Observational Study

Medicine

OPEN

MYC and NCAPG2 as molecular targets of colorectal cancer and gastric cancer in nursing

Xihua Mi, MD[a], Haifeng Shan, MD[a], Chunbo Kang, MD[a,*], Jie Zhang, MD[a], Shiyang Hou, MD[a], Yanfang Gao, MD[a], Lingli Hao, MD[a], Xiaoli Gao, MD[a], Qiang Gao, MD[a], Xiaoqian Chi, MD[a], Qijun Zhang, MD[a]

Abstract

Colorectal cancer is a common malignant tumor in intestinal tract, the early symptoms are not obvious. Gastric cancer is a malignant tumor originating from the gastric mucosal epithelium. However, the role of MYC and non-SMC condensin II complex subunit G2 (NCAPG2) in colorectal cancer and gastric cancer remains unclear. The colorectal cancer datasets GSE49355 and gastric cancer datasets GSE19826 were downloaded from gene expression omnibus database. Differentially expressed genes (DEGs) were screened and weighted gene co-expression network analysis (WGCNA) was performed. Functional enrichment analysis, gene set enrichment analysis (GSEA) and immune infiltration analysis was performed. Construction and analysis of protein-protein interactions (PPI) network. Survival analysis and comparative toxicogenomics database (CTD) were performed. A heat map of gene expression was drawn. A total of 751 DEGs were obtained. According to the gene ontology (GO) analysis, in Biological process (BP) analysis, they are mainly enriched in cell differentiation, cartilage development, and skeletal development. In cellular component (CC) analysis, they are mainly enriched in the cytoskeleton of muscle cells and actin filaments. In molecular function (MF) analysis, they are mainly concentrated in Rho GTPase binding, DNA binding, and fibronectin binding. In Kyoto encyclopedia of genes and genomes (KEGG) pathway analysis, they are mainly enriched in the MAPK signaling pathway, apoptosis, and cancer pathways. The soft threshold power for WGCNA analysis was set to 9, resulting in the generation of 40 modules. Ultimately, 2 core genes (MYC and NCAPG2) were identified. The heatmap of core gene expression showed high expression of MYC and NCAPG2 in colorectal cancer tissue samples and low expression in normal tissue samples, while they were core molecules in gastric cancer. Survival analysis indicated that MYC and NCAPG2 were risk factors, showing an upregulation trend with increasing risk scores. CTD analysis revealed associations of MYC and NCAPG2 with colorectal cancer, gastric cancer, inflammation, and immune system diseases. MYC and NCAPG2 are highly expressed in colorectal cancer. The higher the expression of MYC and NCAPG2, the worse the prognosis. MYC and NCAPG2 are core molecules in gastric cancer.

Abbreviations: CTD = comparative toxicogenomics database, DEGs = differentially expressed genes, FC = fold change, FDR = false discovery rate, GO = gene ontology, GSEA = gene set enrichment analysis, KEGG = Kyoto encyclopedia of genes and genomes, NCAPG2 = non-SMC condensin II complex subunit G2, PPI = protein-protein interactions, STRING = search tool for the retrieval of interacting genes, WGCNA = weighted gene co-expression network analysis.

Keywords: colorectal cancer, gastric cancer, molecular targets, MYC, NCAPG2, nursing

1. Introduction

Colorectal cancer and gastric cancer are 2 common malignant tumors of the digestive system, which differ in terms of pathogenesis, epidemiology, pathological characteristics, harmfulness, and nursing requirements.[1] Colorectal cancer originates within the colon and rectum, typically arising from the cells of the colon or rectal mucosa.[2] Gastric cancer refers to the malignancy formed by the malignant transformation of epithelial cells on the gastric mucosa.[3] Colorectal cancer is common among middle-aged and elderly populations, with the incidence increasing with age, particularly in developed countries.[4] Gastric cancer is more prevalent worldwide, especially in Asian countries such as Japan, China, and Korea.[5] Untreated colorectal cancer may lead to serious consequences such as gastrointestinal bleeding, obstruction, distant metastasis, and even life-threatening conditions.[6] Gastric cancer often presents no symptoms in its early stages, but as it progresses to advanced stages, symptoms like dyspepsia, abdominal pain, and weight loss may occur, and in severe cases, it can lead to cachexia and

XM and HS contributed equally to this work.

The authors have no funding and conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

The data for this paper were obtained exclusively from public databases, thus exempting it from ethical review.

[a] *Gastrointestinal Rehabilitation Center, Beijing Rehabilitation Hospital Affiliated to Capital Medical University, Beijing, China.*

* *Correspondence: Chunbo Kang, Gastrointestinal Rehabilitation Center, Beijing Rehabilitation Hospital Affiliated to Capital Medical University, Xixiazhuang, Badachu, Shijingshan District, Beijing 100144, China (e-mail: kangchunbo220@163.com)*

How to cite this article: Mi X, Shan H, Kang C, Zhang J, Hou S, Gao Y, Hao L, Gao X, Gao Q, Chi X, Zhang Q. MYC and NCAPG2 as molecular targets of colorectal cancer and gastric cancer in nursing. Medicine 2024;103:18(e38029).

Received: 28 August 2023 / Received in final form: 17 March 2024 / Accepted: 5 April 2024

http://dx.doi.org/10.1097/MD.0000000000038029

distant metastasis.[7] Although the exact causes are unclear, genetic factors, chromosomal abnormalities, gene fusions, and other factors may be associated with its development. Therefore, in-depth research into the molecular mechanisms of colorectal and gastric cancers is particularly important as it can provide insights for personalized treatment strategies. Currently, molecular targeted therapy has become an important treatment strategy in nursing protocols. By targeting specific

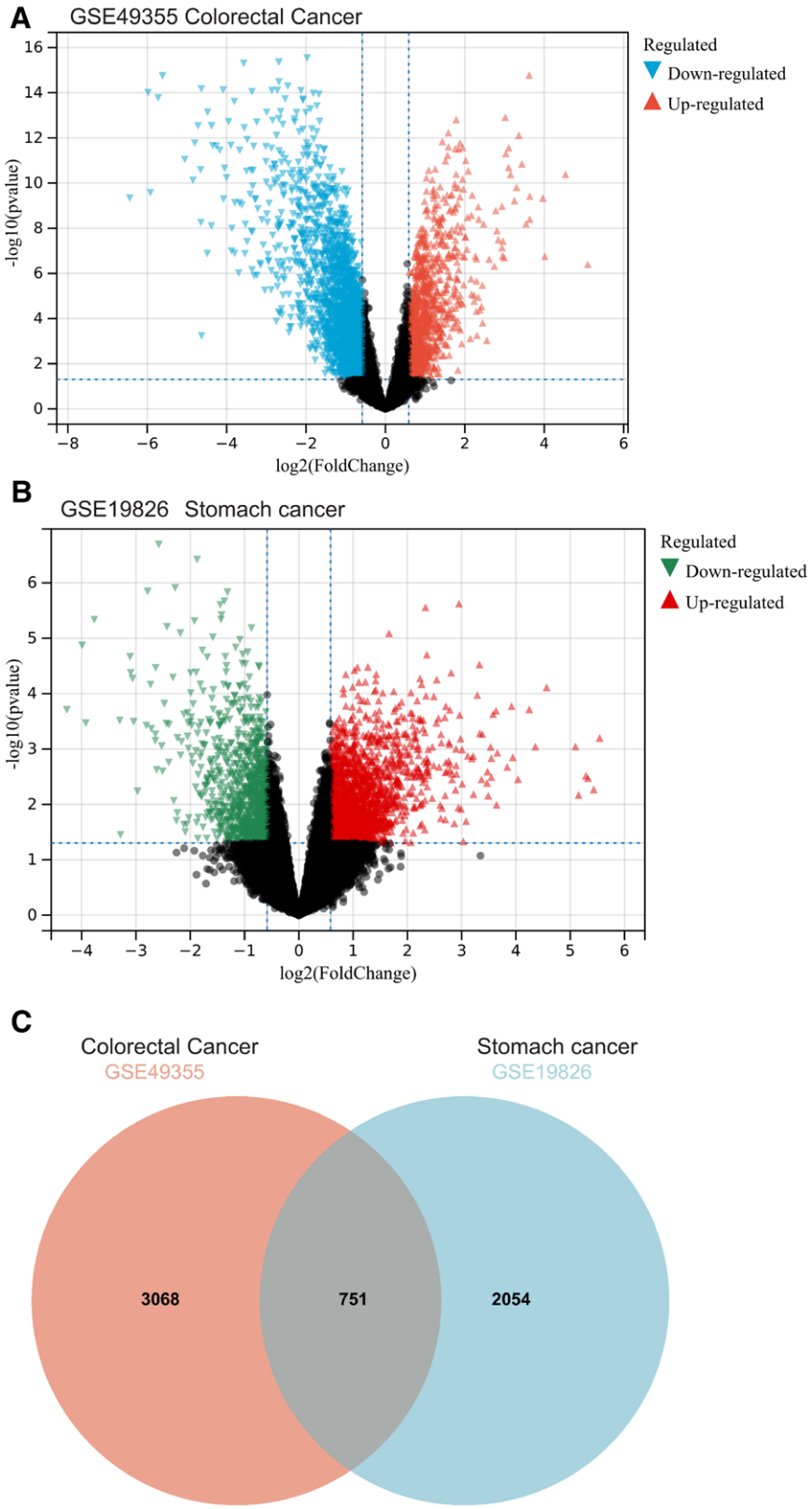

Figure 1. Differential gene analysis. (A) Differential gene of GSE49355. (B) Differential gene of GSE19826. (C) A total of 751 DEGs. DEGs = differentially expressed genes.

molecular targets, tumor growth and spread can be more effectively inhibited, thereby improving treatment outcomes and survival rates.

Biological informatics technology has rapidly developed, providing new perspectives and methods for biological research, with high-throughput sequencing and machine learning at its core.[8] Its efficiency, accuracy, diversity, and comprehensiveness in data processing and analysis make it a powerful tool for deciphering life. In the field of medicine, bioinformatics supports personalized medicine, driving medical progress, and its interdisciplinary applications further promote innovation and collaboration in scientific research.[9,10]

MYC is an important transcription factor that regulates gene expression and is involved in processes such as cell proliferation, differentiation, and apoptosis.[11] Non-SMC condensin II complex subunit G2 (NCAPG2) is a gene encoding a protein involved in cell mitosis, regulating chromosome structure and dynamics.[12] MYC and NCAPG2 play important roles in cell biology and molecular biology research and have certain effects on the occurrence and development of tumors.[13,14] However, the relationship between MYC, NCAPG2 genes, and colorectal/gastric cancer is not yet clear.

Therefore, this study aims to use bioinformatics technology to explore the core genes between colorectal cancer and gastric cancer and normal tissues, and conduct enrichment analysis and pathway analysis. Public datasets will be used to validate the significant roles of MYC and NCAPG2 genes in colorectal and gastric cancers. In terms of nursing for colorectal and gastric cancers, changes in MYC and NCAPG2 indicators will be observed to optimize nursing protocols.

2. Methods

2.1. Colorectal cancer and gastric cancer datasets

In this study, the colorectal cancer dataset GSE49355 and gastric cancer dataset GSE19826 profiles were downloaded from the gene expression omnibus database (http://www.ncbi.nlm.nih.gov/geo/), generated from GPL96 and GPL570 platforms, respectively. GSE49355 comprises 20 colorectal cancer and 18 normal tissue samples, GSE19826 comprises 12 gastric cancer and 12 normal tissue samples, utilized for identifying differentially expressed genes (DEGs) in colorectal and gastric cancers.

2.2. DEGs selection

Log2 transformation was applied to the colorectal cancer dataset GSE49355 and gastric cancer dataset GSE19826. Multivariate linear regression was performed using the lmFit

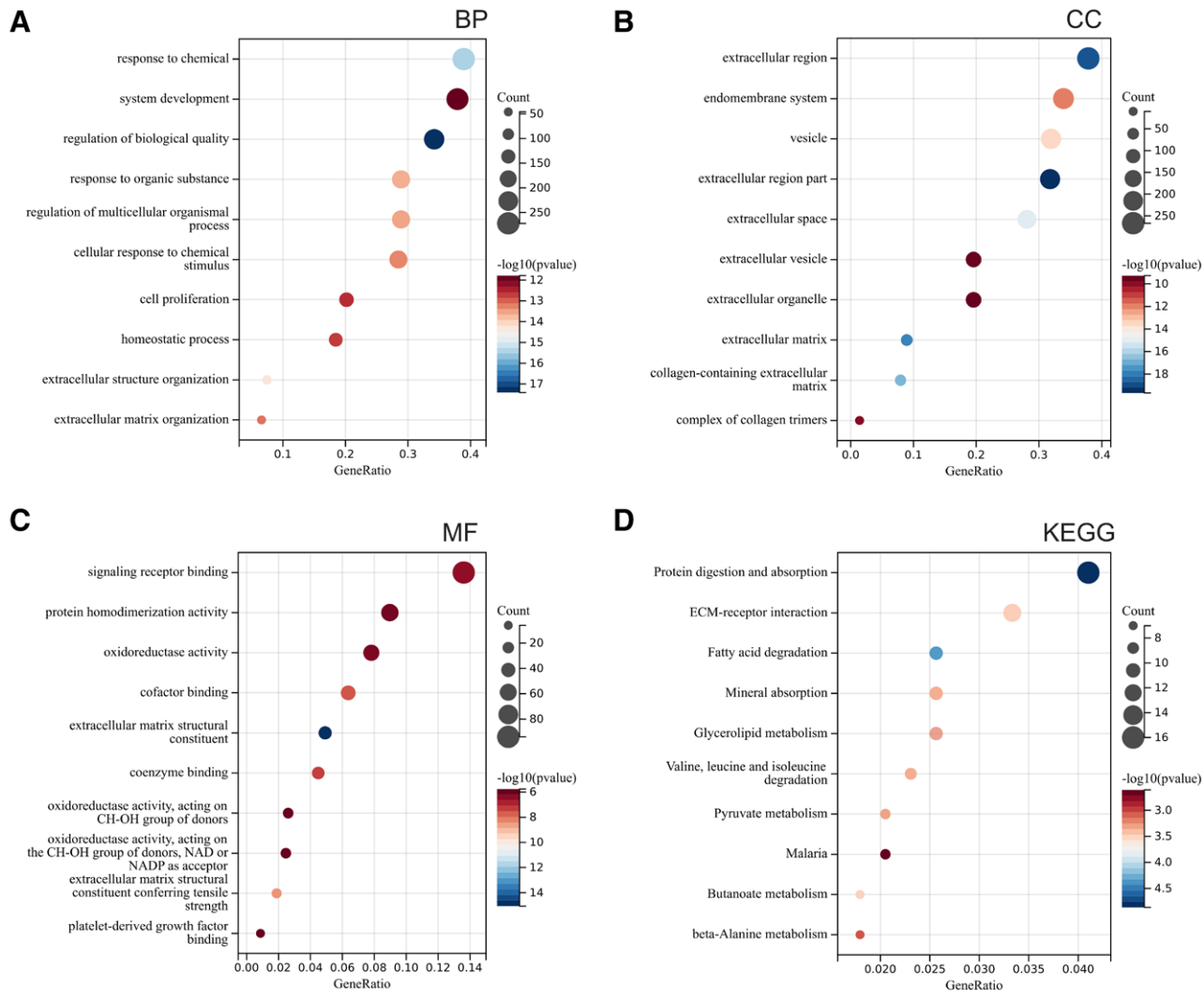

Figure 2. Functional enrichment analysis of DEGs. (A) Biological process analysis. (B) Cellular component analysis. (C) Molecular function analysis. (D) KEGG analysis. KEGG = Kyoto encyclopedia of genes and genomes.

function, adjusting the t-statistic, f-statistic, and log2 fold change (FC) through empirical Bayesian moderation of the standard errors toward a common value. The R package "limma" was employed for probe summarization and background correction of the combined matrix for GSE49355. The Benjamini-Hochberg method was used to adjust the raw *P* values, and FC was calculated using the false discovery rate (FDR). The cutoff criteria for DEGs were set at $P < .05$ and $FC > 1.5$. Significant differences for each gene were obtained, and volcano plots were generated. Subsequently, the intersection of DEGs between the colorectal cancer dataset GSE49355 and gastric cancer dataset GSE19826 was obtained to obtain DEGs.

2.3. Weighted gene co-expression network analysis (WGCNA)

Initially, the median absolute deviation of each gene was calculated from the gene expression profiles, and the bottom 50% of genes with the smallest median absolute deviation were removed. Outliers of genes and samples were removed using the goodSamplesGenes method of the R package WGCNA. Further, a scale-free co-expression network was constructed using WGCNA. Specifically, Pearson correlation matrices and average linkage methods were applied to all pairs of genes. Then, using the power function A_mn = | C_mn | ^β structure weighted adjacency matrix (C_mn = Gene_m and Pearson correlation between Gene_n; A_mn = adjacency between Gene m and Gene n). β is a soft thresholding parameter emphasizing strong correlations and reducing the influence of weak or negative correlations. After selecting a power of 10, the adjacency matrix was transformed into a Topological Overlap Matrix (TOM), measuring the network connectivity of a gene. Modules of genes with similar expression profiles were identified using average linkage hierarchical clustering based on TOM-based dissimilarity measure, with a minimum module size (genes per module) set to 30. The sensitivity was set to 3. For further

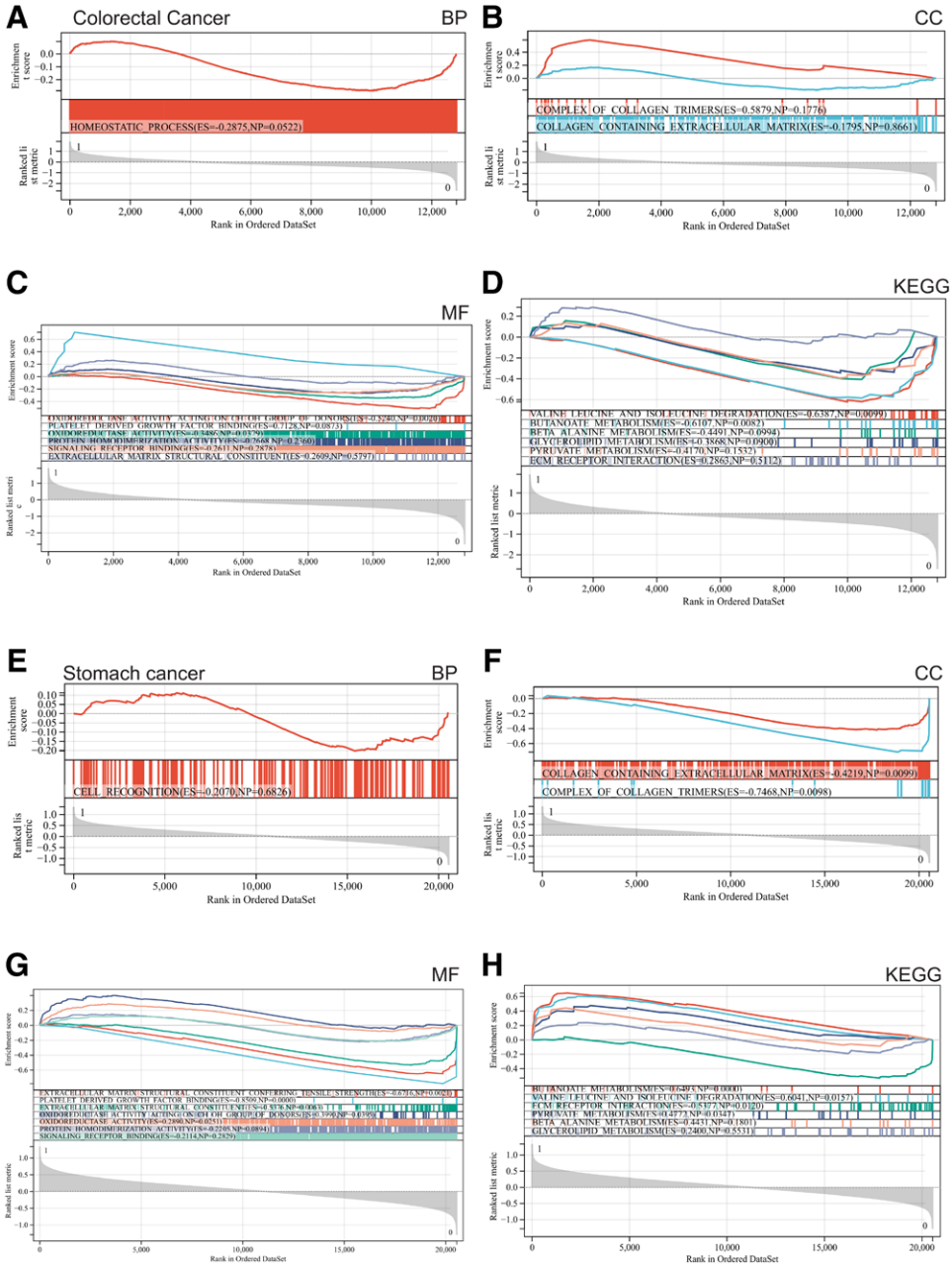

Figure 3. Gene set enrichment analysis. (A–D) GSE49355. (E–H) GSE19826.

analysis of modules, the dissimilarity of module eigengenes was calculated, a cutting line was selected for the module dendrogram, and some modules were merged. Additionally, modules with a distance <0.25 were merged. Notably, the gray module was considered a set of genes that could not be assigned to any module.

2.4. Functional enrichment analysis

Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses are computational methods for evaluating gene function and biological pathways. In this study, the lists of DEGs filtered by Venn diagrams were input into the KEGG API (https://www.kegg.jp/kegg/rest/keggapi.html) to obtain the latest gene annotations for KEGG Pathways. These annotations were used as a background to map genes, and the R package clusterProfiler (version 3.14.3) was utilized for enrichment analysis to obtain results of gene set enrichment. GO annotations for genes were obtained from the R package org.Hs.e.g..db (version 3.1.0) and used as a background to map genes. A minimum gene set size of 5, maximum gene set size of 5000, a P value of < .05, and an FDR of < 0.25 were considered statistically significant metrics.

Additionally, the Metascape database was utilized to provide comprehensive gene list annotations and analysis resources, with visualization export capabilities. We performed functional enrichment analysis on the aforementioned lists of DEGs using the Metascape database (http://metascape.org/gp/index.html).

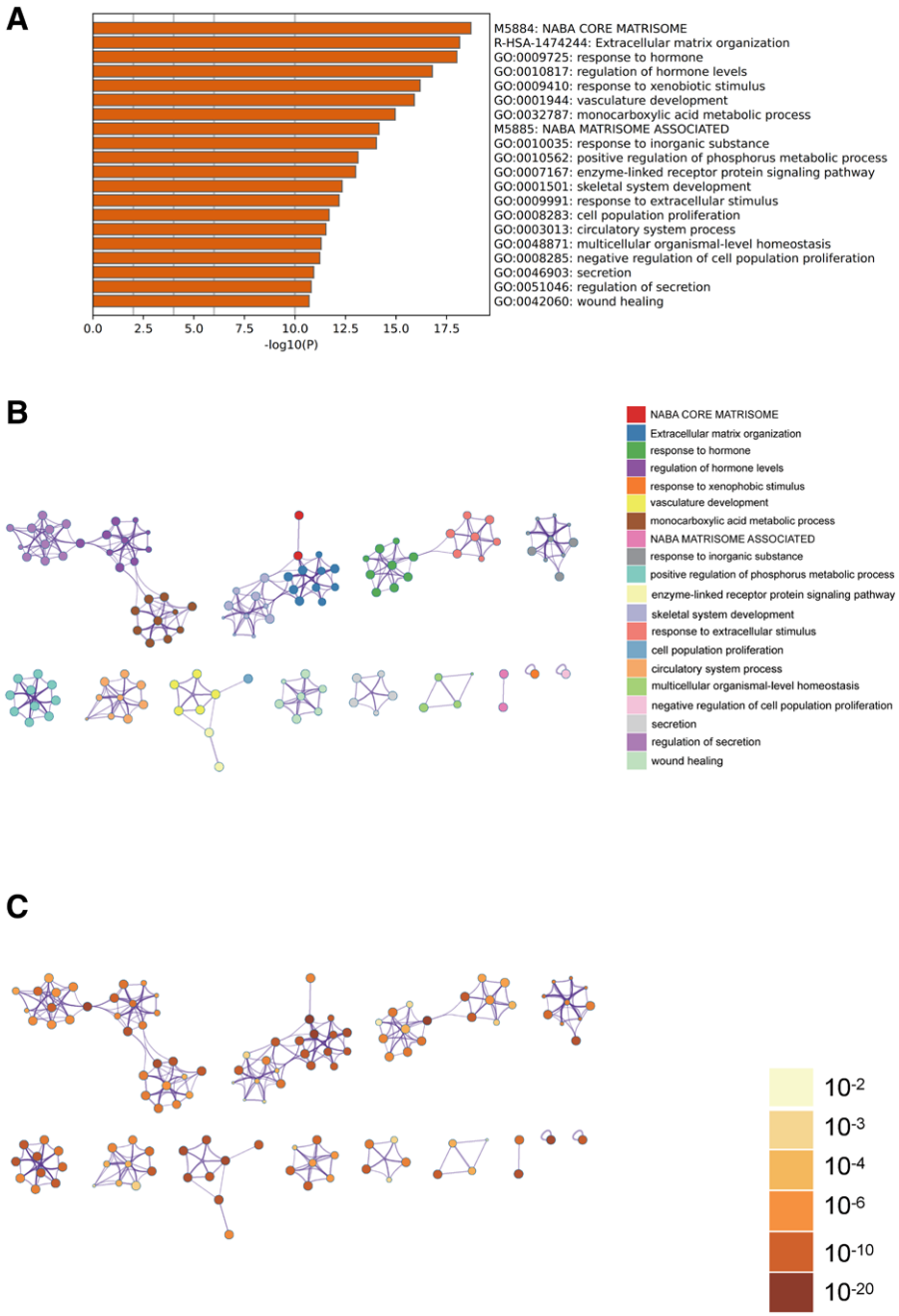

Figure 4. Metascape enrichment analysis. (A) Bar graph of enriched terms across input gene lists, colored by P values. (B) Network of enriched terms: colored by cluster-ID, where nodes that share the same cluster-ID are typically close to each other. (C) Colored by P value, where terms containing more genes tend to have a more significant P value.

2.5. Gene set enrichment analysis (GSEA)

For GSEA, the GSEA software (version 3.0) was obtained from the GSEA website (DOI: 10.1073/pnas.0506580102, http://software.broadinstitute.org/gsea/index.jsp). Samples from both disease and normal tissues were divided into 2 groups, and the Molecular Signatures Database (DOI: 10.1093/bioinformatics/btr260, http://www.gsea-msigdb.org/gsea/downloads.jsp) was used to download the c2.cp.kegg.v7.4.symbols.gmt subset for evaluating relevant pathways and molecular mechanisms. Based on gene expression profiles and phenotype grouping, a minimum gene set size of 5, maximum gene set size of 5000, 1000 permutations, a P value of < .05, and an FDR of < 0.25 were considered statistically significant. Additionally, GO and KEGG analyses were performed on the entire genome, as formulated by GSEA.

2.6. Immune infiltration analysis

CIBERSORT (http://CIBERSORT.stanford.edu/) is a widely used method for calculating immune cell infiltration, with the LM22 gene file defining 22 immune cell subtypes. We applied integrated bioinformatics methods using the CIBERSORT package to analyze gene expression matrices of the colorectal cancer dataset GSE49355 and gastric cancer dataset GSE19826. Linear support vector regression was used to

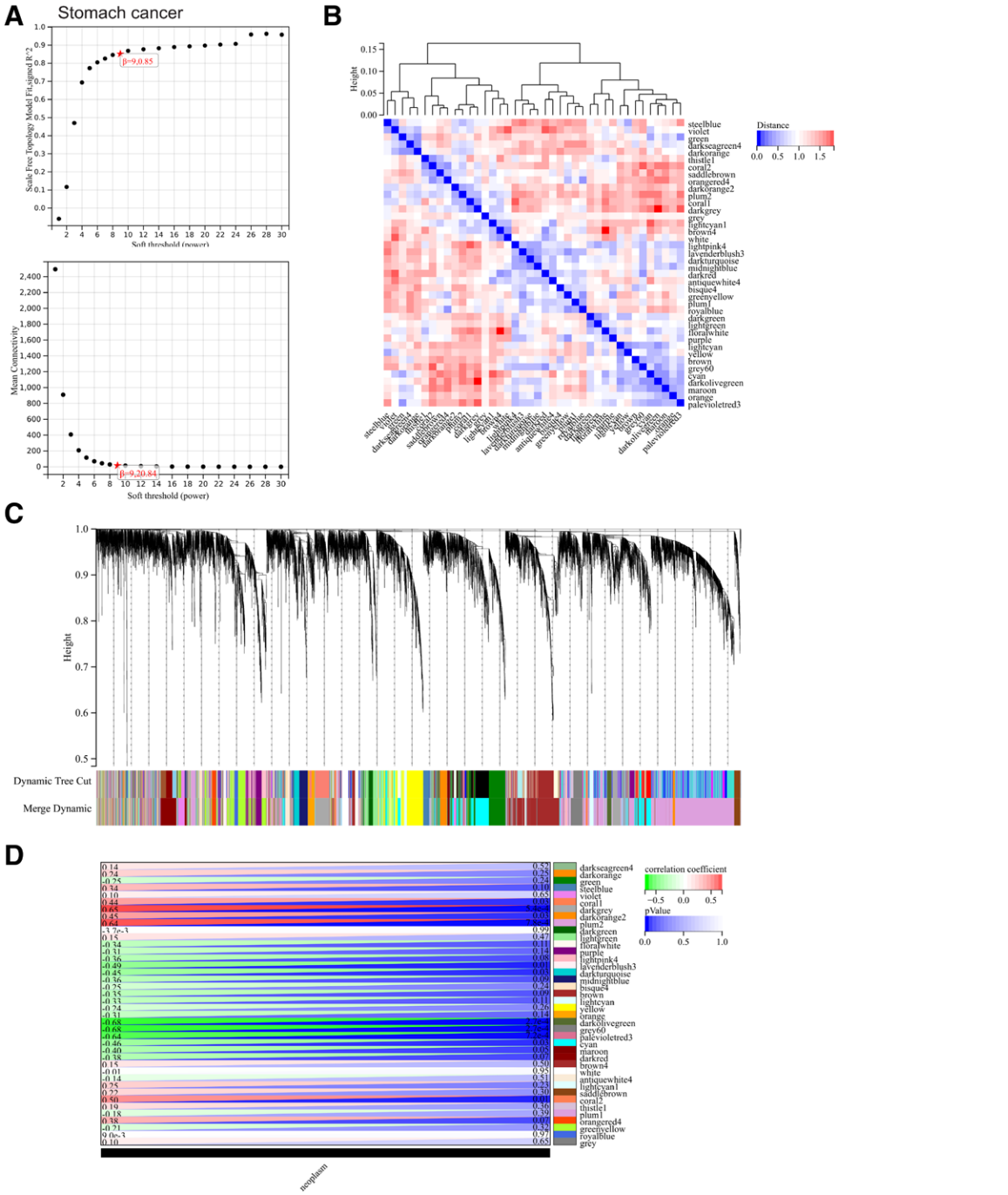

Figure 5. WGCNA. (A) β = 9,0.85. β = 9, 20.84. (B, C) The hierarchical clustering tree of all genes was constructed, and 40 important modules were generated. (D) The heat map of correlation between modules and phenotypes. WGCNA = weighted gene co-expression network analysis.

deconvolute the expression matrix of immune cell subtypes, estimating immune cell abundance. Samples with a confidence $P < .05$ were selected as the cutoff criterion for sufficient confidence.

2.7. Construction and analysis of protein-protein interaction (PPI) networks

The search tool for the retrieval of interacting genes (STRING) database (http://string-db.org/) collects, scores, and integrates all publicly available protein-protein interaction information sources and supplements them with predicted interactions. In this study, the list of DEGs was input into the STRING database to construct a predicted core gene PPI network (confidence > 0.4). Cytoscape software was utilized for biological network analysis and 2-dimensional (2D) visualization. Cytoscape was used to visualize and predict core genes from the PPI network formed by the STRING database. The PPI network was imported into Cytoscape software, and MCODE was used to find the best modules in terms of relevance. Additionally, 3 algorithms (MCC, MNC, DMNC) were used to calculate the top 10 genes of highest relevance, and their intersection was taken. The resulting core gene list was exported after visualization.

2.8. Survival analysis

Clinical survival data for colorectal and gastric cancers were obtained from the cancer genome atlas, and the R package maxstat (version: 0.7–25) was used to calculate the optimal cutoff value of RiskScore for core genes. The minimum sample size for the low and high groups was set >25% and <75%, respectively. Using this cutoff value, patients were divided into high and low groups, and the R package survival survfit function was used to analyze the prognostic differences between the 2 groups. The log-rank test method was employed to assess the significance of prognostic differences between samples in different groups. Additionally, the forest plot was created using the R package forest to observe whether there were significant changes in the prognosis of colorectal and gastric cancers for each independent core gene.

2.9. Gene expression heatmaps

The R package heatmap was used to generate heatmaps of the expression levels of core genes identified from the PPI network in the colorectal cancer dataset GSE49355 and gastric cancer dataset GSE19826, visualizing the expression differences among gastric cancer, colorectal cancer, and normal tissue samples.

2.10. CTD analysis

The comparative toxicogenomics database (CTD) integrates vast amounts of data on interactions between chemicals, genes, phenotypes, and diseases, providing great convenience for research on disease-related environmental exposure factors and potential mechanisms of action of drugs. Core genes were input into the CTD website to find the most relevant diseases associated with core genes, and radar plots of expression differences for each gene were drawn using Excel.

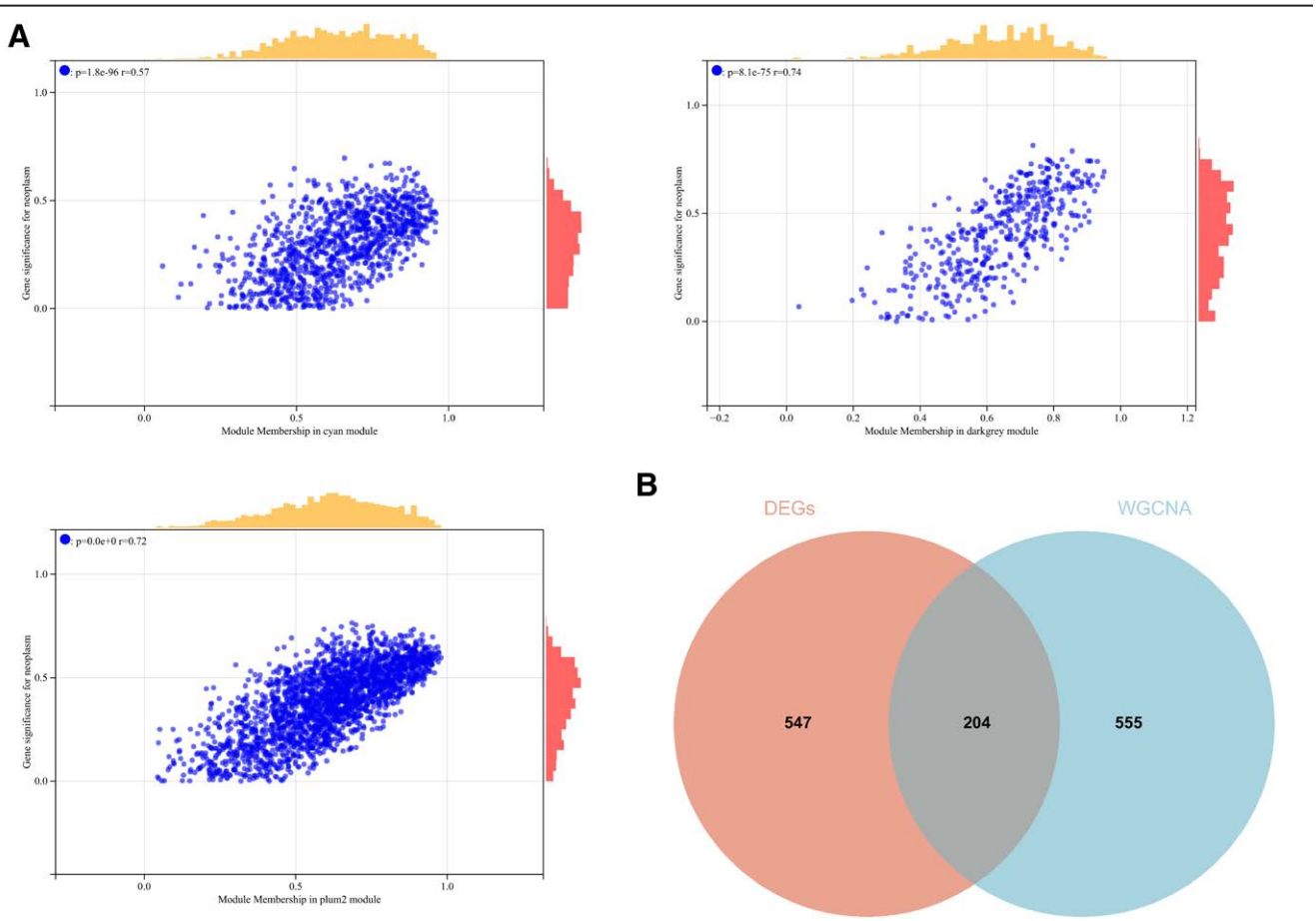

Figure 6. (A) The scatter map of correlation between GS and MM of related hub genes. (B) The DEGs screened by WGCNA and DEGs were used to obtain Venn map. 204 intersection genes were obtained. DEGs = differentially expressed genes, WGCNA = weighted gene co-expression network analysis.

3. Results

3.1. Analysis of DEGs

In this study, differential gene expression was identified from the gene expression matrices of colorectal cancer dataset GSE49355 (Fig. 1A) and gastric cancer dataset GSE19826 (Fig. 1B) using predefined cutoff values. A Venn diagram was utilized to identify the intersecting DEGs between the 2 diseases, resulting in a total of 751 DEGs (Fig. 1C).

3.2. Functional enrichment analysis

3.2.1. DEGs.. We conducted GO and KEGG pathway analyses on these differential expression genes. According to the GO analysis results, in Biological process (BP) analysis, they were mainly enriched in cell differentiation, cartilage development, and skeletal development (Fig. 2A). In cellular component (CC) analysis, they were mainly enriched in the cytoskeleton of muscle cells and actin filaments (Fig. 2B). In molecular function (MF) analysis, they were mainly concentrated in Rho GTPase binding, DNA binding, and fibronectin binding (Fig. 2C). In KEGG pathway analysis, they were mainly enriched in the MAPK signaling pathway, apoptosis, and cancer pathways (Fig. 2D).

3.3. GSEA

Furthermore, we performed GSEA on the entire genome to identify potential enrichments among non-DEGs and to validate the results of DEGs. The intersection of enrichment

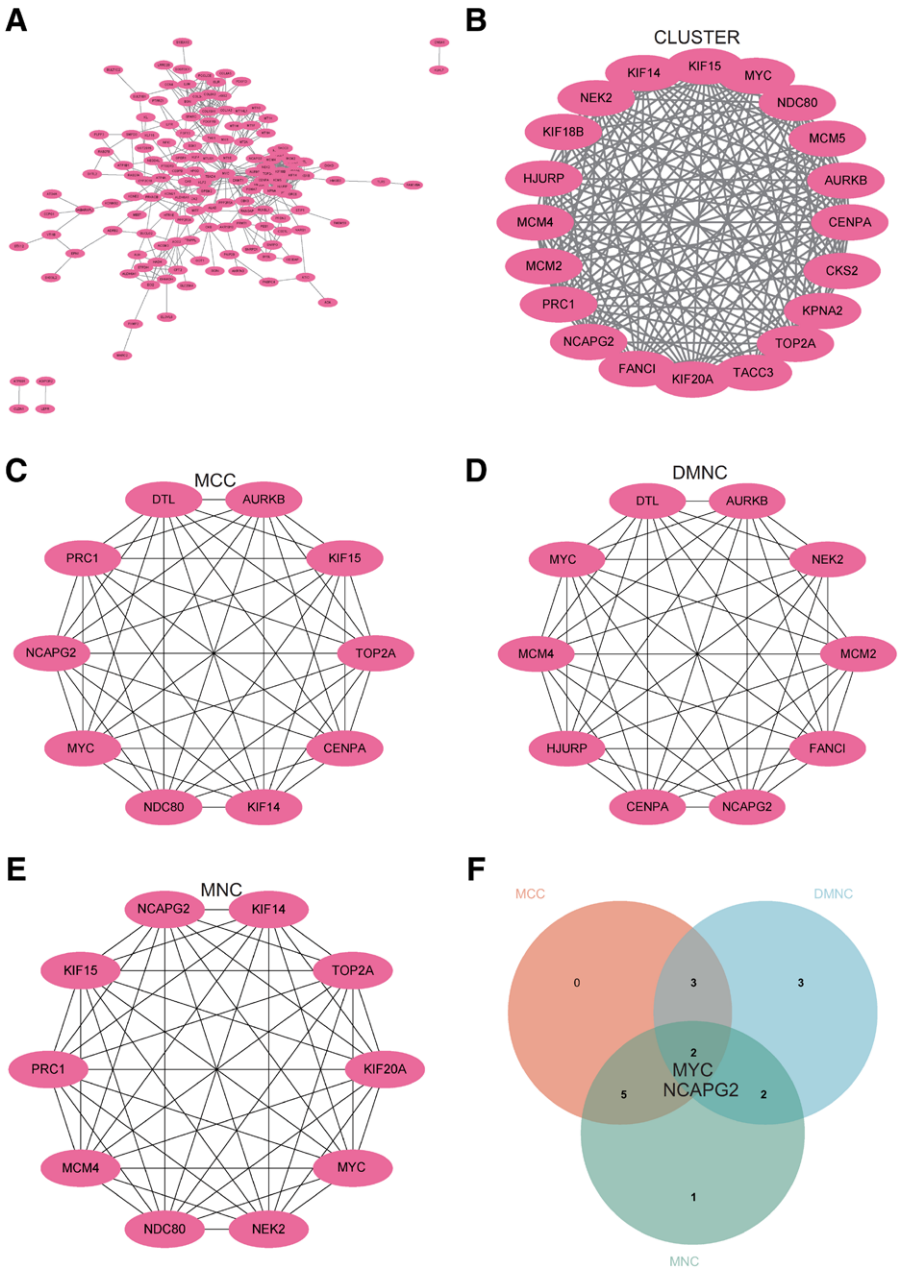

Figure 7. Construction and analysis of protein-protein interaction (PPI) networks. (A) Construct the PPI network of DEGs using STRING online database and utilize Cytoscape software for analysis. (B) CLUSTER was used to identify the central gene. (C) MCC was used to identify the central gene. (D) DMNC was used to identify the central gene. (E) MNC was used to identify the central gene. (F) Two core genes (MYC, NCAPG2) were obtained by merging using Venn diagrams. DEGs = differentially expressed genes, NCAPG2 = non-SMC condensin II complex subunit G2, STRING = search tool for the retrieval of interacting genes.

items with GO and KEGG enrichment items and DEGs is shown in the figures. The enrichment results of both colorectal cancer dataset GSE49355 (Fig. 3A–D) and gastric cancer dataset GSE19826 (Fig. 3E–H) showed that the DEGs were mainly enriched in negative regulation of cell growth factor-stimulated responses, extracellular side of the Plasma membrane, MAPK signaling pathway, and protein-lipid complex binding.

3.4. Metascape enrichment analysis

In Metascape enrichment projects, GO enrichment projects such as skeletal system development were observed (Fig. 4A). Additionally, we also outputted enrichment networks colored and *P* value colored for enrichment items (Fig. 4B and C), visually representing the associations and confidence of each enrichment item.

3.5. WGCNA

The selection of soft threshold power is an important step in WGCNA. We set the soft threshold power to 9 for network topology analysis (Fig. 5A). A hierarchical clustering tree of all genes was constructed, resulting in the generation of 40 modules (Fig. 5B), and the interaction between important modules was analyzed (Fig. 5C). Furthermore, a heatmap of module-trait correlations (Fig. 5D) and scatter plots of correlations between module eigengenes and gene significance were generated (Fig. 6A). We identified 3 highly interconnected genes as hub genes in clinically significant modules based on |MM| > 0.8

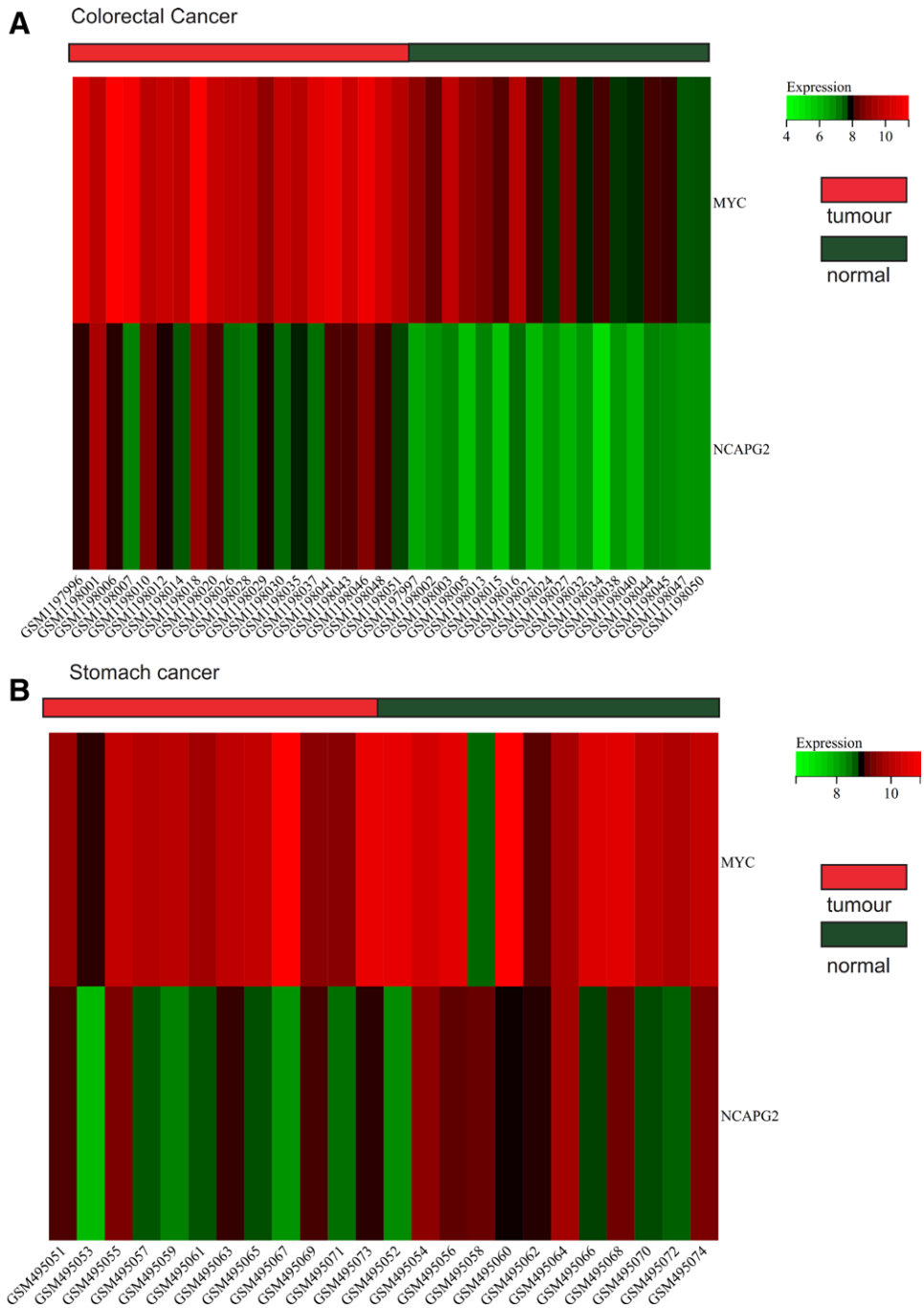

Figure 8. Heat map of gene expression. (A) GSE49355 core gene expression heat map. (B) GSE19826 core gene expression heat map.

cutoff. We also plotted a Venn diagram and obtained the intersection of WGCNA and DEGs to confirm the results (Fig. 6B).

3.6. Construction and analysis of PPI network

The PPI network of DEGs was constructed using the STRING online database and analyzed using Cytoscape software (Fig. 7A). The core gene cluster (Fig. 7B) was obtained and central genes were identified using 3 algorithms (MCC, MNC, DMNC) (Fig. 7C–E). The Venn diagram was used to obtain the union as core genes (Fig. 7F), ultimately identifying 2 core genes (MYC and NCAPG2).

3.7. Heatmap of core gene expression

The expression levels of core genes in colorectal cancer dataset GSE49355 (Fig. 8A) and gastric cancer dataset GSE19826 (Fig. 8B) were visualized and presented in heatmaps. It was observed that core genes (MYC and NCAPG2) were highly expressed in colorectal cancer tissue samples and low expressed in normal tissue samples, while they were core molecules in gastric cancer.

3.8. Survival analysis

Survival data for colorectal cancer and gastric cancer were obtained from the cancer genome atlas. The prognosis score

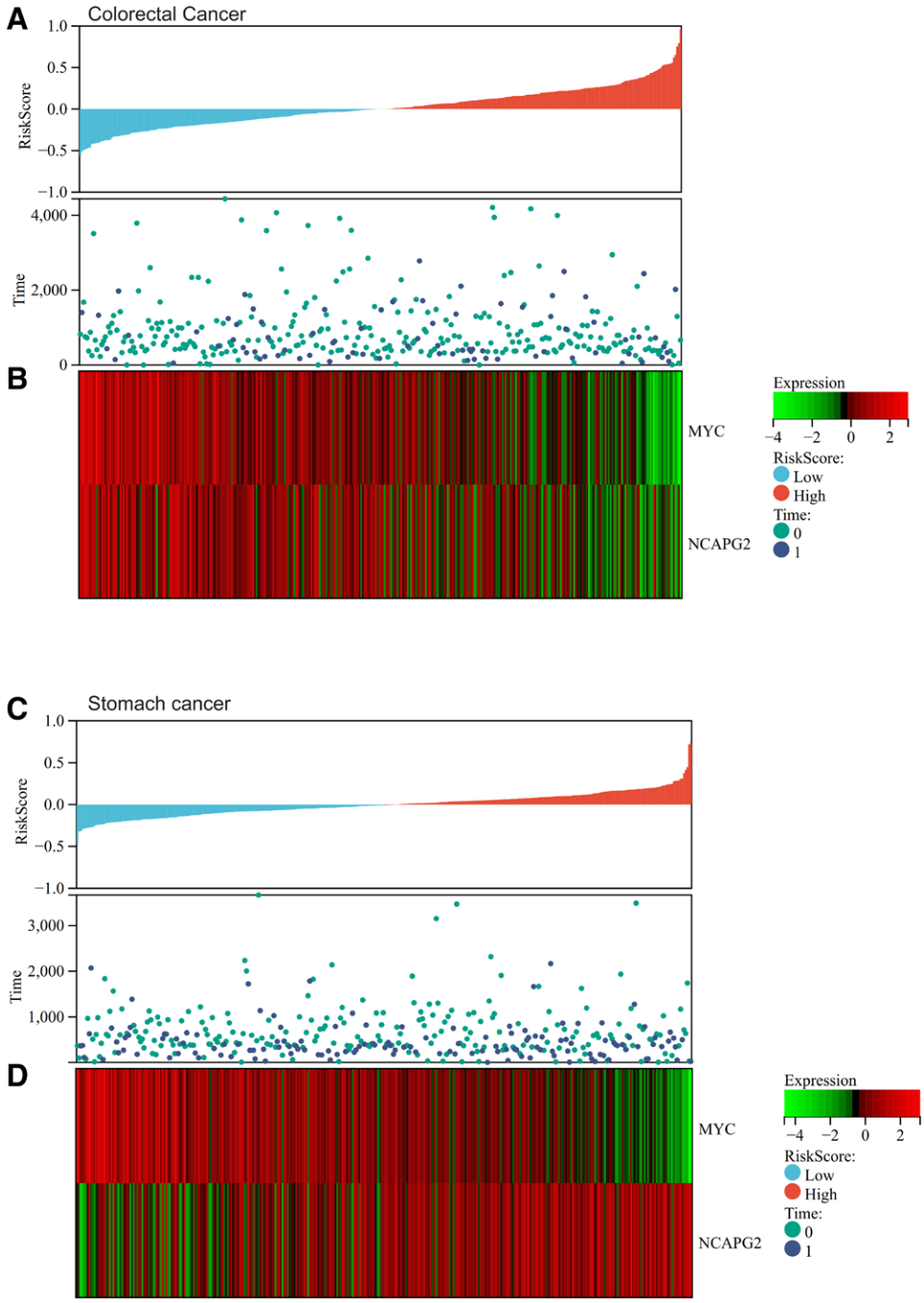

Figure 9. Survival analysis. (A) The effect of GSE49355 on survival time and survival rate in hepatocellular carcinoma patients. (B) Heatmap of GSE49355 core gene expression in liver cancer survival data. (C) The effect of GSE19826 on survival time and survival rate in hepatocellular carcinoma patients. (D) Heatmap of GSE19826 core gene expression in liver cancer survival data.

relationship graph showed a significant decrease in patient survival rates with increasing risk scores, indicating that the low-risk group had significantly higher survival time and rate compared to the high-risk group (Fig. 9A and C). Visualization of the expression heatmap of core genes in colorectal cancer survival data showed that core genes (MYC and NCAPG2) were risk factors, exhibiting an upregulation trend with increasing risk scores (Fig. 9B and D). The results indicated significant differences in core genes (MYC and NCAPG2) among colorectal cancer, gastric cancer, and normal samples. Additionally, forest plots (Fig. 10A and D), receiver operating characteristic curves (Fig. 10B and E), and KM survival curves (Fig, 10C and F) related to core genes (MYC and NCAPG2) in colorectal cancer and gastric cancer were obtained, suggesting a significant correlation between core genes and survival rate, potentially playing a role in the prognosis of colorectal cancer and gastric cancer patients. Boxplots of core genes in colorectal cancer and gastric cancer (Fig. 11A and B) showed significant differences in expression.

3.9. CTD analysis

In this study, we inputted core gene lists into the CTD website to search for diseases related to core genes, enhancing understanding of the association between genes and diseases. It was found that 2 genes (MYC and NCAPG2) were associated with colorectal tumors, gastric tumors, inflammation, and immune system diseases (Fig. 11C).

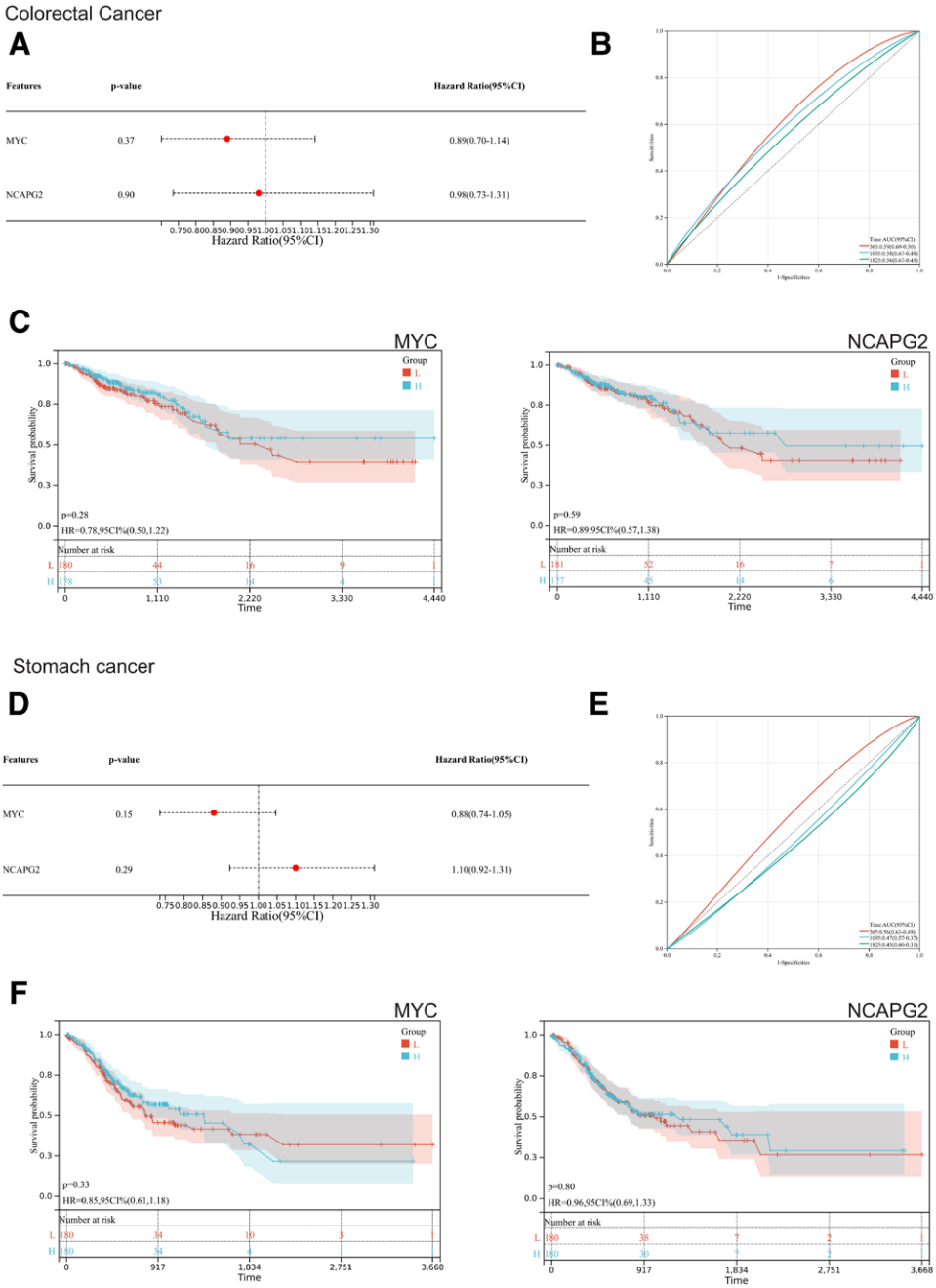

Figure 10. (A) Overall survival of patients with colorectal cancer. (B) ROC curve of risk score. (C) KM survival curve for hazard scores. (D) Overall survival of patients with stomach cancer. (E) ROC curve of risk score. (F) KM survival curve for hazard scores. ROC = receiver operating characteristic.

3.10. Immune infiltration analysis

We used the CIBERSORT package to analyze the gene expression matrices of colorectal cancer dataset GSE49355 and gastric cancer dataset GSE19826. At 95% confidence, we obtained the proportions of immune cells in the full gene expression matrix.

In the colorectal cancer dataset GSE49355, Macrophages M0 were found to be highly represented in the samples (Fig. 12A). Additionally, a heatmap of immune cell expression in the dataset showed high expression of Macrophages M0 in colorectal samples (Fig. 12B). Furthermore, correlation analysis of infiltrating immune cells revealed a co-expression pattern between immune cell components. It was observed that when Dendritic_cells_activated expression was high, Macrophages_M1 expression was low, indicating a highly positive correlation between Dendritic_cells_activated and Macrophages_M1, which may affect the progression of colorectal cancer (Fig. 12C).

The gastric cancer dataset GSE19826 revealed a higher proportion of Plasma cells in the samples (Fig. 13A). Furthermore, an immune cell expression heatmap was generated from the dataset, indicating elevated expression of Plasma cells in gastric cancer (Fig. 13B). Additionally, analysis of the co-expression pattern among infiltrating cell components showed that when Macrophages_M2 expression was high, Plasma_cells expression was low. This suggests a strong positive correlation between Macrophages_M2 and Plasma_cells, which could potentially influence the progression of gastric cancer (Fig. 13C).

4. Discussion

Colorectal cancer and gastric cancer are common malignant tumors of the digestive system. If not treated in a timely manner, they can lead to serious consequences such as bleeding, obstruction, distant metastasis, etc, increasing the difficulty of nursing care. In-depth exploration of the molecular mechanisms of colorectal cancer and gastric cancer is crucial for optimizing nursing care plans. The main results of this study show that the MYC and NCAPG2 genes are highly expressed in colorectal cancer and serve as core molecules in gastric cancer.

The MYC gene, as an important transcription factor, plays a key role in processes such as cell proliferation, differentiation, and apoptosis.[15,16] Currently, foreign literature has extensively studied the relationship between the MYC gene and colorectal cancer. Previous studies have shown that abnormal expression of the MYC gene is closely related to the occurrence and development of colorectal cancer. Its overexpression is associated with the malignancy, prognosis, and chemotherapy resistance of colorectal cancer.[17] In addition, some experimental studies have also revealed the mechanisms by which the MYC gene participates in regulating important biological processes such as proliferation, invasion, and metastasis of colorectal cancer cells.[18,19] Therefore, based on the comprehensive results of foreign literature, it is speculated that the MYC gene may play an important role in the occurrence, development, and metastasis of colorectal cancer. Research has shown that abnormal expression of the MYC gene is closely related to the occurrence and development of gastric cancer. Overexpression of the MYC gene is positively

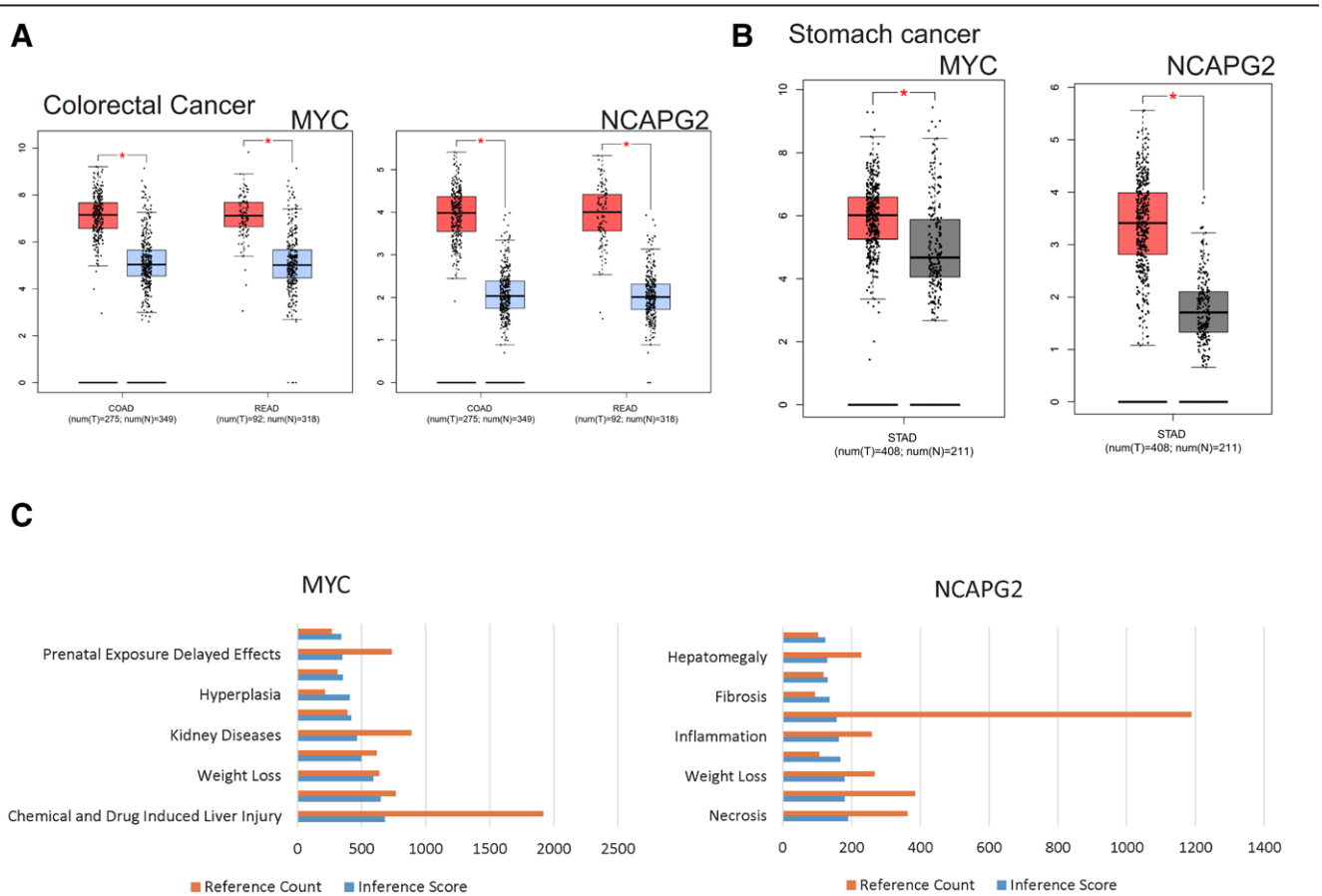

Figure 11. (A) Advances in core gene block diagram of colorectal cancer. (B) Box plots of core genes in gastric cancer. (C) CTD analysis. Two core genes (MYC, NCAPG2) are associated with liver tumors, hepatocellular carcinoma, invasive tumors, and end-stage liver disease. CTD = comparative toxicogenomics database, NCAPG2 = non-SMC condensin II complex subunit G2.

correlated with the incidence and malignancy of gastric cancer, as well as poor prognosis and chemotherapy resistance.[20] Additionally, experimental studies have revealed the mechanisms by which the MYC gene participates in regulating important biological processes such as proliferation, invasion, and metastasis of gastric cancer cells.[21] Based on the comprehensive results of foreign literature, it is speculated that the MYC gene may play an important role in the occurrence, development, and metastasis of gastric cancer.

The NCAPG2 gene encodes proteins related to cell mitosis and is involved in the orderly separation and condensation of chromosomes during nuclear replication.[22] The importance of the NCAPG2 gene in the process of cell mitosis is widely recognized. In recent years, more and more studies have begun to focus on the role of the NCAPG2 gene in tumors.

NCAPG2 is expressed in various types of cells, and its abnormal elevation can affect the assembly and separation of chromosomes in cells, thus playing a role in cancer cell proliferation and metastasis.[23] NCAPG2 is associated with the infiltration of immune cells such as B cells, CD4 + T cells, neutrophils, and macrophages.[24] It is necessary for centromere localization and cell cycle progression, making it a potential target for immune infiltration therapy. Numerous studies have indicated that NCAPG2 is highly expressed in tumors and participates in tumor proliferation, metastasis, and invasion.

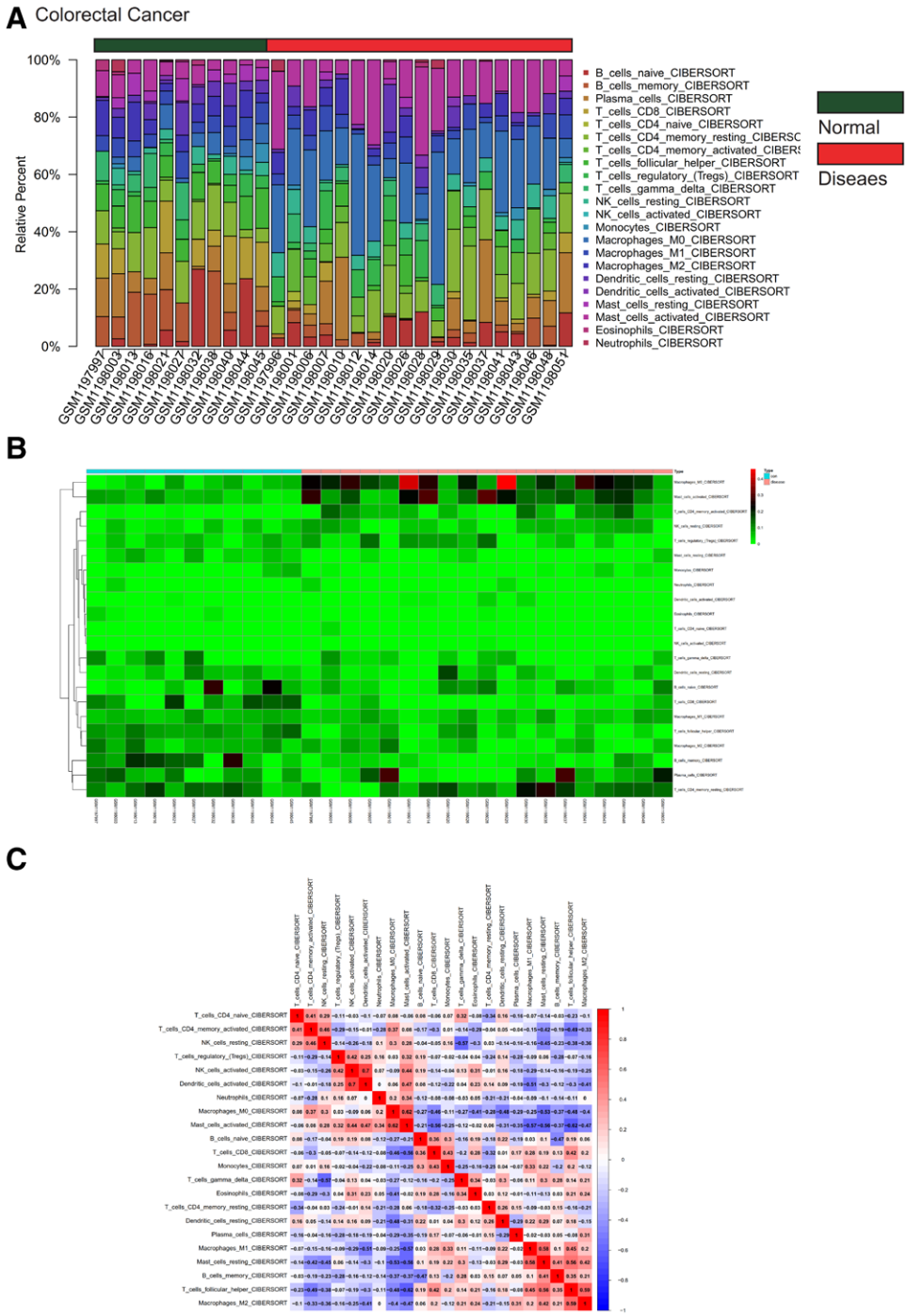

Figure 12. Analysis of immune infiltration in colorectal carcinoma. (A) Whole gene expression matrix results in proportion of immune cells. (B) Heat map of immune cell expression in data set. (C) Map of co-expression patterns between immune cell components.

NCAPG2 promotes the progression of glioblastoma cells[25] and is a novel breast cancer initiator.[26] High expression of NCAPG2 is associated with tumor immune infiltration and the progression of non-small cell lung cancer.[27] Transcription of NCAPG2 promotes the proliferation and migration of liver cancer cells.[28] NCAPG2 is associated with the tumor immune-suppressive microenvironment. Therefore, NCAPG2 may play an important role in the pathological progression of colorectal cancer and gastric cancer.

Despite the rigorous bioinformatics analysis, there are still some shortcomings. In this study, no animal experiments with overexpression or knockdown of the gene were performed to further verify its function. Therefore, we should conduct in-depth exploration in this aspect in future research.

5. Conclusion

MYC and NCAPG2 genes are highly expressed in colorectal cancer, and the higher the expression level, the worse the prognosis. MYC and NCAPG2 are the core molecules in gastric cancer and may be the core targets of colorectal cancer and gastric cancer. In the nursing of colorectal cancer and gastric cancer, we should focus on the changes of these 2 molecules, so as to provide a theoretical basis for the optimization of nursing programs.

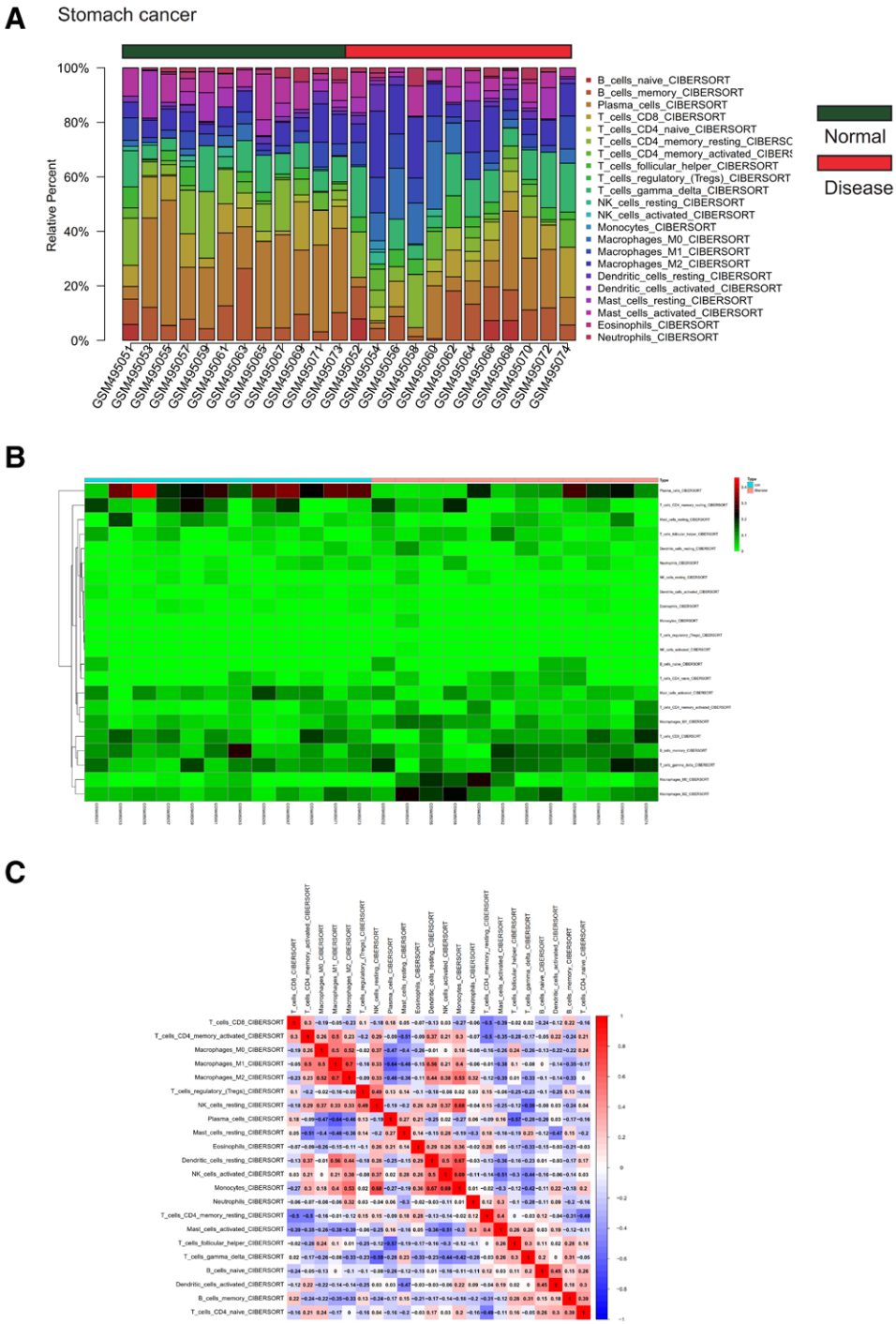

Figure 13. Analysis of immune infiltration in stomach carcinoma. (A) Whole gene expression matrix results in proportion of immune cells. (B) Heat map of immune cell expression in data set. (C) Map of co-expression patterns between immune cell components.

Author contributions

Conceptualization: Xihua Mi, Jie Zhang, Yanfang Gao, Qiang Gao.
Data curation: Chunbo Kang, Shiyang Hou, Yanfang Gao, Lingli Hao, Xiaoli Gao, Qijun Zhang.
Formal analysis: Chunbo Kang, Shiyang Hou, Yanfang Gao, Lingli Hao, Xiaoli Gao, Qijun Zhang, Haifeng Shan.
Investigation: Jie Zhang, Qiang Gao, Xiaoqian Chi, Haifeng Shan.
Methodology: Xihua Mi, Chunbo Kang, Lingli Hao, Xiaoqian Chi, Qijun Zhang, Haifeng Shan.
Project administration: Jie Zhang, Qiang Gao.
Supervision: Jie Zhang, Qiang Gao.
Software: Shiyang Hou, Xiaoli Gao, Xiaoqian Chi.
Writing – review & editing: Xihua Mi, Chunbo Kang, Jie Zhang, Qiang Gao.
Writing – original draft: Chunbo Kang.

References

[1] Chen Y, Bai B, Ying K, et al. Anti-PD-1 combined with targeted therapy: theory and practice in gastric and colorectal cancer. Biochim Biophys Acta Rev Cancer. 2022;1877:188775.
[2] Dekker E, Tanis PJ, Vleugels J, et al. Colorectal cancer. Lancet. 2019;394:1467–80.
[3] Smyth EC, Nilsson M, Grabsch HI, et al. Gastric cancer. Lancet. 2020;396:635–48.
[4] Mahmoud NN. Colorectal cancer: preoperative evaluation and staging. Surg Oncol Clin N Am. 2022;31:127–41.
[5] Machlowska J, Baj J, Sitarz M, et al. Gastric cancer: epidemiology, risk factors, classification, genomic characteristics and treatment strategies. Int J Mol Sci. 2020;21:4012.
[6] Sinha R. Colorectal cancer. Clin Radiol. 2021;76:870.
[7] Sexton RE, Al Hallak MN, Diab M, et al. Gastric cancer: a comprehensive review of current and future treatment strategies. Cancer Metastasis Rev. 2020;39:1179–203.
[8] Gauthier J, Vincent AT, Charette SJ, et al. A brief history of bioinformatics. Brief Bioinform. 2019;20:1981–96.
[9] Uesaka K, Oka H, Kato R, et al. Bioinformatics in bioscience and bioengineering: recent advances, applications, and perspectives. J Biosci Bioeng. 2022;134:363–73.
[10] Holtsträter C, Schrörs B, Bukur T, et al. Bioinformatics for cancer immunotherapy. Methods Mol Biol. 2020;2120:1–9.
[11] Das SK, Lewis BA, Levens D. MYC: a complex problem. Trends Cell Biol. 2023;33:235–46.
[12] Feng H, Chen N, Li H, et al. Pan-cancer analysis of the immune aspects and prognostic value of NCAPG2. Heliyon. 2023;9:e18051.
[13] Duffy MJ, O'Grady S, Tang M, et al. MYC as a target for cancer treatment. Cancer Treat Rev. 2021;94:102154.
[14] Wang Q, Li Z, Zhou S, et al. NCAPG2 could be an immunological and prognostic biomarker: from pan-cancer analysis to pancreatic cancer validation. Front Immunol. 2023;14:1097403.
[15] Purhonen J, Klefström J, Kallijärvi J. MYC-an emerging player in mitochondrial diseases. Front Cell Dev Biol. 2023;11:1257651.
[16] Devaiah BN, Mu J, Akman B, et al. MYC protein stability is negatively regulated by BRD4. Proc Natl Acad Sci U S A. 2020;117:13457–67.
[17] Saeed H, Leibowitz BJ, Zhang L, et al. Targeting Myc-driven stress addiction in colorectal cancer. Drug Resist Updat. 2023;69:100963.
[18] Tan L, Peng D, Cheng Y. Significant position of C-myc in colorectal cancer: a promising therapeutic target. Clin Transl Oncol. 2022;24:2295–304.
[19] Wang H, Yang W, Qin Q, et al. E3 ubiquitin ligase MAGI3 degrades c-Myc and acts as a predictor for chemotherapy response in colorectal cancer. Mol Cancer. 2022;21:151.
[20] Anauate AC, Leal MF, Calcagno DQ, et al. The complex network between MYC oncogene and microRNAs in gastric cancer: an overview. Int J Mol Sci. 2020;21:1782.
[21] Liu M, Yao B, Gui T, et al. PRMT5-dependent transcriptional repression of c-Myc target genes promotes gastric cancer progression. Theranostics. 2020;10:4437–52.
[22] Feng Z, Zhang L, Liu Y, et al. NCAPG2 contributes to the progression of malignant melanoma through regulating proliferation and metastasis. Biochem Cell Biol. 2022;100:473–84.
[23] Kim JH, Shim J, Ji MJ, et al. The condensin component NCAPG2 regulates microtubule-kinetochore attachment through recruitment of Polo-like kinase 1 to kinetochores. Nat Commun. 2014;5:4588.
[24] Meng F, Zhang S, Song R, et al. NCAPG2 overexpression promotes hepatocellular carcinoma proliferation and metastasis through activating the STAT3 and NF-κB/miR-188-3p pathways. EBioMedicine. 2019;44:237–49.
[25] Wu J, Li L, Jiang G, et al. NCAPG2 facilitates glioblastoma cells' malignancy and xenograft tumor growth via HBO1 activation by phosphorylation. Cell Tissue Res. 2021;383:693–706.
[26] Han S, Jin X, Hu T, et al. The mRNA stability of NCAPG2, a novel contributor to breast invasive carcinoma, is enhanced by the RNA-binding protein PCBP2. Cell Signal. 2023;110:110844.
[27] Chen X, Guo J, Ren W, et al. LncRNA-AL035458.2/hsa-miR-181a-5p axis-mediated high expression of NCAPG2 correlates with tumor immune infiltration and non-small cell lung cancer progression. Front Oncol. 2022;12:910437.
[28] Li S, Lu Y, Xu Y, et al. Brachyury promotes proliferation and migration of hepatocellular carcinoma via facilitating the transcription of NCAPG2. Am J Cancer Res. 2022;12:3625–43.