

Identification of hub genes in colon cancer via bioinformatics analysis

Journal of International Medical Research
48(9) 1–14

© The Author(s) 2020

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0300060520953234

journals.sagepub.com/home/imr



Jun Liu^{1,2,*}, Gui-Li Sun^{3,*}, Shang-Ling Pan⁴,
Meng-Bin Qin¹, Rong Ouyang^{1,2} and
Jie-An Huang¹ 

Abstract

Objectives: This study aimed to investigate hub genes and their prognostic value in colon cancer via bioinformatics analysis.

Methods: Differentially expressed genes (DEGs) of expression profiles (GSE33113, GSE20916, and GSE37364) obtained from Gene Expression Omnibus (GEO) were identified using the GEO2R tool and Venn diagram software. Function and pathway enrichment analyses were performed, and a protein–protein interaction (PPI) network was constructed. Hub genes were verified based on The Cancer Genome Atlas (TCGA) and Human Protein Atlas (HPA) databases.

Results: We identified 207 DEGs, 62 upregulated and 145 downregulated genes, enriched in Gene Ontology terms “organic anion transport,” “extracellular matrix,” and “receptor ligand activity”, and in the Kyoto Encyclopedia of Genes and Genomes pathway “cytokine-cytokine receptor interaction.” The PPI network was constructed and nine hub genes were selected by survival analysis and expression validation. We verified these genes in the TCGA database and selected three potential predictors (*ZG16*, *TIMP1*, and *BGN*) that met the independent predictive criteria. *TIMP1* and *BGN* were upregulated in patients with a high cancer risk, whereas *ZG16* was downregulated. The immunostaining results from HPA supported these findings.

Conclusion: This study indicates that these hub genes may be promising prognostic indicators or therapeutic targets for colon cancer.

¹Department of Gastroenterology, the Second Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi, P.R. China

²Department of Gastroenterology, the Fourth Affiliated Hospital of Guangxi Medical University/Liu Zhou Worker’s Hospital, Liuzhou, Guangxi, P.R. China

³Department of Clinical Nutrition, The Third Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi, P.R. China

⁴Departments of Pathophysiology, Guangxi Medical University, Nanning, Guangxi, P.R. China

*These authors contributed equally to this work.

Corresponding author:

Jiean Huang, Department of Gastroenterology, the Second Affiliated Hospital of Guangxi Medical University, 166 Daxuedong Road, Nanning, Guangxi 530007, P.R. China.

Email: hjagxmu@163.com



Keywords

Bioinformatic analysis, colon cancer, hub genes, *ZGI6*, *TIMPI*, *BGN*

Date received: 28 March 2020; accepted: 28 July 2020

Introduction

Colon cancer has become a major global public health concern because of its high incidence and mortality rates; more than 1.8 million new cases and 881,000 deaths were expected in 2018.¹ Several risk factors have been linked to this disease;² however, predicting prognosis in colon cancer is imperfect. The search for effective prognostic biomarkers will aid in predicting prognosis and improving colon cancer treatment.

Microarray technology can simultaneously scan hybridization signals of various gene probes and rapidly detect differentially expressed genes (DEGs).³ It also provides a platform for public databases and exploring the molecular mechanisms of multiple cancers. A high-quality microarray could potentially link molecular biomarkers to the development, diagnosis, and treatment of colon cancer.

In this study, we identified DEGs in expression profiles GSE33113, GSE20916, and GSE37364 from the Gene Expression Omnibus (GEO) database. We performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional and pathway analyses and constructed a protein–protein interaction (PPI) network. Survival analysis and expression validation of hub genes were determined by Gene Expression Profiling Interactive Analysis (GEPIA). Hub genes and the clinical prognosis of patients with colon cancer were compared and analyzed with the objective of identifying novel prognostic biomarkers and therapeutic targets.

Materials and methods

Ethical approval

The Second Affiliated Hospital of Guangxi Medical University Ethical Review Committee approved this study (Approval Number: 2020-KY[0097]). All study material was from publicly available databases and therefore patient consent was not applicable.

Data collection

Expression profiles of GSE33113, GSE20916, and GSE37364 based on the GPL570 platform were obtained from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>). The three datasets included 90, 66, and 27 colon cancer tissues and 6, 34, and 38 normal tissues, respectively.

DEG identification

DEGs of expression profiles were identified via GEO2R online tools (www.ncbi.nlm.nih.gov/geo/geo2r) and Venn diagrams (bioinformatics.psb.ugent.be/webtools/Venn). The cut-off criteria were $|\log \text{fold-change}| > 2$ and adjusted P -value < 0.05 .

GO and KEGG pathway analyses

To elucidate potential gene functional annotation and pathway enrichment, GO^{4,5} and KEGG pathway analyses^{6,7} were performed using the “enrichplot” package in R software (<https://www.R-project.org/>). GO annotation categorizes functions into biological processes (BP), molecular functions (MF), and cellular components (CC).

PPI network construction

DEGs were evaluated using the Search Tool for the Retrieval of Interacting Genes (STRING, <https://string-db.org/>), and results with a minimum interaction score of 0.4 were visualized in Cytoscape software (version 3.6.1; <https://cytoscape.org>). In addition, the degree of each protein node in the PPI network was calculated using cytoHubba,⁸ and the top 30 genes were selected as potential hub genes.

Survival analysis and expression validation of hub genes

GEPIA is a web server that analyzes cancerous and normal gene expression profiles and interactions in The Cancer Genome Atlas (TCGA; <https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>) and the Genotype-Tissue Expression (GTEx; <https://gtexportal.org/home/>) projects.⁹ The overall survival of colon cancer patients and the expression level validation of hub genes were evaluated using GEPIA. $P < 0.05$ was defined as the significant threshold to indicate candidate genes.

Verification of hub genes with TCGA database

The 30 potential hub genes were analyzed by survival analysis and expression validation using GEPIA; nine of these genes were analyzed by univariate Cox regression, in which candidate genes were selected if they satisfied the screening condition of $P < 0.05$. Thereafter, we used LASSO regression for high-dimensional data to select the most useful prognostic factors using the “glmnet” package in R software.¹⁰ By this process, three hub genes were selected and the related risk score was calculated. Patients were divided into high-risk and low-risk groups based on the median expression levels of these three genes.

Univariate and multivariate Cox regression analyses were performed to identify prognostic factors for colon cancer, and a heatmap of these three hub genes and clinical risk factors was constructed using the R package “pheatmap.”

Protein levels of hub genes in the HPA database

In our study, the protein levels of hub genes in tumors and normal tissues were explored using the HPA database (<https://www.proteinatlas.org/>), which contains immunohistochemistry-based expression data for specific human tissues.¹¹

Statistical analyses

Statistical analysis and graphs were prepared with R software (version 3.8; <https://www.R-project.org/>). Data are represented as the mean \pm standard deviation, and the *t*-test was used for comparisons between the two groups. R packages “survival” and “forestplot” were used for univariate and multivariate Cox regression analyses, and package “glmnet” was used for LASSO regression analysis. $P < 0.05$ was considered statistically significant.

Results

Identification of DEGs

The three profile datasets included 183 colon cancer tissues and 78 normal tissues. According to the cut-off criteria, 272, 184, and 301 DEGs were extracted from GSE33113, GSE20916, and GSE37364, respectively, via the GEO2R online tools. Subsequently, 207 DEGs were identified by Venn diagram software, including 62 upregulated and 145 downregulated genes (Table 1 and Figure 1).

Table I. Differentially expressed genes (DEGs) in three Gene Expression Omnibus datasets.

DEGs	Gene Name
Upregulated	SLCO4A1, UBD, PDPN, CXCL3, GDF15, TRIB3, SERPINB5, SULF1, FOXQ1, MMP3, FAM3B, SLCO1B3, CXCL5, ASCL2, INHBA, SCD, AJUBA, CXCL6, ETV4, CCL20, SPPI, COL1A1, MIR675, CXCL2, CLDN1, IL1B, SRPX2, LCN2, PSAT1, TMPRSS3, LRP8, NFE2L3, TREM1, CEMIP, SLC6A6, CTHRC1, SLC7A11, TACSTD2, FCGR3B, TESC, EGFL6, REG3A, PHLDA1, TIMP1, THBS2, SLC6A14, BGN, CXCL1, COL1A2, CXCL11, MMP1, HCAR3, ADAMTS2, WISPI, COL8A1, MMP12, CXCL8, CRNDE, LOC101928195, REG1A, PRRX1, MMP11
Downregulated	LGALS2, GREM2, SPIB, CNTN3, HSD17B2, ABCG2, MMP28, HMGCS2, FOXP2, ZG16, GUCA2B, FLJ22763, EFNA5, LDHD, UGT2B17, ITLN1, CHP2, OGN, IGLJ3, SCARA5, CLCA4, DHRS11, BTNL8, INSM1, AKR1B10, HSD11B2, SLC28A2, TUBAL3, CA4, TRPM6, IGH, ADAMDECI, PTGDR, FABP1, IGKC, CES2, ANO5, PYY, DNASE1L3, BCHE, SCIN, UGT1A1, SLC26A3, CD36, CXCL12, SLC30A10, NR5A2, B3GALT5, TSPAN7, HHLA2, RGS13, CA2, AHCYL2, DPP10-AS1, IL6R, FCGBP, CHGA, MT1M, SLC26A2, PKIB, ANPEP, CDHR5, CEACAM7, PADI2, PBLD, ADTRP, VSTM2A, NRIH4, MT1G, CLCA1, ABCA8, AMN, SLC13A2, SLC16A9, BEST2, CLDN23, INSL5, SLC51B, EDN3, ADH1B, AQP8, BRINP3, UGT1A3, GCG, SCUBE2, LYPD8, SOSTDC1, TNFRSF17, CD177, GBA3, SCGN, CLDN8, MS4A12, TPH1, FXYD1, PCK1, VSIG2, ADHIC, TMEM72, ABI3BP, SLC17A4, HEPACAM2, UGT2A3, GCNT2, SST, RUND3B, BMS1P20, ABCC13, PLAC8, SCNN1B, NXPE1, C2orf88, PLP1, EPB41L3, CNNM2, CDKN2B, RETNLB, HPGD, LAMA1, LOC101929340, ZBTB7C, CWH43, KANK4, MAMDC2, BEST4, CA1, STMN2, TEX11, SCGB2A1, SLC4A4, IGK, SFRP1, SELENBP1, MOGAT2, CA12, SI, SLC51A, GUCA2A, UGT1A3, TMEM100, HAGLR, DHRS9, HAPLN1, CA7, ENTPD5

DEG, differentially expressed gene; GEO, Gene Expression Omnibus.

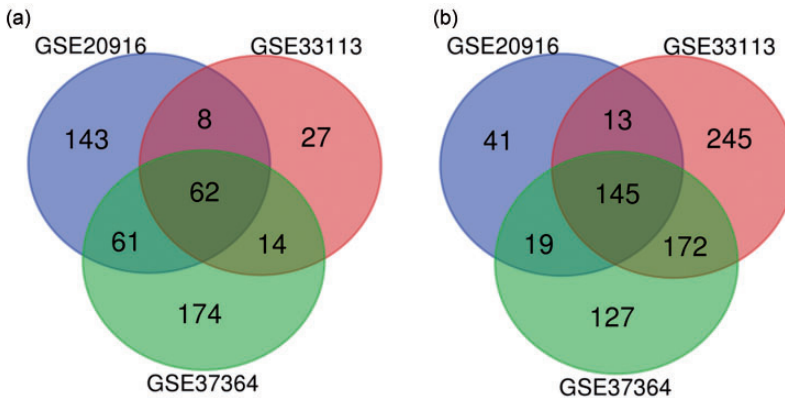


Figure 1. Venn diagram of DEGs from three GEO datasets. (a) Upregulated genes (n = 62), (b) downregulated genes (n = 145).

Functional enrichment analysis of DEGs

To elucidate potential gene functional annotation and pathway enrichment, GO

categories BP, CC, and MF were analyzed using R software; the top 10 GO terms among 207 DEGs are shown in Figure 2A. Regarding the KEGG analysis,

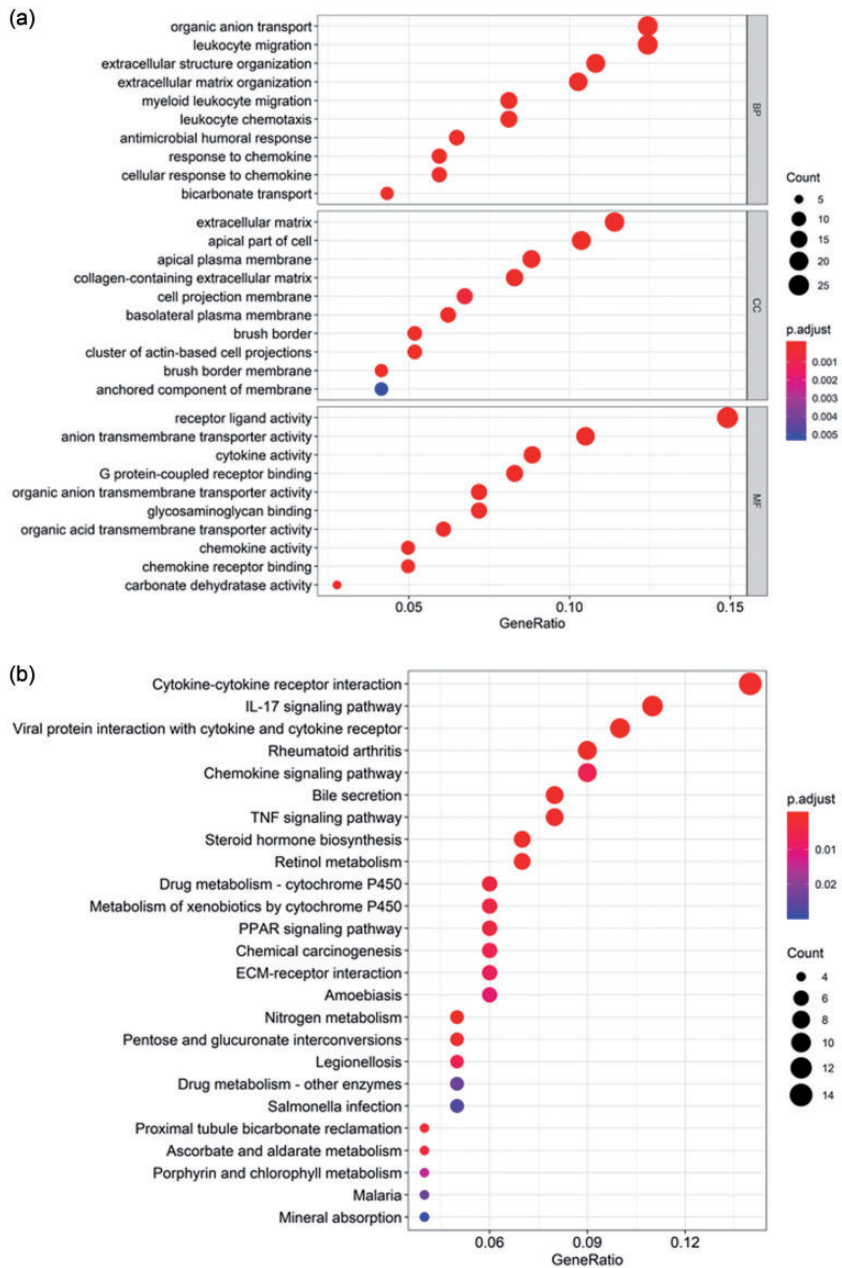


Figure 2. GO and KEGG enrichment analyses. (a) Top 10 GO term enrichment analysis of DEGs, including BP, CC, and MF; (b) KEGG pathway enrichment analysis of DEGs.

GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEG, differentially expressed gene; BP, biological process; CC, cellular component; MF, molecular function.

the 25 main terms are shown in Figure 2B, and these indicated that DEGs were particularly enriched in cytokine–cytokine receptor interaction, IL-17 signaling pathway, viral protein interaction with cytokines and cytokine receptors, rheumatoid arthritis, and chemokine signaling pathway ($P < 0.05$).

PPI network construction

A PPI network complex that included 189 nodes and 424 edges was constructed using STRING. In this PPI network complex, there were 152 DEGs, with 48 upregulated and 104 downregulated genes (Figure 3A). Finally, the top 30 genes were identified by cytoHubba as potential hub genes (Figure 3B).

Survival analysis and expression levels of hub genes

Survival analysis and expression level analysis of the 30 hub genes were performed using GEPIA (<http://gepia.cancer-pku.cn/index.html>). The survival analysis of 10 genes was statistically significant ($P < 0.05$). Patients with high expression of *CLCA1*, *CXCL2*, *CXCL3*, *CXCL8*, *IL1B*, and *ZG16* demonstrated better survival; in contrast, patients with high expression of *BGN*, *COL1A2*, *THBS2*, and *TIMP1* demonstrated poor survival (Figure 4). Results from the GEPIA database revealed that the expression levels of nine genes differed significantly between normal tissues and tumors ($P < 0.05$). The expression levels of *CXCL2*, *CXCL3*, *CXCL8*, *IL1B*, *BGN*, *THBS2*, and *TIMP1* significantly increased in tumors, whereas those of *CLCA1* and *ZG16* decreased ($P < 0.05$) (Figure 5).

Construction of LASSO model

We used univariate Cox regression to analyze the nine genes, and three candidate

genes were selected after corresponding to a screening condition of $P < 0.05$ (Figure 6A). The LASSO Cox regression model was adopted to select the most predictive genes as prognostic indicators; λ was selected when the median of the sum of squared residuals was the smallest. Three hub genes (*BGN*, *TIMP1*, and *ZG16*) were identified as prognostic factors for colon cancer (Figure 6B, C). The risk scores of these three genes were also calculated for further univariate and multivariate Cox regression analyses.

Prognostic values of the three hub genes

Univariate analysis showed that age, clinical stage, T stage, M stage, N stage, and risk score of hub genes could affect the prognosis of patients ($P < 0.05$). Sex was not correlated with colon cancer prognosis (Figure 7A). The multivariate regression analysis demonstrated that age, clinical stage, and risk score of the three hub genes were independent prognostic factors in colon cancer ($P < 0.05$) (Figure 7B). As shown in Figure 7C, the protective *ZG16* gene had a tendency to be upregulated in low-risk patients, whereas *TIMP1* and *BGN* had a tendency to be highly expressed in high-risk patients. M stage and survival status were linked with the degree of risk ($P < 0.05$), whereas T stage, N stage, clinical stage, sex, and age were not significant.

Validation of the three hub genes with the HPA database

As shown in Figure 8, expression levels of the three hub genes, *ZG16*, *TIMP1*, and *BGN*, were validated by immunostaining data from the HPA database. Compared with normal tissue, expression levels of *BGN* and *TIMP1* increased in colon cancer tissue, whereas that of *ZG16* decreased.

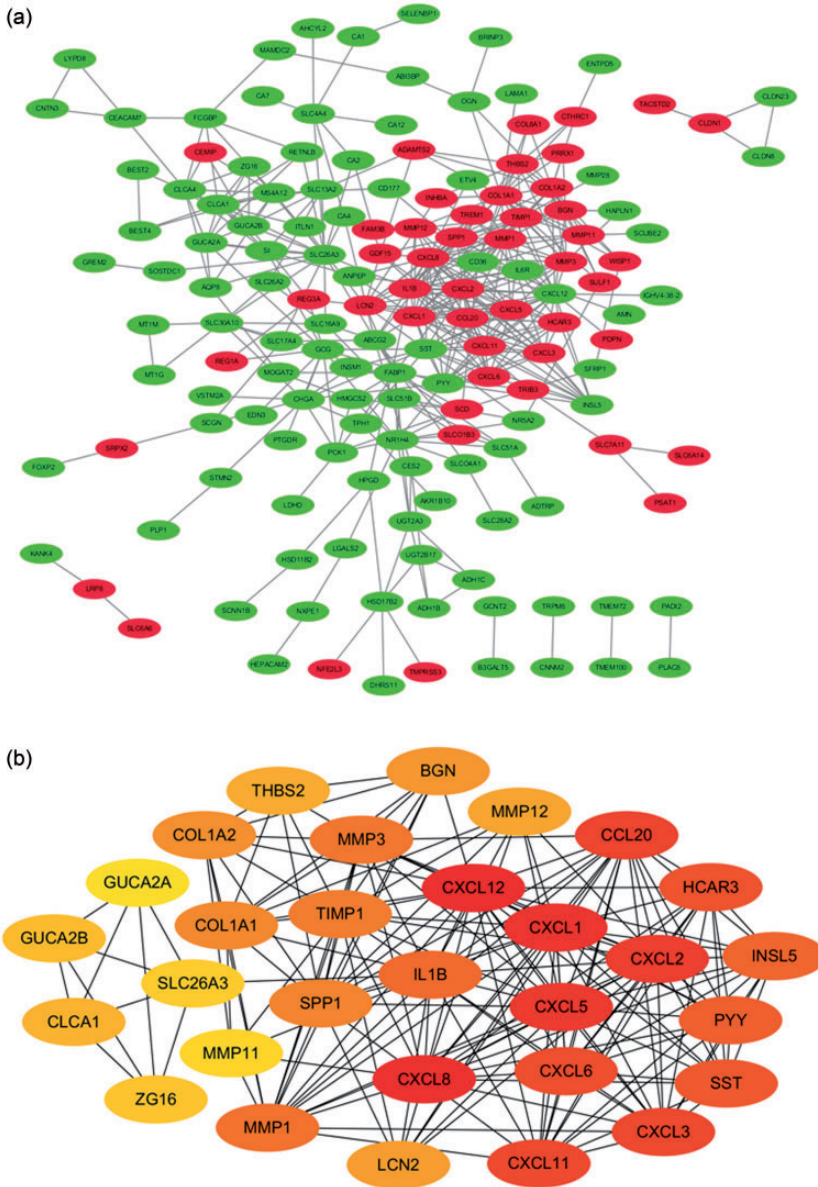


Figure 3. PPI network and hub gene identification. (a) PPI networks of DEGs: 48 upregulated genes are shown in red and 104 downregulated genes are shown in green; (b) the top 30 genes were selected from the PPI network.

PPI, protein–protein interaction; DEG, differentially expressed gene.

Discussion

To discover more effective prognostic biomarkers for colon cancer, a worldwide

public health problem, we analyzed three datasets (GSE33113, GSE20916, and GSE37364) from the GEO database. A total of 207 DEGs were identified; these

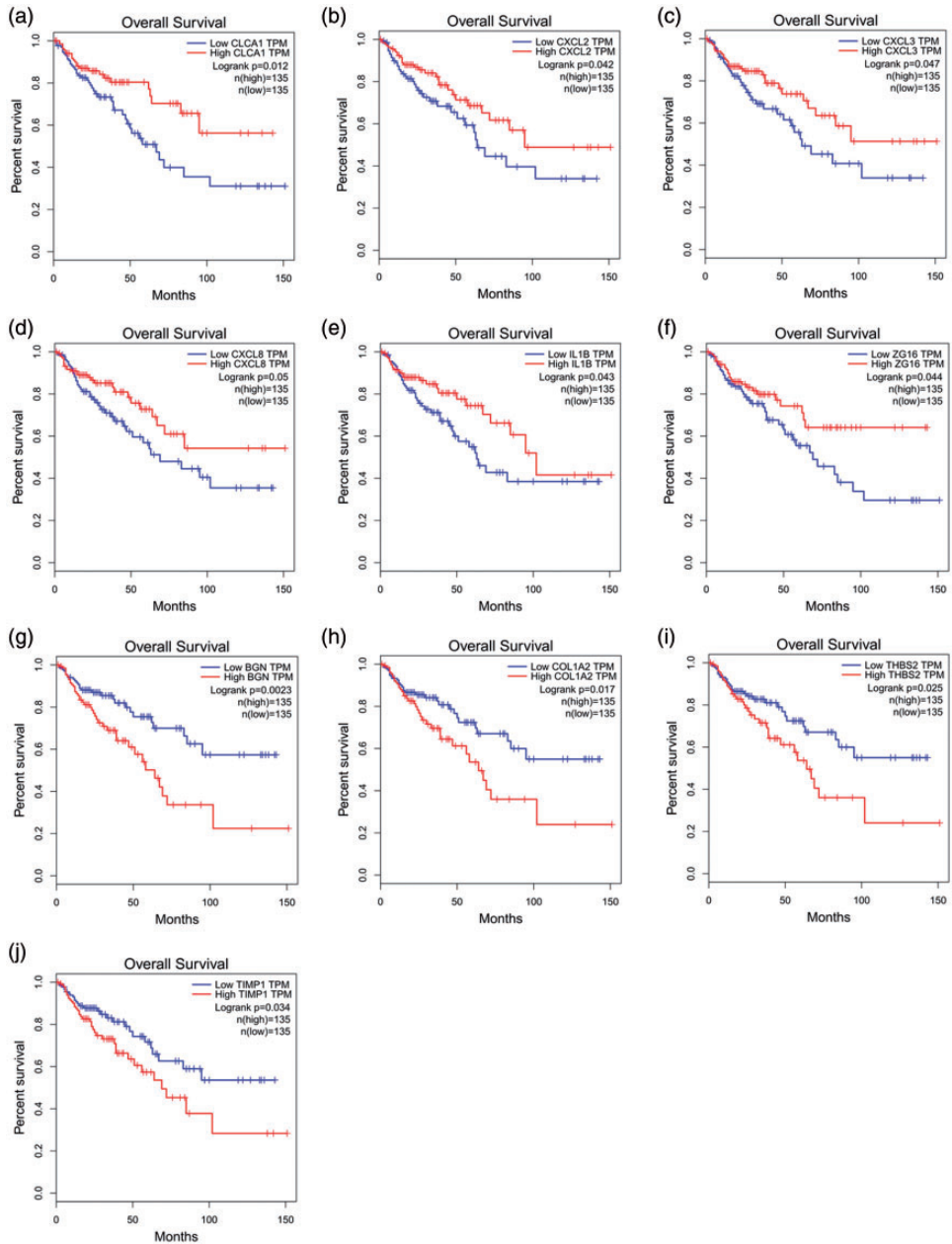


Figure 4. Survival analysis of hub genes by GEPIA. Six genes (*CLCA1*, *CXCL2*, *CXCL3*, *CXCL8*, *IL1B*, and *ZG16*) were related to better survival (a–e), and four genes (*BGN*, *COL1A2*, *THBS2*, and *TIMP1*) showed poor survival (f–j) ($P < 0.05$).

GEPIA, Gene Expression Profiling Interactive Analysis; TPM, transcripts per million.

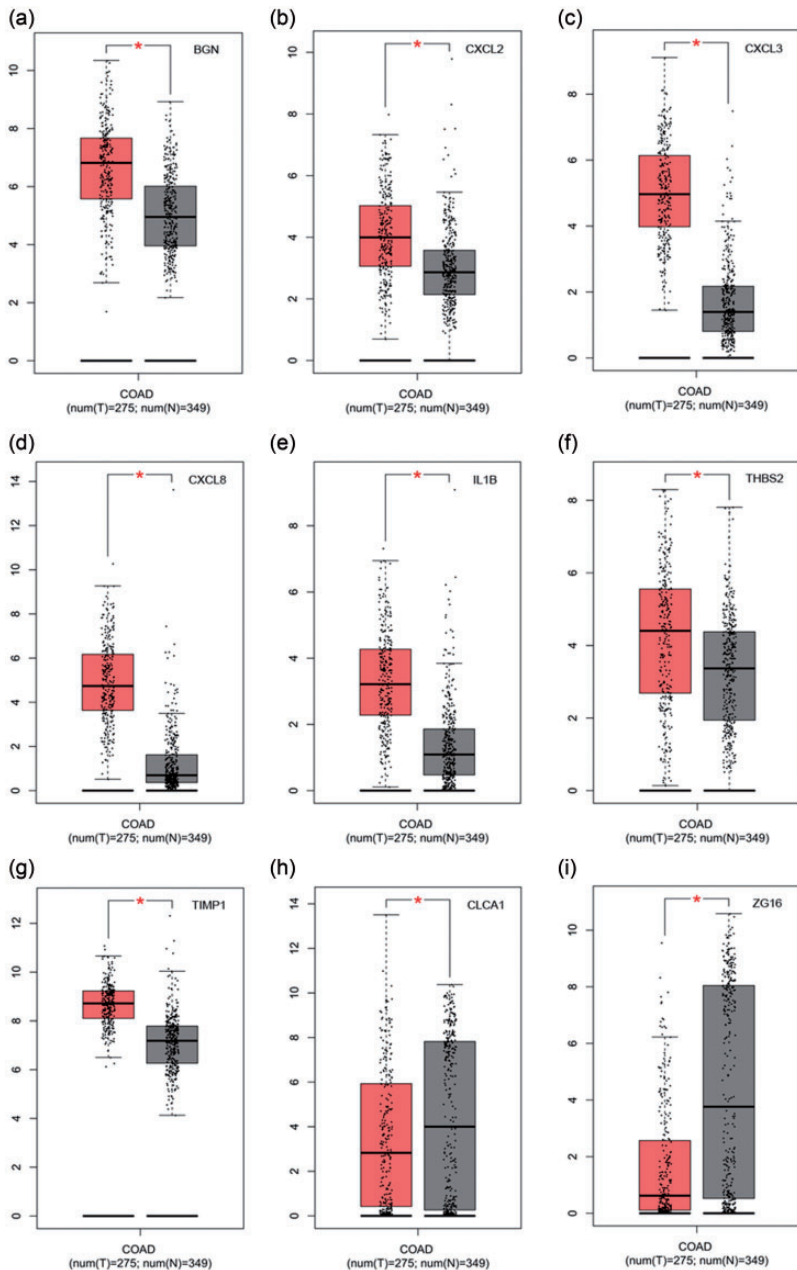


Figure 5. Expression of hub genes in different tissues by GEPIA. Red indicates tumor tissue and gray indicates normal tissue. (a) *BGN*, (b) *CXCL2*, (c) *CXCL3*, (d) *CXCL8*, (e) *IL1B*, (f) *THBS2*, (g) *TIMP1*, (h) *CLCA1*, and (i) *ZG16*. The box represents the interquartile range of data, the middle line is the median, and the upper and lower whiskers indicate the maximum and minimum values, respectively. Each dot represents a patient sample. Red asterisks indicate $P < 0.05$.

GEPIA, Gene Expression Profiling Interactive Analysis; COAD, colorectal adenocarcinoma.

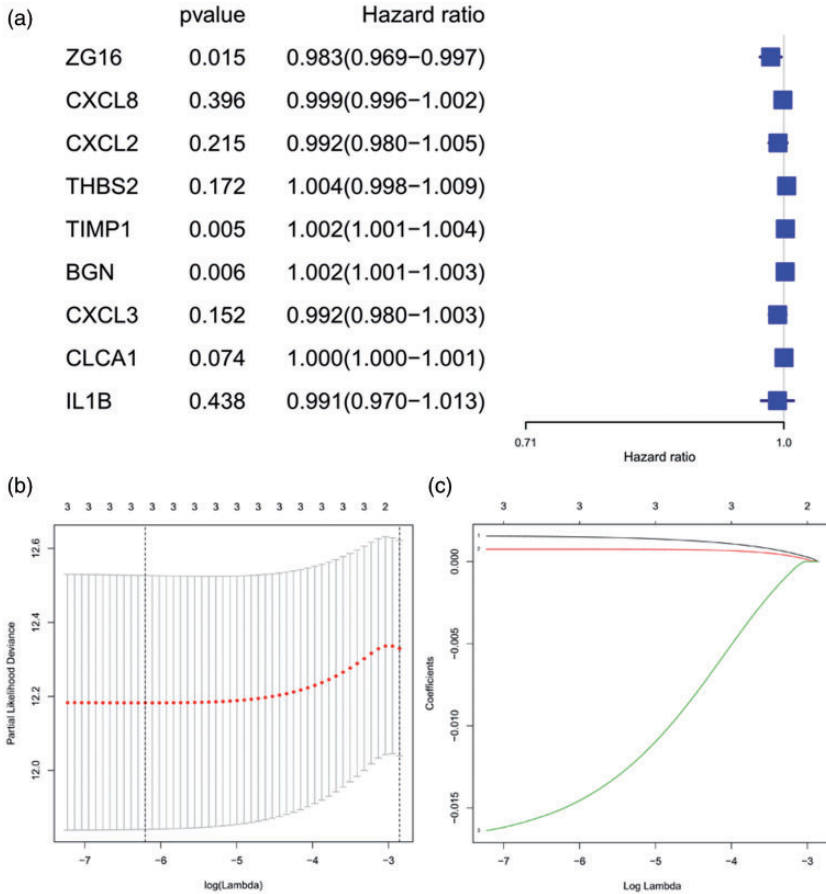


Figure 6. Selection of genes in colon cancer prognosis prediction. (a) Forest plot for hazard ratios of survival-associated hub genes in colon cancer; (b) partial likelihood deviance versus $\log(\lambda)$ was drawn using a LASSO Cox regression model; (c) coefficients of selected features shown by the lambda parameter.

were significantly enriched in the GO terms “organic anion transport” (BP), “extracellular matrix” (CC), and “receptor ligand activity” (MF). In the KEGG pathway analysis, DEGs were primarily enriched in “cytokine-cytokine receptor interaction” and “IL-17 signaling pathway.” Cytokines are crucial components of the extracellular environment and they influence tumor cell proliferation, invasion, and migration.¹² In recent years, the development of colon cancer treatment methods based on cytokines such as interleukin (IL-17) has attracted much attention.¹³

Using integrated bioinformatic analyses, nine hub genes were selected and verified based on TCGA. Ultimately, we found that three genes (*ZG16*, *TIMP1*, and *BGN*) could be considered potential targets for prediction of colon cancer prognosis. High expression levels of *TIMP1* and *BGN* were related to worse clinical outcomes, whereas higher expression of *ZG16* might predict better survival in colon cancer patients. Expression levels of these three genes were validated with immunostaining data from the HPA database.

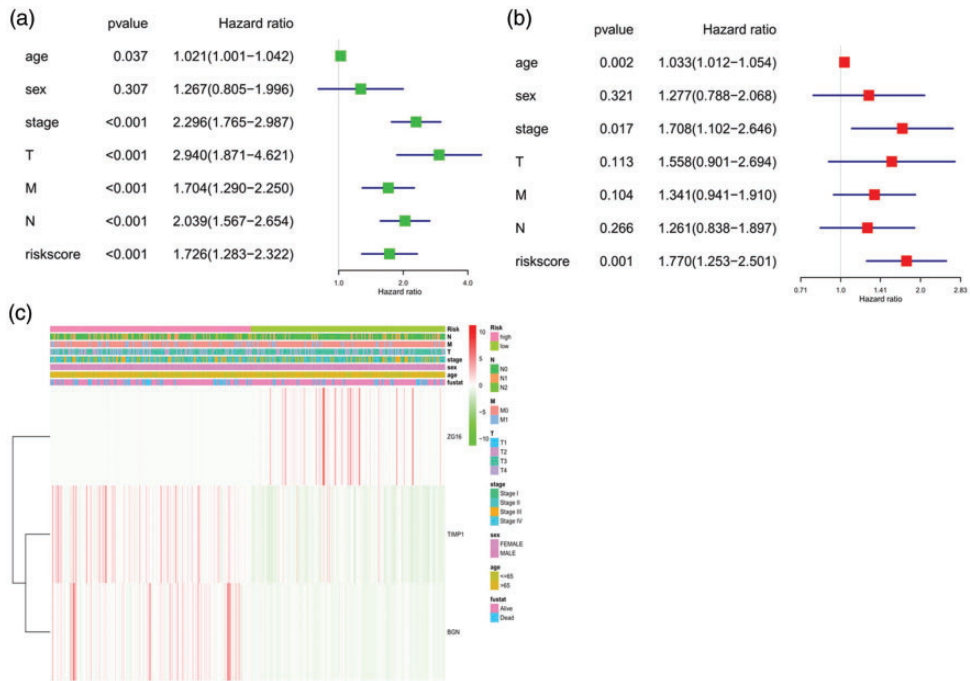


Figure 7. Forest plot and heatmap of hub genes and clinical risk factors (age, sex, stage, TMN stages, and risk score). (a) Forest plot of the univariate Cox regression analysis in colon cancer; (b) forest plot of the multivariate Cox regression analysis in colon cancer; (c) heatmap of hub genes *ZG16*, *TIMP1*, and *BGN* and clinical risk factors.

Zymogen granule protein 16 (*ZG16*) is expressed in mucus-secreting cells of the digestive system such as in the liver, colon, small intestine, and pancreas.¹⁴ *ZG16* is an important component of the protective mucus layer, which keeps bacteria away from host colon epithelium.¹⁵ The lectin-like domain of *ZG16* can stimulate the T cells of the intestinal immune system and participate in tumor recognition. Compared with that in normal colon tissue, *ZG16* expression in colon cancer is significantly decreased. Meng et al.¹⁶ reported that low levels of *ZG16* mRNA were correlated with poor survival in colon cancer. Previous studies have found that *ZG16* was further downregulated in patients with distant metastasis and this loss of protein expression was regulated

by the microRNA miR-196a.¹⁷ However, the role of *ZG16* in colon cancer is still poorly documented and further research is needed.

Tissue inhibitor of metalloproteinases (*TIMP-1*), one of the most important regulators of matrix metalloproteinases (*MMPs*), is markedly increased in malignant tissue¹⁸ and expressed in fibroblast-like cells located at the invasive front.¹⁹ *TIMP-1* is also associated with poor prognosis in colon cancer patients²⁰ and positively correlates with lymph node metastasis and Dukes classification.²¹ *TIMP-1* can promote the migration and invasion of cancer cells via reduction of cell apoptosis and activation of cell growth. Song et al.²² found that downregulation of *TIMP1* decreased proliferation

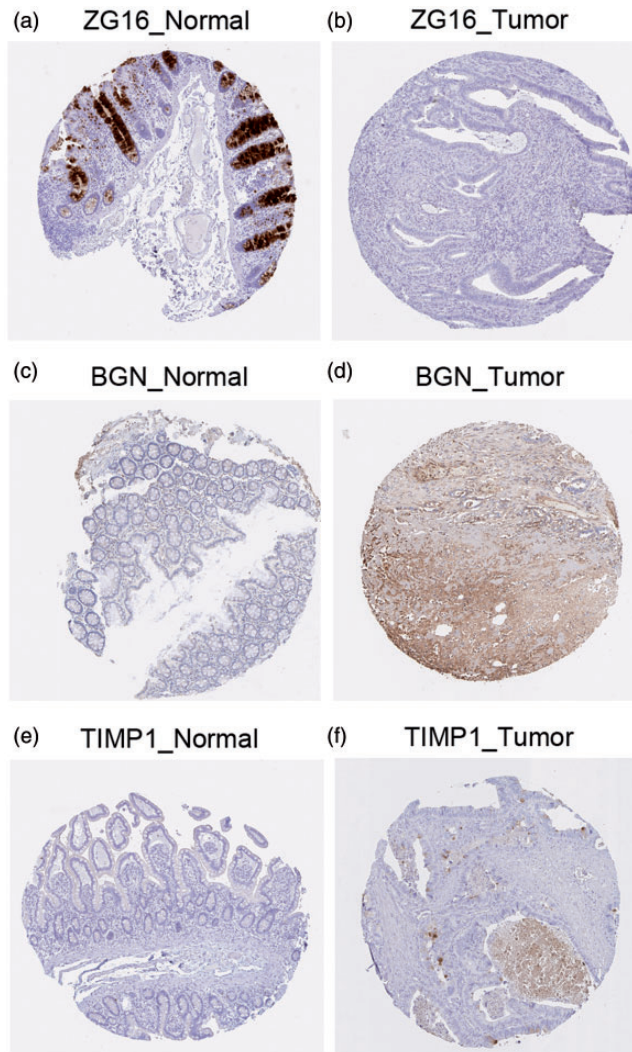


Figure 8. Validation of the three hub genes: (a, b) *ZG16*, (c, d) *BGN*, and (e, f) *TIMP1* with immunohistochemistry of normal and tumor tissue from the Human Protein Atlas database.

and metastasis in human colon cancer by inducing the focal adhesion kinase (FAK)-phosphatidylinositol 3-kinase (PI3K)-Akt signaling pathway and mitogen-activated protein kinase (MAPK) pathway. Therefore, a high TIMP-1 level can predict poor prognosis of colon cancer, which is consistent with our findings.

Biglycan (BGN) is a proteoglycan of the extracellular matrix that plays an important

role in the biological function of epithelial cells²³ and is the downstream target of a variety of signal transduction pathways.²⁴ The upregulation of *BGN* has been reported to lead to poor prognosis in several malignant epithelial tumors. Previous reports have demonstrated that BGN is associated with lymph node metastasis and distant metastasis of colon cancer.²⁵ BGN may be involved in the epithelial-mesenchymal

transition process by the integrated pathway of transforming growth factor- β /Snail with tumor necrosis factor- α /nuclear factor- κ B,²⁶ and it exerts an antiapoptotic effect via the p38 MAPK signaling pathway in colon cancer.²⁷ Our study demonstrated that *BGN* is also an indicator of poor prognosis in colon cancer patients.

Conclusions

Through bioinformatics analyses, we demonstrated that *ZG16*, *TIMP-1*, and *BGN* are hub genes related to clinical status in patients with colon cancer, and these genes are potential biomarkers for prognosis of colon cancer. Although the possible biological mechanisms through which these relationships occur require verification by *in vivo* and *in vitro* analyses, our study provides a new direction for further studies on colon cancer.

Acknowledgements

The authors thank the staff members of Department of Gastroenterology, the Second Affiliated Hospital of Guangxi Medical University, for their technical assistance and kind encouragement.

Declaration of conflicting interest

The authors declare that there is no conflict of interest.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by grants from Guangxi Health Department Scientific Research Program (No. Z20200153).

ORCID iD

Jie-An Huang  <https://orcid.org/0000-0003-0431-1888>

References

1. 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. DOI: 10.3322/caac.21492.
2. Siegel RL, Miller KD, Fedewa SA, et al. Colorectal cancer statistics, 2017. *CA Cancer J Clin* 2017; 67: 177–193. DOI: 10.3322/caac.21395.
3. Vogelstein B, Papadopoulos N, Velculescu VE, et al. Cancer genome landscapes. *Science* 2013; 339: 1546–1558. DOI: 10.1126/science.1235122.
4. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000; 25: 25–29. DOI: 10.1038/75556.
5. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res* 2019; 47: D330–D338. DOI: 10.1093/nar/gky1055.
6. Kanehisa M, Sato Y, Furumichi M, et al. New approach for understanding genome variations in KEGG. *Nucleic Acids Res* 2019; 47: D590–D595. DOI: 10.1093/nar/gky962.
7. Kanehisa M, Furumichi M, Tanabe M, et al. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* 2017; 45: D353–D361. DOI: 10.1093/nar/gkw1092.
8. Chin CH, Chen SH, Wu HH, et al. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol* 2014; 8: S11. DOI: 10.1186/1752-0509-8-s4-s11.
9. 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. DOI: 10.1093/nar/gkx247.
10. Sauerbrei W, Royston P and Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. *Stat Med* 2007; 26: 5512–5528. DOI: 10.1002/sim.3148.
11. Asplund A, Edqvist PH, Schwenk JM, et al. Antibodies for profiling the human

- proteome-The Human Protein Atlas as a resource for cancer research. *Proteomics* 2012; 12: 2067–2077. DOI: 10.1002/pmic.201100504.
12. Wang X, Yang L, Huang F, et al. Inflammatory cytokines IL-17 and TNF- α up-regulate PD-L1 expression in human prostate and colon cancer cells. *Immunol Lett* 2017; 184: 7–14. DOI: 10.1016/j.imlet.2017.02.006.
 13. Razi S, Baradaran Noveiry B, Keshavarz-Fathi M, et al. IL-17 and colorectal cancer: From carcinogenesis to treatment. *Cytokine* 2019; 116: 7–12. DOI: 10.1016/j.cyto.2018.12.021.
 14. Tateno H, Yabe R, Sato T, et al. Human ZG16p recognizes pathogenic fungi through non-self polyvalent mannose in the digestive system. *Glycobiology* 2012; 22: 210–220. DOI: 10.1093/glycob/cwr130.
 15. Bergstrom JH, Birchenough GM, Katona G, et al. Gram-positive bacteria are held at a distance in the colon mucus by the lectin-like protein ZG16. *Proc Natl Acad Sci U S A* 2016; 113: 13833–13838. DOI: 10.1073/pnas.1611400113.
 16. Meng H, Li W, Boardman LA, et al. Loss of ZG16 is associated with molecular and clinicopathological phenotypes of colorectal cancer. *BMC Cancer* 2018; 18: 433. DOI: 10.1186/s12885-018-4337-2.
 17. Chen X, Du P, She J, et al. Loss of ZG16 is regulated by miR-196a and contributes to stemness and progression of colorectal cancer. *Oncotarget* 2016; 7: 86695–86703. DOI: 10.18632/oncotarget.13435.
 18. Seo DW, Li H, Qu CK, et al. Shp-1 mediates the antiproliferative activity of tissue inhibitor of metalloproteinase-2 in human microvascular endothelial cells. *J Biol Chem* 2006; 281: 3711–3721. DOI: 10.1074/jbc.M509932200.
 19. Holten-Andersen MN, Hansen U, Brunner N, et al. Localization of tissue inhibitor of metalloproteinases 1 (TIMP-1) in human colorectal adenoma and adenocarcinoma. *Int J Cancer* 2005; 113: 198–206. DOI: 10.1002/ijc.20566.
 20. Bockelman C, Beilmann-Lehtonen I, Kaprio T, et al. Serum MMP-8 and TIMP-1 predict prognosis in colorectal cancer. *BMC Cancer* 2018; 18: 679. DOI: 10.1186/s12885-018-4589-x.
 21. Mroczko B, Groblewska M, Okulczyk B, et al. The diagnostic value of matrix metalloproteinase 9 (MMP-9) and tissue inhibitor of matrix metalloproteinases 1 (TIMP-1) determination in the sera of colorectal adenoma and cancer patients. *Int J Colorectal Dis* 2010; 25: 1177–1184. DOI: 10.1007/s00384-010-0991-9.
 22. Song G, Xu S, Zhang H, et al. TIMP1 is a prognostic marker for the progression and metastasis of colon cancer through FAK-PI3K/AKT and MAPK pathway. *J Exp Clin Cancer Res* 2016; 35: 148. DOI: 10.1186/s13046-016-0427-7.
 23. Frantz C, Stewart KM and Weaver VM. The extracellular matrix at a glance. *J Cell Sci* 2010; 123: 4195–4200. DOI: 10.1242/jcs.023820.
 24. Wang H, Sun W, Ma J, et al. Biglycan mediates suture expansion osteogenesis via potentiation of Wnt/ β -catenin signaling. *J Biomech* 2015; 48: 432–440. DOI: 10.1016/j.jbiomech.2014.12.032.
 25. Gu X, Ma Y, Xiao J, et al. Up-regulated biglycan expression correlates with the malignancy in human colorectal cancers. *Clin Exp Med* 2012; 12: 195–199. DOI: 10.1007/s10238-011-0155-4.
 26. Li H, Zhong A, Li S, et al. The integrated pathway of TGF β /Snail with TNF α /NF κ B may facilitate the tumor-stroma interaction in the EMT process and colorectal cancer prognosis. *Sci Rep* 2017; 7: 4915. DOI: 10.1038/s41598-017-05280-6.
 27. Xing X, Gu X and Ma T. Knockdown of biglycan expression by RNA interference inhibits the proliferation and invasion of, and induces apoptosis in, the HCT116 colon cancer cell line. *Mol Med Rep* 2015; 12: 7538–7544. DOI: 10.3892/mmr.2015.4383.