Oncology and Translational Medicine
DOI 10.1007/s10330-021-0535-5
October 2022, Vol. 8, No. 5, P232–P238

ORIGINAL ARTICLE

Bioinformatics analysis of potential hub genes associated with biological characteristics and survival in patients with gastric cancer*

Libo Feng[1], Liu Yu[2], Xiaolong Chen[1] (✉)

[1] Department of Gastrointestinal Surgery, The Affiliated Hospital of Southwest Medical University, Luzhou 646000, China
[2] Department of Electrocardiogram, The Affiliated Hospital of Southwest Medical University, Luzhou 646000, China

Abstract

Objective Gastric cancer (GC) is a serious threat to human health. In this study, we aimed to explore the differentially expressed genes (DEGs) and identify potential targets for the treatment of GC.

Methods The gene expression profile of GSE79973 which compared tissue samples from gastric cancer patients and healthy individuals, downloaded from the GEO database, was submitted to the GCBI online analysis platform to screen for DEGs. Gene ontology (GO) analysis, pathway analysis, and construction of networks, including gene signal and gene co-expression networks, were performed to identify the core DEGs. Survival analysis was performed to determine the relationship between these genes and patient survival time.

Results Nine hundred eighty-three genes were identified as DEGs ($P < 0.001$; FC > 2). GO analysis showed that DEGs were primarily involved in processes such as angiogenesis, cell metabolism, cell adhesion, redox processes, and cell migration. The metabolism of xenobiotics by cytochrome P450, ECM-receptor interaction, drug metabolism by cytochrome P450, metabolic pathways, and the PI3K-Akt signaling pathway were significantly enriched in pathway analysis. Genes such as *UGT2B15*, Hepatocyte growth factor (HGF), Nidogen-2 (NID2), Follistatin-like protein 1 (FSTL1), and Inhibin beta A chain (INHBA) were closely linked to other genes in the network. Survival analyses indicated that *HGF*, *NID2*, *FSTL1*, and *INHBA* expression levels were inversely correlated with survival time in patients with gastric cancer.

Conclusion *HGF*, *NID2*, *FSTL1*, and *INHBA* may be potential key genes associated with the biological characteristics and survival in patients with gastric cancer.

Received: 16 November 2021
Revised: 21 January 2022
Accepted: 15 March 2022

Key words: gastric cancer; differentially expressed genes; enrichment analysis; bioinformatics

Tumors are a global health problem. Gastric cancer is the third leading cause of cancer-related deaths worldwide [1]. In China, the incidence of gastric cancer is approximately 679.1/100,000, and the mortality rate is as high as 498/100,000, making gastric cancer the second most common malignant cancer in China [2]. In the United States, 26,370 new cases of gastric cancer were estimated in 2016, of which 10,370 people were estimated to have died [3]. Presently, the preferred treatment for gastric cancer is radical surgery, usually combined with systemic chemotherapy in the perioperative period [4, 5]. The lack of specific performance resulted in most patients being in the late stage of diagnosis. Although, the five-year mortality rate of early gastric cancer has reduced in recent years, it remains 30%–50% in advanced gastric cancer cases [6]. The rapid development of genomics and gene chip technology has led to the occurrence, development, and prognosis of gastric cancer being studied at the gene level. Bioinformatics technology uses multiple analytical methods (including gene ontology (GO) and pathway analyses) and biological networks (including co-expression, signal, and protein interaction networks) to identify potential core genes from thousands of differentially expressed genes (DEGs).

✉ Correspondence to: Xiaolong Chen. Email: 1870747218@qq.com
*Supported by a grant from the Project of Southwest Medical University-Three Affiliated Hospitals (No. 2017-ZRQN-028).

In this study, DEGs were screened based on the gene expression profiles of gastric cancer and normal tissues, which were downloaded from the gene expression omnibus (GEO) database. Various bioinformatics analyses and biological networks were applied to further filter out the core DEGs that may serve as molecular markers and potential therapeutic targets for gastric cancer, and this is also helpful in understanding the molecular mechanisms underlying gastric cancer development.

Materials and methods

Microarray data

The gene expression data set GSE79973 [7] was downloaded from the GEO data repository (http://www.ncbi.nlm.nih.gov/geo/). Twenty samples were included in this dataset: 10 gastric cancer and 10 normal tissues. Gene expression profiles were obtained using the GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array.

Data preprocessing and screening of DEGs

Gene-Cloud of Biotechnology Information (GCBI) (https://www.gcbi.com.cn/gclib/html/index), an online genetic data analysis software based on the R language, was used to normalize and analyze the gene chip data. The cut-offs to filter out differentially expressed genes were set to $P < 0.001$ and fold change (FC) > 2.

Function and pathway enrichment of DEGs

Gene ontology analysis, which annotates and classifies genes according to biological pathways, molecular functions, and cell locations [8], is commonly used in functional studies [9]. DEGs were enriched in various biological functions, pathways, and cell localization through GO analysis and were easily dertermined by GO analysis. Gene expression data of gastric cancer were submitted to GCBI for functional enrichment analysis to uncover biological processes. DEGs with a false discovery rate (FDR) < 0.05 and $P < 0.05$ were considered significant. The Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.genome.jp/kegg/), established by the Bioinformatics Center at Kyoto University, Japan, is a useful database resource for genome sequencing and other high-throughput experimental techniques generated from molecular-level information, especially large molecular data sets. KEGG combines genomic information with high-level functional information and systematically analyzes the function of genes by computerizing known biological processes within the cell and standardizing existing gene function interpretations [10, 11]. Pathway analysis was used to identify the pathway entries in which DEGs were enriched and to determine the related changes in cellular pathways. Statistical significance was set at $P < 0.05$.

Signal network and co-expression network analysis

Biological networks reflect the interrelationship between genes or genes and other functions or pathways. Signal network analysis conducts interaction analysis through relational or predictive relationships in the GCBI database and identifies important nodes from the network diagram parameters. A co-expression network was constructed based on gene expression similarity. This network displayed the similarities between genes in a clear and hierarchical manner and helped identify key regulatory genes and interactions.

Survival analysis

Gene expression profiling interactive analysis (GEPIA) (http://gepia.cancer-pku.cn/) is an online analytical tool based on TCGA database. This included various modules, including differential expression analysis, profile mapping, correlation analysis, and patient survival analysis. Survival analysis was used to compare different genes based on gene expression levels. A survival curve was plotted based on the survival time of patients with respect to different gene expression levels, to determine the relationship between the gene and patient survival time.

Results

Differential gene expression between gastric cancer and normal tissues

A total of 983 DEGs were identified between normal and gastric cancer samples according to the following criteria: $P < 0.001$ and FC > 2. This included 547 upregulated and 436 downregulated genes (Fig. 1).

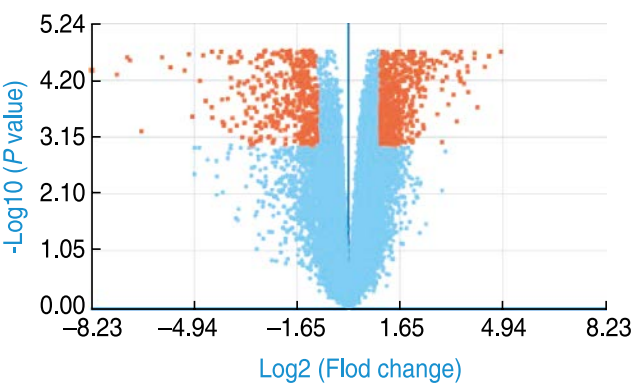

Fig. 1 Volcano plot of DEGs

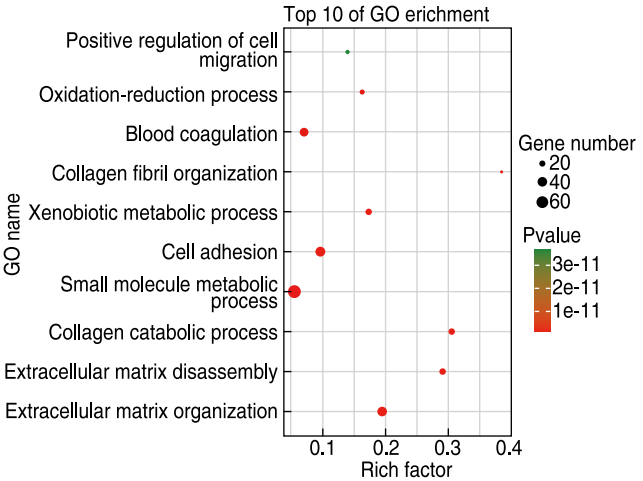

Fig. 2 The top 10 most enriched GO analysis categories

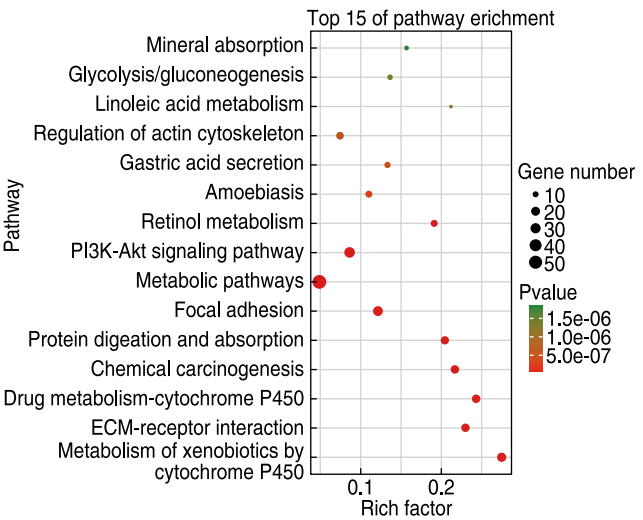

Fig. 3 The top 15 most enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways

Function and pathway enrichment analysis

Gastric cancer gene expression was investigated at a functional level. A total of 983 DEGs (FDR < 0.05) were classified into 232 GO terms. Based on the decreasing FDR values the top categories were (Fig. 2) extracellular matrix organization, extracellular matrix disassembly, collagen catabolic process, small molecule metabolic process, cell adhesion, xenobiotic metabolic process, collagen fibril organization, blood coagulation, oxidation-reduction process, and positive regulation of cell migration. Pathway enrichment analysis was used to identify significant biological pathways related to the DEGs. A total of 50 enriched pathways were identified with an FDR < 0.05. The results revealed that biological processes such as metabolism of xenobiotics by cytochrome P450s (FDR = 2.5×10^{-19}), ECM-receptor interaction (FDR = 4.65

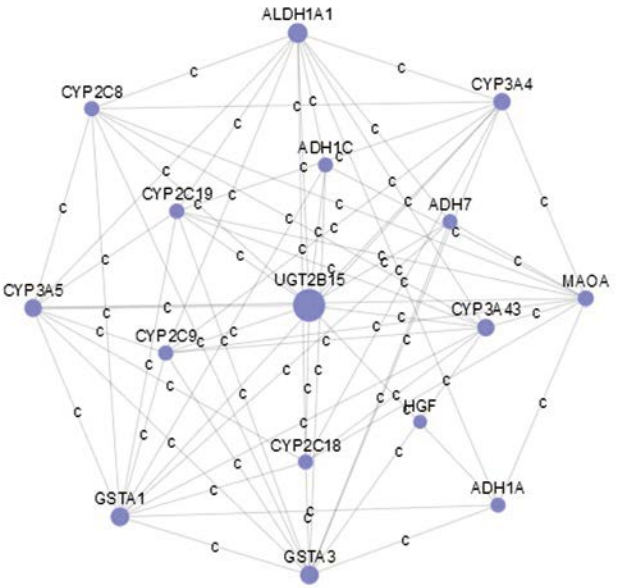

Fig. 4 Signal network of DEGs

$\times 10^{-18}$), drug metabolism by cytochrome P450s (FDR = 7.81×10^{-17}), metabolic pathways (FDR = 3.2×10^{-13}), and the PI3K-Akt signaling pathway (FDR = 9.75×10^{-13}) were significantly enriched (Fig. 3).

Construction of signal and co-expression networks

Genes interact with and regulate one another. Networks visualize the relationships of DEGs to determine the upstream and downstream regulatory relationships between genes, which enables the identification of core genes. The signal and co-expression networks of DEGs are shown in Fig. 4 and Fig. 5, respectively. The cytochrome P450 family genes (such as *CYP2C8*, *CYP2C9*, *CYP2C18*, *CYP2C19*, *CYP3A4*, *CYP3A5*, and *CYP3A43*) and several members of the alcohol dehydrogenase family (such as *ADH1A*, *ADH1C*, and *ADH7*), are related to gastric cancer (Fig. 4). Furthermore, the UDP glucuronosyl transferase 2 family polypeptide B15 (*UGT2B15*) was most closely linked with other genes in the network. The relationships between the top 49 significant DEGs, including 39 upregulated and 10 downregulated genes, is shown in Fig. 5.

Survival analysis

Survival analyses were performed to determine the association between gene expression levels and patient survival time. Patients with gastric cancer and high expression of genes such as *HGF*, *NID2*, *FSTL1*, and *INHBA* had shorter survival times than those with lower expression of these genes (Fig. 6–9). The *P* value of each of the four genes was less than 0.05 (*P* = 0.0041, 0.0013, 0.023, and 0.028, respectively).

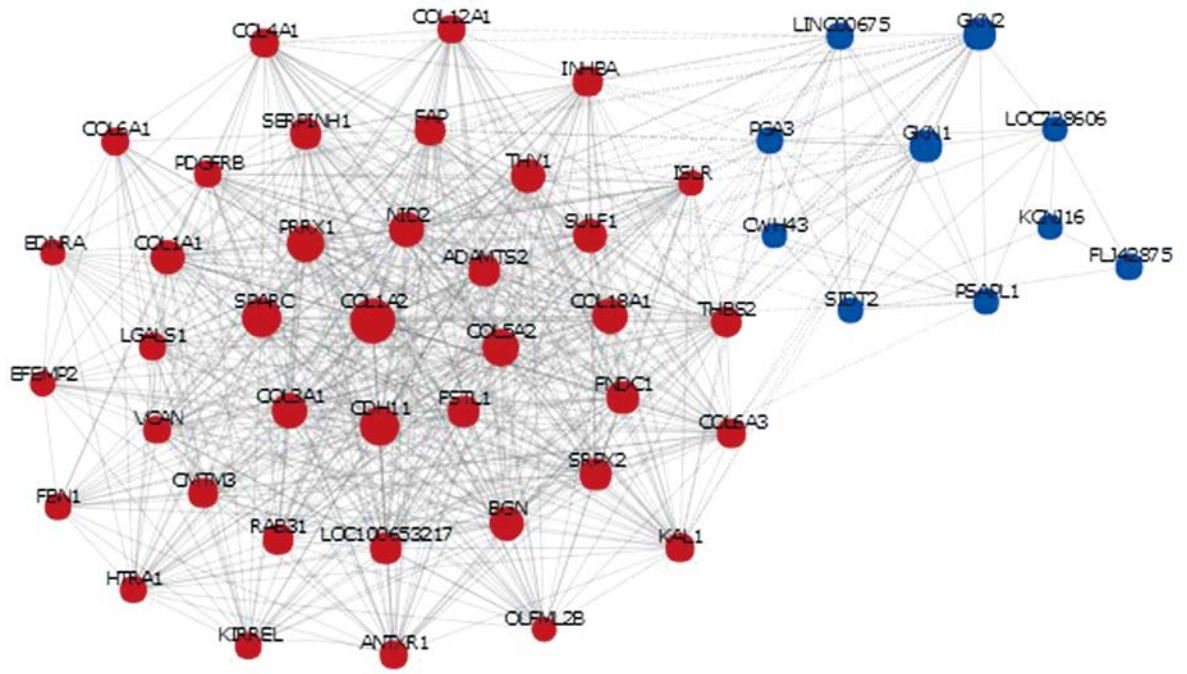

Fig. 5 Co-expression network of DEGs. The red nodes represent upregulated genes, and the blue nodes represent downregulated genes

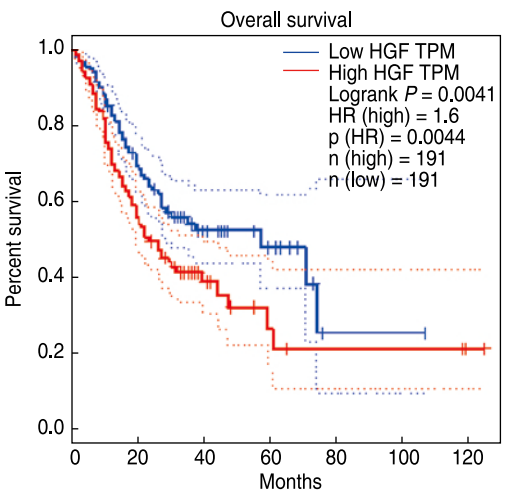

Fig. 6 Survival curve of HGF

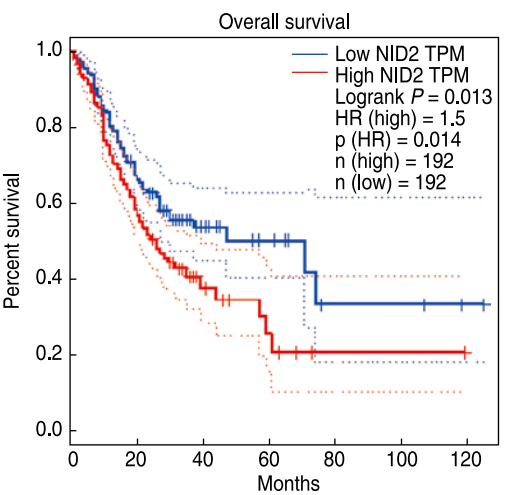

Fig. 7 Survival curve of NID2

Discussion

Gastric cancer is the fourth most common malignancy and the third largest cause of cancer-related deaths worldwide. Although radical surgery combined with systemic chemotherapy has improved the survival rate in patients with gastric cancer, most patients are diagnosed at an advanced stage. Radical resection has a high recurrence rate and results in poor response to treatment [12, 13]. To improve the survival rate in patients with gastric cancer, medical scientists have been committed to understanding gastric cancer based on molecular mechanisms and seeking new therapeutic targets [14]. Genomics, epigenetics, and proteomics have been used to elucidate the molecular mechanisms underlying gastric cancer and identify biomarkers that are associated with poor prognosis and possess curative effects [15, 16]. These biomarkers may serve as therapeutic targets in patients with advanced gastric cancer. Therefore, based on bioinformatics analyses, core DEGs were identified and gastric cancer was better understood at the genetic level, which may provide a new direction for the future

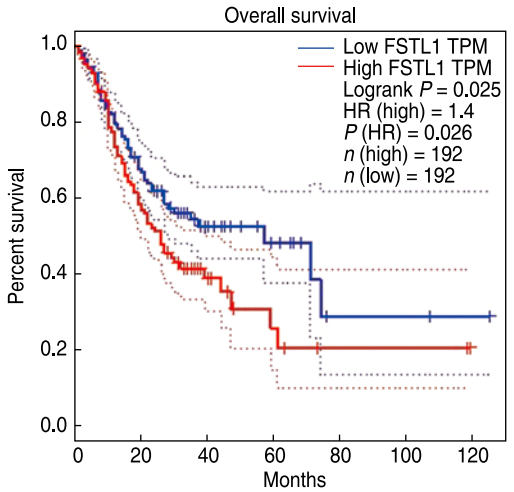

Fig. 8 Survival curve of FSTL1

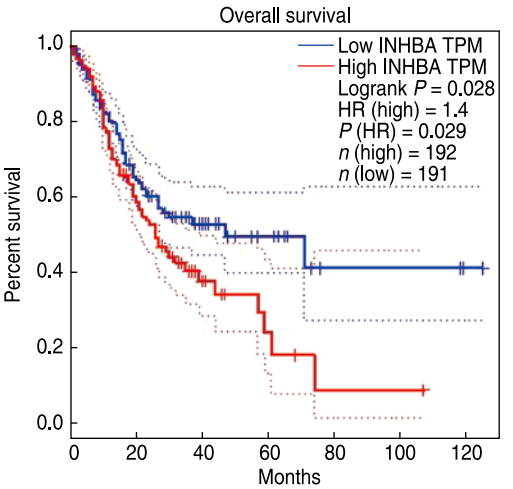

Fig. 9 Survival curve of INGBA

treatment of gastric cancer.

In this study, 983 DEGs were screened by comparing tissue samples from patients with gastric cancer and healthy individuals. Among them, 547 genes were upregulated and 436 were downregulated. The most significant genes were primarily involved in signal transduction, cell proliferation, apoptosis, biosynthesis, gene expression, hormone secretion, and biological metabolism. GO enrichment analysis classifies genes according to their functions and identifies significantly different biological functions by counting the degree of enrichment of genes involved with different functions. The most significant functions were in processes such as cell metabolism, substance metabolism, cell adhesion, coagulation reactions, redox processes, cell migration, and angiogenesis. Additionally, 76, 44, and 20 genes were enriched in the metabolic pathway of small molecules, cell adhesion, and angiogenesis, respectively. Recent evidence has indicated that neovascularization is a necessary condition for tumor growth and metastasis [17]. Cell adhesion, molecular metabolism, and cell migration are all associated with tumor development and progression [18]. Pathway enrichment analysis and pathway relation network construction of DEGs indicated biological processes such as pathways in cancer, adherens junction, focal adhesion, the WNT signaling pathway, and the pentose phosphate pathway to be significant.

The signal networks suggest that *UGT2B15*, *ALDH1A1*, and cytochrome P450 family genes (such as *CYP2C8*, *CYP2C9*, *CYP2C18*, *CYP2C19*, *CYP3A4*, *CYP3A5*, and *CYP3A43*), and hepatocyte growth factor (*HGF*) are closely related to gastric cancer. *HGF* is a ubiquitous cytokine that is involved in multiple biological processes. *HGF* stimulates tumor cells by activating the homologous receptor c-Met [19, 20]. *HGF/c-MET* signaling strongly participates in angiogenesis, proliferation, progression, and metastasis in many human cancers, including head and neck, non-small cell lung carcinoma, pancreatic, cervical, and prostate cancers [21–25]. *HGF/Met* expression levels are inversely correlated with survival time in patients with late-stage nasopharyngeal carcinoma [26]. Recently, research has shown that *HGF* expression was enhanced in gastric cancer tissues, and the overexpression of *HGF* stimulated an increase in vascular endothelial growth factor (VEGF) expression, which could significantly promote proliferation and increase cell migration [27]. It has been reported that cells overexpressing *HGF* are less sensitive to the suppression of biological processes such as cell growth, cell migration, cell invasion, colony formation, and cell cycle regulation by the *C-Met* inhibitor [28]. Therefore, *HGF*, which acts as a cancer promoter, plays a vital role in gastric cancer and may be a promising therapeutic target [27]. After co-expression analysis of the DEGs, we found that genes, including *COL1A2*, *CDH11*, *SPARC*, *COL18A1*, *COL3A1*, *NID2*, *BGN*, *THY1*, and *FSTL1* were closely related to the rest of the genes. These genes are primarily involved in various signaling pathways, such as angiogenesis, cell adhesion, and cell proliferation, and play important roles in tumorigenesis and metastasis.

Potential genes, such as *HGF*, *NID2*, *FSTL1*, and *INHBA*, were further confirmed to be related to gastric cancer by drawing their survival curves. The results indicated that the expression levels of these genes were negatively correlated with the overall survival in patients with gastric cancer. This correlation may be due to these genes playing an substantial role in promoting metastasis and the development of gastric cancer. Therefore, these genes may be potential targets for the treatment of gastric cancer, and interventions to inhibit their expression may help improve the survival rate in patients with gastric cancer.

Conclusions

Bioinformatics analyses have been increasingly applied in research on various clinical diseases. In this study, we identified the core DEGs, including *HGF*, *NID2*, *FSTL1*, and *INHBA*, which may be involved in proliferation, progression, metastasis of gastric cancer and survivial of patients with gastric cancer.. This finding verifies the reliability of DEGs and provides a theoretical basis for further related research. These genes may provide new insights into the diagnosis, treatment, prognosis, and prevention of gastric cancer. However, as this study was based on minimal bioinformatics analyses of gene chip data, further experiments are needed to validate the potential of these genes.

Acknowledgments

Not applicable.

Funding

Supported by a grant from the Project of Southwest Medical University-Three Affiliated Hospitals (No. 2017-ZRQN-028).

Conflicts of interest

The authors indicated no potential conflicts of interest.

Author contributions

Not applicable.

Data availability statement

Not applicable.

Ethical approval

Not applicable.

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(6):394-424.
2. Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015. CA Cancer J Clin. 2016;66(2):115-132.
3. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. 2019;69(1):7-34.
4. Cunningham D, Allum WH, Stenning SP, et al. Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. N Engl J Med. 2006;355(1):11-20.
5. Ychou M, Boige V, Pignon JP, et al. Perioperative chemotherapy compared with surgery alone for resectable gastroesophageal adenocarcinoma: an FNCLCC and FFCD multicenter phase III trial. J Clin Oncol. 2011;29(13):1715-1721.
6. Hamashima C, Shabana M, Okada K, et al. Mortality reduction from gastric cancer by endoscopic and radiographic screening. Cancer Sci. 2015;106(12):1744-1749.
7. He J, Jin Y, Chen Y, et al. Downregulation of ALDOB is associated with poor prognosis of patients with gastric cancer. Oncol Targets Ther. 2016;9:6099-6109.
8. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25(1):25-29.
9. Hulsegge I, Kommadath A, Smits MA. Globaltest and GOEAST: two different approaches for Gene Ontology analysis. Bmc Proceedings. 2009;3(S4):S10.
10. Kanehisa M, Araki M, Goto S, et al. KEGG for linking genomes to life and the environment. Nucleic Acids Res. 2008;36(Database issue): D480-484.
11. Kanehisa M, Furumichi M, Tanabe M, et al. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 2017;5(D1):D353-D361.
12. Fujii M, Kochi M, Takayama T. Recent advances in chemotherapy for advanced gastric cancer in Japan. Surg Today. 2010;40(4):295-300.
13. Kang H, Kauh JS. Chemotherapy in the treatment of metastatic gastric cancer: is there a global standard? Curr Treat Options Oncol. 2011;12(1):96-106.
14. Bang YJ, Kim YW, Yang HK, et al. CLASSIC trial investigators. Adjuvant capecitabine and oxaliplatin for gastric cancer after D2 gastrectomy (CLASSIC): a phase 3 open-label, randomised controlled trial. Lancet. 2012;379(9813):315-321.
15. Lin LL, Huang HC, Juan HF. Discovery of biomarkers for gastric cancer: a proteomics approach. J Proteomics. 2012;75(11):3081.
16. Yamashita K, Sakuramoto S, Watanabe M. Genomic and epigenetic profiles of gastric cancer: Potential diagnostic and therapeutic applications. Surg Today. 2011;41(1):24.
17. Rak J, Mitsuhashi Y, Bayko L, et al. Mutant ras oncogenes upregulate VEGF/VPF expression: implications for induction and inhibition of tumor angiogenesis. Cancer Res. 1995;55(20):4575-4580.
18. Berger H, Marques MS, Zietlow R, et al. Gastric cancer pathogenesis. Helicobacter. 2016;21 Suppl 1:34-38.
19. Owusu BY, Galemmo R, Janetka J, et al. Hepatocyte growth factor, a key tumor-promoting factor in the tumor microenvironment. Cancers (Basel). 2017;9(4):35.
20. Spina A, De PV, Cerulo G, et al. HGF/c-MET axis in tumor microenvironment and metastasis formation. Biomedicines. 2015;3(1):71-88.
21. Arnold L, Enders J, Thomas SM. Activated HGF-c-met axis in head and neck cancer. Cancers (Basel). 2017;9(12).
22. Boromand N, Hasanzadeh M, Shahidsales S, et al. Clinical and prognostic value of the C-Met/HGF signaling pathway in cervical cancer. J Cell Physiol. 2018;233(6):4490-4496.
23. Martin TA, Mason MD, Jiang WG. HGF and the regulation of tight junctions in human prostate cancer cells. Oncol Rep. 2014;32(1): 213-224.
24. Pothula SP, Xu Z, Goldstein D, et al. Targeting the HGF/c-MET pathway: stromal remodelling in pancreatic cancer. Oncotarget, 2017;8(44):76722-76739.
25. Tsuta K, Kozu Y, Mimae T, et al. C-MET/Phospho-MET protein expression and MET gene copy number in non-small cell lung carcinomas. J Thorac Oncol. 2012 Feb;7(2):331-339.
26. Qian CN, Guo X, Cao B, et al. Met protein expression level correlates with survival in patients with late-stage nasopharyngeal carcinoma. Cancer Res. 2002;62(2):589.
27. Si Y, Zhang H, Ning T, et al. miR-26a/b inhibit tumor growth and angiogenesis by targeting the HGF-VEGF Axis in gastric carcinoma. Cell Physiol Biochem. 2017;42(4):1670-1683.
28. Ahn SY, Kim J, Kim MA, et al. Increased HGF expression induces resistance to c-MET Tyrosine Kinase inhibitors in gastric cancer.

Anticancer Res. 2017;37(3):1127-1138.

DOI 10.1007/s10330-021-0535-5
Cite this article as: Feng LB, Yu L, Chen XL. Bioinformatics analysis of potential hub genes associated with biological characteristics and survival in patients with gastric cancer. Oncol Transl Med. 2021;8(5):232–238.