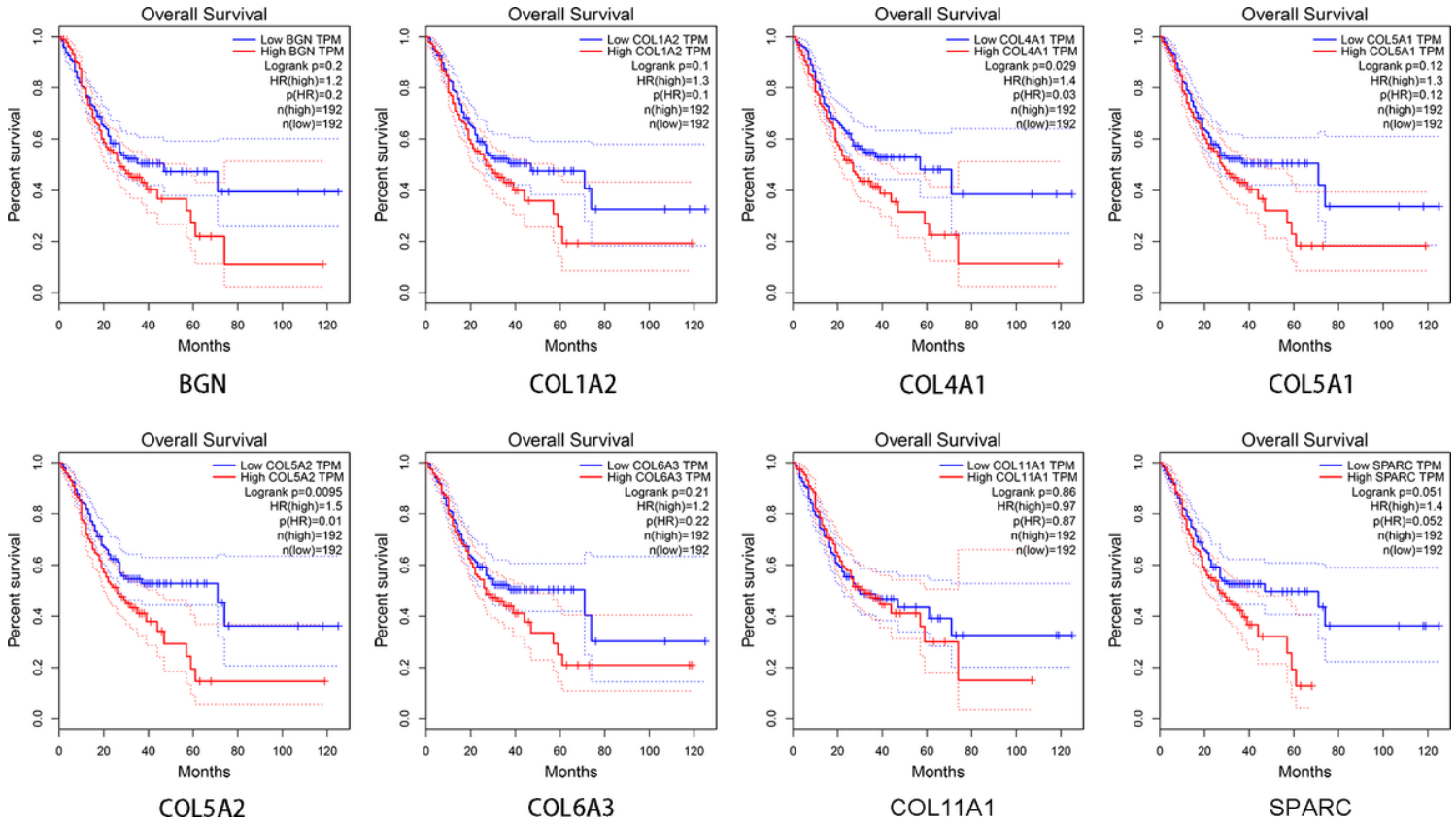

Figure 5

Survival analysis of hub genes in gastric cancer.