

Identification of Key Genes for Controlling Cancer Stem Cell Characteristics in Gastric Cancer

Chao Huang

Nanchang University Second Affiliated Hospital

Jiefeng Zhao

Nanchang University Second Affiliated Hospital

Zhikun Ning

Nanchang University Second Affiliated Hospital

Jun Huang

Nanchang University Second Affiliated Hospital

Zhengming Zhu (✉ zzm8654@163.com)

Nanchang University Second Affiliated Hospital <https://orcid.org/0000-0003-1078-9538>

Primary research

Keywords: gastric cancer, cancer stem cell, key gene, TCGA, WGCNA, mRNAsi

Posted Date: June 2nd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-31510/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Self-renewal of gastric cancer stem cells (GCSCs) is considered to be the underlying cause of the metastasis, drug resistance and recurrence of gastric cancer (GC). The purpose of this study was to characterize the expression of stem cell-related genes in GC.

Methods

The RNA sequencing results and clinical data for gastric adenoma and adenocarcinoma samples were obtained from the Cancer Genome Atlas (TCGA) database, and the results of the GC mRNA expression-based stemness index (mRNAsi) were analyzed. Weighted gene coexpression network analysis (WGCNA) was then used to find modules of interest and their key genes. Survival analysis of key genes was performed using the online tool Kaplan-Meier Plotter, and the online database Oncomine was used to assess the expression of key genes in GC.

Results

mRNAsi was significantly upregulated in GC tissues compared to normal gastric tissues ($p < 0.0001$). A total of 16 modules were obtained from the gene coexpression network; the brown module was most positively correlated with mRNAsi. Sixteen key genes (BUB1, BUB1B, NCAPH, KIF14, RACGAP1, RAD54L, TPX2, KIF15, KIF18B, CENPF, TTK, KIF4A, SGOL2, PLK4, XRCC2 and C1orf112) were identified in the brown module. Survival analysis and Oncomine analysis revealed that the prognosis of patients with GC and the expression of three genes (RAD54L, TPX2 and XRCC2) were consistently related.

Conclusion

Sixteen key genes play an important role in the maintenance of GCSCs. RAD54L, TPX2 and XRCC2 are the most likely therapeutic targets for inhibiting the stemness characteristics of GC cells.

Background

Gastric cancer (GC) is a common malignant tumor, and its mortality rate ranks third among cancers globally [1]. In China, GC is also the major cause of cancer-related death, and patients with advanced disease have lower survival rates and higher recurrence rates [2]. Self-renewal of gastric cancer stem cells (GCSCs) is considered to be the underlying cause of the metastasis, drug resistance and recurrence of GC [3–4]. The identification and targeted therapy of cancer stem cells was of great significance in the treatment of GC [5]. The mRNAsi, which can be used as a quantitative representation of cancer cell stemness, is an indicator describing the degree of similarity between tumor cells and stem cells.

Moreover, inhibiting DNA methyltransferase can reduce the accumulation of tumorigenic ability of GCSCs [6]. The transcriptomes and methylomes were analyzed on multiple platforms to quantify stemness, and the DNA methylation based stemness index (mDNAsi) and mRNAsi were obtained; this analysis was also applied to the TCGA dataset to calculate their scores of those samples [7]. Weighted gene coexpression network analysis (WGCNA) determines the correlation between genes based on systematic biological methods; it is used not only to build gene networks and detect modules but also to identify key genes [8–9]. WGCNA has been widely used to screen key genes related to the clinical characteristics of many tumour types, such as pancreatic cancer and kidney cancer [10–11]. In addition, the genes that regulate the maintenance and proliferation of GCSCs remain unknown. This study aimed to identify key genes related to stemness by combining WGCNA with GC mRNAsi in TCGA, thus providing new ideas for the treatment of GC.

Materials And Methods

Data download and processing

RNA sequencing results of 373 tissues and 348 human gastric adenomas and adenocarcinoma samples were obtained from TCGA database (<https://portal.gdc.cancer.gov>). The RNA-seq results of 30 normal samples and 343 cancer samples were merged into a matrix file using a script in the Perl language (<http://www.perl.org/>). We then converted the Ensembl ID in the matrix file to the gene name using the Ensembl database (<http://www.ensembl.org/index.html>) and the Perl language script. In addition, 406 pieces of clinical data were downloaded, and the relevant clinical data were collated and extracted using scripts in the Perl language. The calculation of mRNAsi was performed from the molecular spectrum of normal cells with different degrees of stemness [7].

Screening of differentially expressed genes (DEG)

The statistical software R (version 3.6.0, <https://www.r-project.org/>), limma package and Wilcoxon test were used to screen DEG expression data between normal gastric tissue and GC samples. The values of the genes with the same name were averaged and the genes with expression levels < 0.2 were deleted. The false discovery rate (FDR) < 0.05 and $|\log_2 \text{fold change}| > 1$ were used as screening criteria.

WGCNA

Construction of WGCNA

The WGCNA R package was used to perform the WGCNA. The normal sample was deleted first, and the data were checked next for missing values. If a value was missing, the data was deleted. First, the similarity matrix was constructed by calculating the correlation of all genes. Second, the WGCNA software package was used to select appropriate soft-thresholding power β to improve coexpression similarity and achieve a scale-free topology. Third, the adjacency of genes was transformed into a topological overlap matrix (TOM), and then the corresponding dissimilarity was calculated. Average

linkage hierarchical clustering was conducted with TOM-based dissimilarity measurements, and the minimum size of the gene dendrogram was 30. Finally, the dissimilarity of the genes was calculated and module dendrograms were built.

Identification of important modules and key genes

Gene significance (GS) refers to the correlation between genes and sample traits. Statistical significance was determined using the associated p values. To reduce the number of modules, we set a threshold (< 0.25) to merge some modules that were highly similar. The gene modules were then analyzed in combination with the clinical phenotype. Module membership (MM) refers to the correlation between a module's own genes and the gene expression profile. After selecting the module of interest, GS and MM for each key gene were calculated. We defined cor. gene MM > 0.8 and cor. gene GS > 0.5 as the threshold for the key genes of the module.

Differential expression and correlation of key genes

We used the ggpubr package of the R language to draw a box plot of differential gene expression to understand the differential expression of key genes. The pheatmap package was used to create heat maps of key genes. Correlation analysis of key genes was performed using the corrplot package.

Protein-protein Interaction (PPI) network analysis of key genes

To explore the interaction between key genes, we imported the key genes into the Search Tool for the Retrieval of Interacting Genes (STRING, <http://string.embl.de/>) database [12]. The PPI network of key genes, which included genes with a combined score of > 0.4 , was constructed using STRING.

Functional and pathway enrichment analyses of key genes

To study the biological functions of the module genes and key genes, we used the clusterProfiler package for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses [13]. $P < 0.05$ and $FDR < 0.05$ were used as thresholds for identifying significant GO terms and pathways. A histogram of GO enrichment results for the key genes was plotted using the ggplot2 package. GO includes cellular component (CC), biological process (BP) and molecular function (MF).

Prognostic analysis of key genes

To understand the relationship between the expression of key genes and prognosis, we used the online tool Kaplan–Meier Plotter (<http://kmplot.com/>) for survival analysis. A log-rank $P < 0.05$ was considered statistically significant. In addition, we used the online database Oncomine (<http://www.oncomine.com>) to assess the expression of key genes in GC.

Results

Differential analysis of mRNAsi and screening of DEG

The Wilcoxon test showed a significant difference in the mRNAsi of normal gastric tissue and that of gastric tumor tissue ($p < 0.0001$, Fig. 1A). A total of 6,685 DEGs were identified by Wilcoxon test analysis, among which 5,479 were upregulated and 1,206 were downregulated (**Supplementary Table 1**). The 20 with the most significant upregulation and the 20 with the most significant downregulated were mapped using the pheatmap package (Fig. 1B).

WGCNA: Identification of the modules of interest

To identify biologically important gene modules and better understand genes associated with gastric cancer stemness, we used the WGCNA package to construct a gene coexpression network. Ultimately, we obtained 16 modules (Fig. 1C); the brown module was most positively correlated with mRNAsi with a correlation coefficient of 0.75. In addition, the turquoise module exhibited a high negative correlation with mRNAsi with a correlation coefficient of -0.77 (Fig. 1D). Therefore, the brown module was used as the most interesting module, so we further analyzed this module.

Identification of key genes

According to the threshold of key gene screening, we screened 16 key genes in the mRNAsi group, namely, BUB1, BUB1B, NCAPH, KIF14, RACGAP1, RAD54L, TPX2, KIF15, KIF18B, CENPF, TTK, KIF4A, SGOL2, PLK4, XRCC2 and C1orf112 (Fig. 1E).

Differential expression and correlation analysis of key genes

The expression of key genes differed significantly between normal gastric tissue and tumor tissue. The expression of key genes in gastric tumors was significantly higher than that in normal gastric tissues (p all < 0.001 , Fig. 2A). There was a clear correlation between the expressions of key genes, and CENPF had the highest correlation with KIF14 (Fig. 2B).

PPI network analysis of key genes

The PPI network of key genes consisted of 16 nodes and 94 edges, including 16 upregulated genes (Fig. 2C). In the PPI network, TTK, TPX2, NCAPH, KIF15, CENPF and BUB1 all had 14 connection nodes (Fig. 2D).

GO and KEGG analyses of key genes

GO and KEGG pathway enrichment analysis was performed using the clusterProfiler package. The most significantly enriched GO terms were spindle (ontology: CC), sister chromatid segregation (ontology: BP) and motor activity (ontology: MF) (Fig. 3). The significantly enriched KEGG pathways have cell cycle and homologous recombination (Table 1).

Table 1
Significantly enriched KEGG pathways

Category	Term	Description	p.adjust	Gene	Gene counts
hsa04110	KEGG	Cell cycle	0.0004	BUB1/BUB1B/TTK	3
hsa03440	KEGG	Homologous recombination	0.001	RAD54L/XRCC2	2

Prognostic analysis of key genes

Survival analysis of key genes was performed using Kaplan-Meier curves, which showed that all key genes except KIF14 and KIF18B were related to prognosis. GC patients with high expression of RAD54L, TPX2 and XRCC2 had short survival times, while GC patients with high expression of other key genes had long survival times (Fig. 4). However, differential analysis showed that 16 key genes were highly expressed in GC tissues. We further analyzed 16 key genes using the online database Oncomine, which showed that 16 key genes were highly expressed in GC tissues (Fig. 5).

Discussion

Metastasis, drug resistance and recurrence are the main causes of the low survival rate of patients with GC. Self-renewal of GCSCs is considered to be the underlying cause of the metastasis, drug resistance and recurrence of GC [3–4]. After conventional surgical treatment or adjuvant chemotherapy, the number of GCSCs does not decrease but stemness is enriched, leading to metastasis, drug resistance and recurrence of GC [3]. Moreover, Bekaii-Saab et al. [5] indicated that identification and targeted therapy of cancer stem cells had important significance in the treatment of GC. Therefore, targeted therapy of GCSCs is crucial. This study used WGCNA based on the mRNAsi score to identify key genes related to GCSC characteristics. This study found that the mRNAsi of gastric tumor tissue was significantly higher than that of normal gastric tissue. In addition, we used the WGCNA package to construct a gene coexpression network. Finally, we obtained a total of 16 modules; the brown module was the most positively correlated with mRNAsi, and the turquoise module showed a high negative correlation with mRNAsi. Therefore, we selected the brown module as the module of most interest and identified 16 key genes that are associated with the characteristics of GCSCs: BUB1, BUB1B, NCAPH, KIF14, RACGAP1, RAD54L, TPX2, KIF15, KIF18B, CENPF, TTK, KIF4A, SGOL2, PLK4, XRCC2 and C1orf112. These genes were upregulated in GC and had a certain correlation. The most significantly enriched GO terms of these key genes were the spindle cellular component, the sister chromatid segregation biological process and the motor activity molecular function. The significantly enriched KEGG pathways were cell cycle and homologous recombination. Except for KIF14 and KIF18B, the expression of all key genes was related to the prognosis of patients with GC.

BUB1 encodes a protein that is essential for the function of the mitotic spindle check-point [14]. Grabsch et al. [15] showed that BUB1 was overexpressed in up to 80% of GCs and was associated with tumor cell proliferation. Meanwhile, related studies indicated that high expression of BUB1B was related to the invasion, lymph node metastasis, liver metastasis and recurrence of GC [16]. Kinesins are a family of

molecular motors that play important roles in intracellular transport and cell division [17]. Previous studies demonstrated that KIF14 can enhance tumor adhesion, invasion and chemical resistance during tumor development [18]. Tong et al. [19] also revealed that KIF14 was overexpressed in GC and that this was related to tumor progression, invasion and metastasis. Spheroid colony formation is an effective model for the characterization of cancer stem cells [4]. Oue et al. [20] showed that KIF15 and KIF4A were upregulated in GC spheroid cells. A previous study also demonstrated that patients with high expression of KIF14 mRNA had a higher risk of malignancy, recurrence and metastasis than did those without high expression of KIF14 mRNA [21]. Overexpression of KIF18B increases the proliferation of hepatocellular carcinoma cells, and downregulation of KIF18B in vitro can inhibit the migration and invasion of cervical cancer cells [22]. KIF4A is considered to be an oncogene, indicating its involvement in malignancy, and its expression was indeed related to the occurrence, development and metastasis of GC [23]. Rac GTPase accelerating protein1 (RacGAP1), which belongs to the GTPase activation protein family, not only induces cytokinesis but also interferes with mitotic spindle assembly, and thus participating in the regulation of cell proliferation [24]. Studies have indicated that RacGAP1 is involved in cell transformation, movement, migration and metastasis [25] and is positively correlated with the proliferation marker Ki67 [26]. In addition, RacGAP1 is highly expressed in GC and significantly associated with tumor progression and poor prognosis [27]. CENPF is a member of the mitochondrial family and regulates the proliferation of various tumor cells [28]. Chen et al. [29] showed that CENPF was overexpressed in GC and promoted its proliferation and metastasis. TTK is also known as a mitotic kinase [30], which is highly expressed in many types of tumors and promotes the proliferation of cancer cells [31–32]. Park et al. [33] found that TTK plays an indispensable role in the development and maintenance of tumor stem cells. PLK4 is a major regulator of centromeric repeats, and its overexpression in somatic cells can disrupt spindle formation, while its depletion can inhibit cell proliferation [34]. Shinmura et al. [35] revealed that PLK4 is overexpressed in GC and plays an important role in cell proliferation, tumorigenesis, invasion and drug resistance. C1ORF112 is overexpressed in various tumors, including GC, and plays an important role in the growth of cancer cells, but it is still unclear whether C1ORF112 is particularly necessary for the proliferation of cancer cells [36]. Condensins are multiprotein complexes that play a central role in chromosome assembly and segregation during mitosis and meiosis [37]. Non-SMC condensin I complex subunit H (NCAPH) is one of the three non-SMC subunits of condensation I. Studies have indicated that non-SMC condensin I complex subunit G (NCAPG) is overexpressed in colon cancer and hepatocellular carcinoma and promotes the proliferation and migration of tumor cells [38].

In addition, survival analysis and Oncomine analysis revealed that the prognosis of patients with GC and the expression of three genes (RAD54L, TPX2 and XRCC2) were consistently related. XRCC2 plays a crucial role in DNA repair and chromosome arrangement, and its dysfunction may lead to tumor development and progression [39–40]. Wang et al. [41] found that XRCC2 is overexpressed in GC. RAD54L and XRCC2 are homologous recombination factors that are overexpressed in lung cancer and GC, and there is a significant correlation between the expression levels of these two proteins [42]. The inactivation of these factors leads to an obvious deficiency in homologous recombination and DNA repair in all eukaryotes [43]. In addition, related studies indicated that abnormal function of homologous

recombination leads to genomic instability, resulting in the occurrence of cancer [44–45]. In this study, KEGG enrichment analysis indicated that RAD54L and XRCC2 were significantly enriched in homologous recombination. TPX2 is a microtubule-associated protein that activates the cell cycle kinase protein Aurora-A and then plays an important role in the formation of spindles in mitosis [46]. Liang et al. [47] showed that TPX2 was overexpressed in GC and was related to poor progression and prognosis of GC.

In conclusion, this study found that the 16 key genes identified play an important role in the maintenance of GCSCs. In addition, RAD54L, TPX2 and XRCC2 are most likely to be therapeutic targets for inhibiting the stemness characteristics of GC, providing a new idea for the treatment of GC. However, further research is needed to verify this hypothesis.

Abbreviations

GCSCs: gastric cancer stem cells; GC: gastric cancer; TCGA: the Cancer Genome Atlas; mRNAsi: mRNA expression-based stemness index; WGCNA: weighted gene coexpression network analysis; mDNAsi: DNA methylation based stemness index; DEG: differentially expressed genes; TOM: topological overlap matrix; GS: gene significance; MM: module membership; PPI: protein-protein interaction; GO: gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; CC: cellular component; BP: biological process; MF: molecular function; RacGAP1: Rac [GTPase](#) accelerating protein1; NCAPH: Non-SMC condensin I complex subunit H; NCAPG: non-SMC condensin I complex subunit G

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The results shown here are in whole or part based upon data generated by the TCGA Research Network: <http://cancergenome.nih.gov/>.

Competing interests

The authors declare no competing financial interests.

Funding

This present study was supported by the National Natural Science Foundation of China (Grant Number: 81560389) and Key Research and Development Program of Jiangxi Province (Grant Number:

Author contributors

CH designed the study. CH, JFZ, and ZKN collected data. CH and JH analysed the data. CH wrote the manuscript with contribution from all authors. All authors read and approved the final version of the paper.

Acknowledgments

Not applicable.

Author details

¹ Department of Gastrointestinal Surgery, The Second Affiliated Hospital of Nanchang University, Nanchang, 330006, Jiangxi, China.

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. *Cancer J Clin*. 2018;68:394–424.
2. Zong L, Abe M, Seto Y, Ji J. The challenge of screening for early gastric cancer in China. *The Lancet*. 2016;388:2606.
3. Hayakawa Y, Ariyama H, Stancikova J, Sakitani K, Asfaha S, Renz BW, Dubeykovskaya ZA, Shibata W, Wang H, et al. Mist1 expressing gastric stem cells maintain the normal and neoplastic gastric epithelium and are supported by a perivascular stem cell niche. *Cancer cell*. 2015;28:800–14.
4. Takaishi S, Okumura T, Wang TC. Gastric cancer stem cells. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*. 2008;26:2876–82.
5. Bekaii-Saab T, El-Rayes B. Identifying and targeting cancer stem cells in the treatment of gastric cancer. *Cancer*. 2017;123:1303–12.
6. Norollahi SE, Mansour-Ghanaei F, Joukar F, Ghadarjani S, Mojtahedi K, Nejad KG, Hemmati H, Gharibpoor F, Khaksar R, et al. Therapeutic approach of Cancer stem cells (CSCs) in gastric adenocarcinoma; DNA methyltransferases enzymes in cancer targeted therapy. *Biomed Pharmacother*. 2019;115:108958.
7. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, Kamińska B, Huelsken J, Omberg L, et al. Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell*. 2018;173:338–54. e315.
8. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform*. 2008;9:559.

9. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Statistical applications in genetics and molecular biology* 2005; 4: Article 17.
10. Yuan L, Zeng G, Chen L, Wang G, Wang X, Cao X, Lu M, Liu X, Qian G, et al. Identification of key genes and pathways in human clear cell renal cell carcinoma (ccRCC) by co-expression analysis. *Int J Biol Sci.* 2018;14:266.
11. Zhou Z, Cheng Y, Jiang Y, Liu S, Zhang M, Liu J, Zhao Q. Ten hub genes associated with progression and prognosis of pancreatic carcinoma identified by co-expression analysis. *Int J Biol Sci.* 2018;14:124–36.
12. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, et al. STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic acids research.* 2014;43:D447–52.
13. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: a journal of integrative biology.* 2012;16:284–7.
14. Sharp-Baker H, Chen R-H. Spindle checkpoint protein Bub1 is required for kinetochore localization of Mad1, Mad2, Bub3, and CENP-E, independently of its kinase activity. *J Cell Biol.* 2001;153:1239–50.
15. Grabsch H, Takeno S, Parsons WJ, Pomjanski N, Boecking A, Gabbert HE, Mueller W. Overexpression of the mitotic checkpoint genes BUB1, BUBR1, and BUB3 in gastric cancer—association with tumour cell proliferation. *The Journal of Pathology: A Journal of the Pathological Society of Great Britain Ireland.* 2003;200:16–22.
16. Ando K, Kakeji Y, Kitao H, Imori M, Zhao Y, Yoshida R, Oki E, Yoshinaga K, Matumoto T, et al. High expression of BUBR1 is one of the factors for inducing DNA aneuploidy and progression in gastric cancer. *Cancer Sci.* 2010;101:639–45.
17. Hirokawa N. Kinesin and dynein superfamily proteins and the mechanism of organelle transport. *Science.* 1998;279:519–26.
18. Ahmed SM, Thériault BL, Uppalapati M, Chiu CW, Gallie BL, Sidhu SS, Angers S. KIF14 negatively regulates Rap1a–Radil signaling during breast cancer progression. *J Cell Biol.* 2012;199:951–67.
19. Tong H, Wang J, Chen H, Wang Z, Fan H, Ni Z. Transcriptomic analysis of gene expression profiles of stomach carcinoma reveal abnormal expression of mitotic components. *Life sciences.* 2017;170:41–9.
20. Oue N, Mukai S, Imai T, Pham TT, Oshima T, Sentani K, Sakamoto N, Yoshida K and Yasui W. Induction of KIFC1 expression in gastric cancer spheroids. *Oncol Rep.* 2016;36:349–55.
21. Nakamura Y, Tanaka F, Haraguchi N, Mimori K, Matsumoto T, Inoue H, Yanaga K, Mori M. Clinicopathological and biological significance of mitotic centromere-associated kinesin overexpression in human gastric cancer. *British journal of cancer.* 2007;97:543–9.
22. Wu Y, Wang A, Zhu B, Huang J, Lu E, Xu H, Xia W, Dong G, Jiang F, et al. KIF18B promotes tumor progression through activating the Wnt/ β -catenin pathway in cervical cancer. *OncoTargets therapy.* 2018;11:1707.

23. Taniwaki M, Takano A, Ishikawa N, Yasui W, Inai K, Nishimura H, Tsuchiya E, Kohno N, Nakamura Y, et al. Activation of KIF4A as a prognostic biomarker and therapeutic target for lung cancer. *Clin Cancer Res.* 2007;13:6624–31.
24. Zhao W-m, Fang G. MgcRacGAP controls the assembly of the contractile ring and the initiation of cytokinesis. *Proceedings of the National Academy of Sciences* 2005; 102: 13158–13163.
25. Sanz-Moreno V, Gadea G, Ahn J, Paterson H, Marra P, Pinner S, Sahai E, Marshall CJ. Rac activation and inactivation control plasticity of tumor cell movement. *Cell.* 2008;135:510–23.
26. Ke H-L, Ke R-H, Li S-T, Li B, Lu H-T, Wang X-Q. Expression of RACGAP1 in high grade meningiomas: a potential role in cancer progression. *Journal of neuro-oncology.* 2013;113:327–32.
27. Saigusa S, Tanaka K, Mohri Y, Ohi M, Shimura T, Kitajima T, Kondo S, Okugawa Y, Toiyama Y, et al. Clinical significance of RacGAP1 expression at the invasive front of gastric cancer. *Gastric cancer.* 2015;18:84–92.
28. Aytes A, Mitrofanova A, Lefebvre C, Alvarez MJ, Castillo-Martin M, Zheng T, Eastham JA, Gopalan A, Pienta KJ, et al. Cross-species regulatory network analysis identifies a synergistic interaction between FOXM1 and CENPF that drives prostate cancer malignancy. *Cancer cell.* 2014;25:638–51.
29. Chen E-B, Qin X, Peng K, Li Q, Tang C, Cheng X, Wei Y-C, Yu S, Gan L, et al HnRNPR/CCNB1/CENPF Axis Contributes to Gastric Cancer Proliferation and Metastasis. *Aging (Albany NY)* 2019; 11, 7473–7491.
30. Dominguez-Brauer C, Thu KL, Mason JM, Blaser H, Bray MR, Mak TW. Targeting mitosis in cancer: emerging strategies. *Molecular cell.* 2015;60:524–36.
31. Jemaa M, Galluzzi L, Kepp O, Senovilla L, Brands M, Boemer U, Koppitz M, Lienau P, Prechtel S, et al. Characterization of novel MPS1 inhibitors with preclinical anticancer activity. *Cell death differentiation.* 2013;20:1532–45.
32. Maia AR, de Man J, Boon U, Janssen A, Song J-Y, Omerzu M, Sterrenburg J, Prinsen M, Willemsen-Seegers N, et al. Inhibition of the spindle assembly checkpoint kinase TTK enhances the efficacy of docetaxel in a triple-negative breast cancer model. *Ann Oncol.* 2015;26:2180–92.
33. Park J-H, Chung S, Matsuo Y, Nakamura Y. Development of small molecular compounds targeting cancer stem cells. *MedChemComm.* 2017;8:73–80.
34. Coelho PA, Bury L, Sharif B, Riparbelli MG, Fu J, Callaini G, Glover DM, Zernicka-Goetz M. Spindle formation in the mouse embryo requires Plk4 in the absence of centrioles. *Developmental cell.* 2013;27:586–97.
35. Shinmura K, Kurabe N, Goto M, Yamada H, Natsume H, Konno H, Sugimura H. PLK4 overexpression and its effect on centrosome regulation and chromosome stability in human gastric cancer. *Molecular biology reports.* 2014;41:6635–44.
36. Cordeiro R. In vitro characterization of a new cancer-related gene. *ResearchGate.* 2012.
37. Wood AJ, Severson AF, Meyer BJ. Condensin and cohesin complexity: the expanding repertoire of functions. *Nat Rev Genet.* 2010;11:391–404.

38. Zhang Q, Su R, Shan C, Gao C, Wu P. Non-SMC condensin I complex, subunit G (NCAPG) is a novel mitotic gene required for hepatocellular cancer cell proliferation and migration. *Oncology Research Featuring Preclinical Clinical Cancer Therapeutics*. 2018;26:269–76.
39. Curtin K, Lin W-Y, George R, Katory M, Shorto J, Cannon-Albright LA, Smith G, Bishop DT, Cox A, et al. Genetic variants in XRCC2: new insights into colorectal cancer tumorigenesis. *Cancer Epidemiology Prevention Biomarkers*. 2009;18:2476–84.
40. Griffin CS, Simpson PJ, Wilson CR, Thacker J. Mammalian recombination-repair genes XRCC2 and XRCC3 promote correct chromosome segregation. *Nat Cell Biol*. 2000;2:757.
41. Wang N, Dong X-j, Zhou R-m, Guo W, Zhang X-j, Li Y. An investigation on the polymorphisms of two DNA repair genes and susceptibility to ESCC and GCA of high-incidence region in northern China. *Molecular biology reports*. 2009;36:357–64.
42. Oji Y, Tatsumi N, Kobayashi J, Fukuda M, Ueda T, Nakano E, Saito C, Shibata S, Sumikawa M, et al. Wilms' tumor gene WT1 promotes homologous recombination-mediated DNA damage repair. *Molecular carcinogenesis*. 2015;54:1758–71.
43. Essers J, Hendriks RW, Swagemakers SM, Troelstra C, de Wit J, Bootsma D, Hoeijmakers JH, Kanaar R. Disruption of mouse RAD54 reduces ionizing radiation resistance and homologous recombination. *Cell*. 1997;89:195–204.
44. Jasin M. Homologous repair of DNA damage and tumorigenesis: the BRCA connection. *Oncogene*. 2002;21:8981–93.
45. Venkitaraman AR. Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell*. 2002;108:171–82.
46. Özlü N, Srayko M, Kinoshita K, Habermann B, O'Toole ET, Müller-Reichert T, Schmalz N, Desai A, Hyman AA. An essential function of the *C. elegans* ortholog of TPX2 is to localize activated aurora A kinase to mitotic spindles. *Developmental cell*. 2005;9:237–48.
47. Liang B, Zheng W, Fang L, Wu L, Zhou F, Yin X, Yu X, Zou Z. Overexpressed targeting protein for Xklp2 (TPX2) serves as a promising prognostic marker and therapeutic target for gastric cancer. *Cancer Biol Ther*. 2016;17:824–32.

Figures

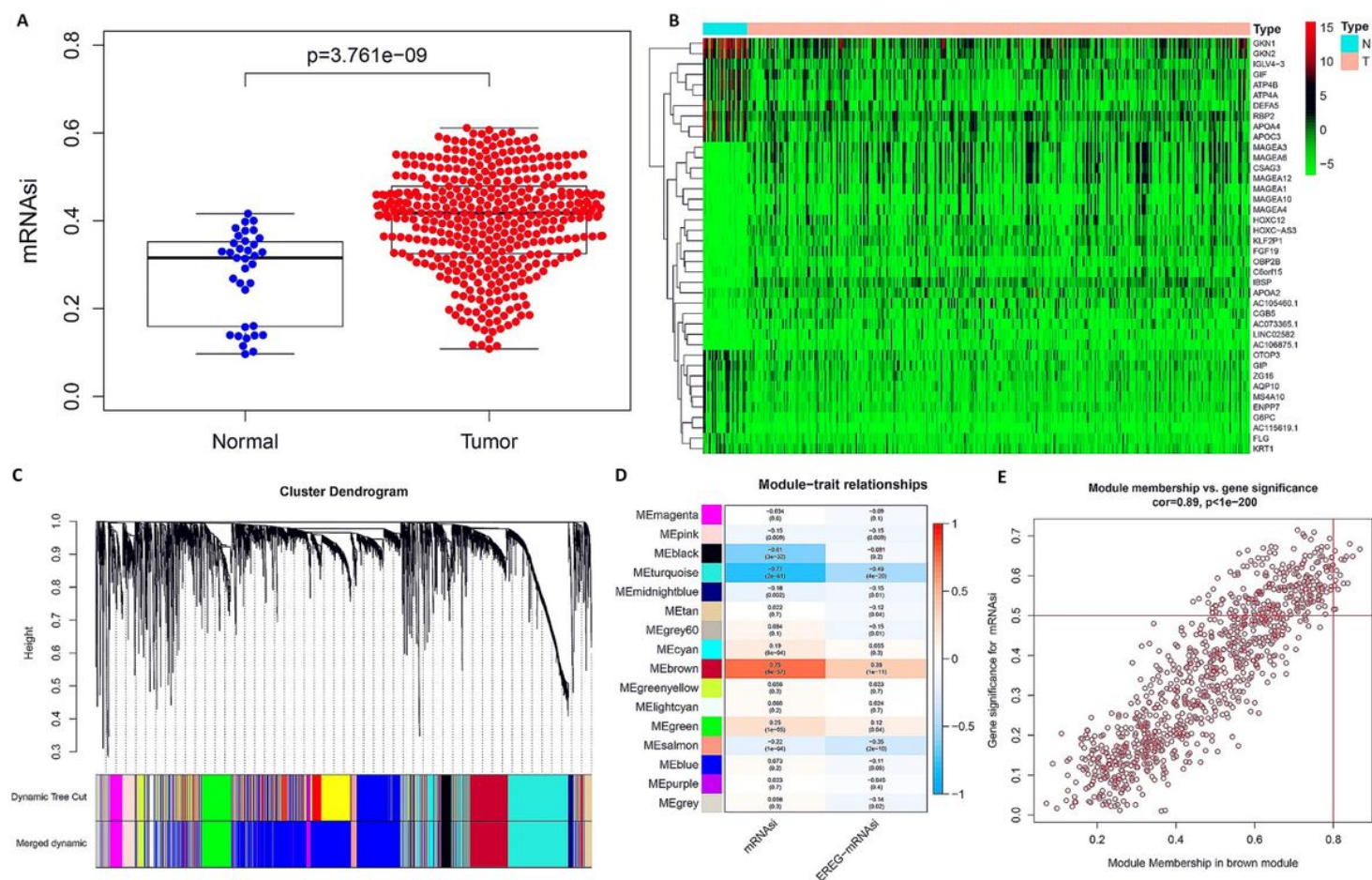


Figure 1

(A) Expression of mRNA in gastric tumor tissues and normal gastric tissues. (B) Significant differentially expressed genes in gastric tumor tissues and normal gastric tissues. (C) Clustering dendrograms of differentially expressed genes. Each piece of the leaves on the cluster dendrogram corresponds to a gene. The first color band indicates the modules detected by dynamic tree cut methods. The second color band indicates the modules after merging similar modules. (D) Correlation between the gene module and mRNA. The correlation coefficient in each cell represents the correlation between the gene module and mRNA, which increases in size from blue to red. The corresponding P-value is also annotated. (E) Scatter plot of module eigengenes in brown. The top right corner is where the key genes are located.

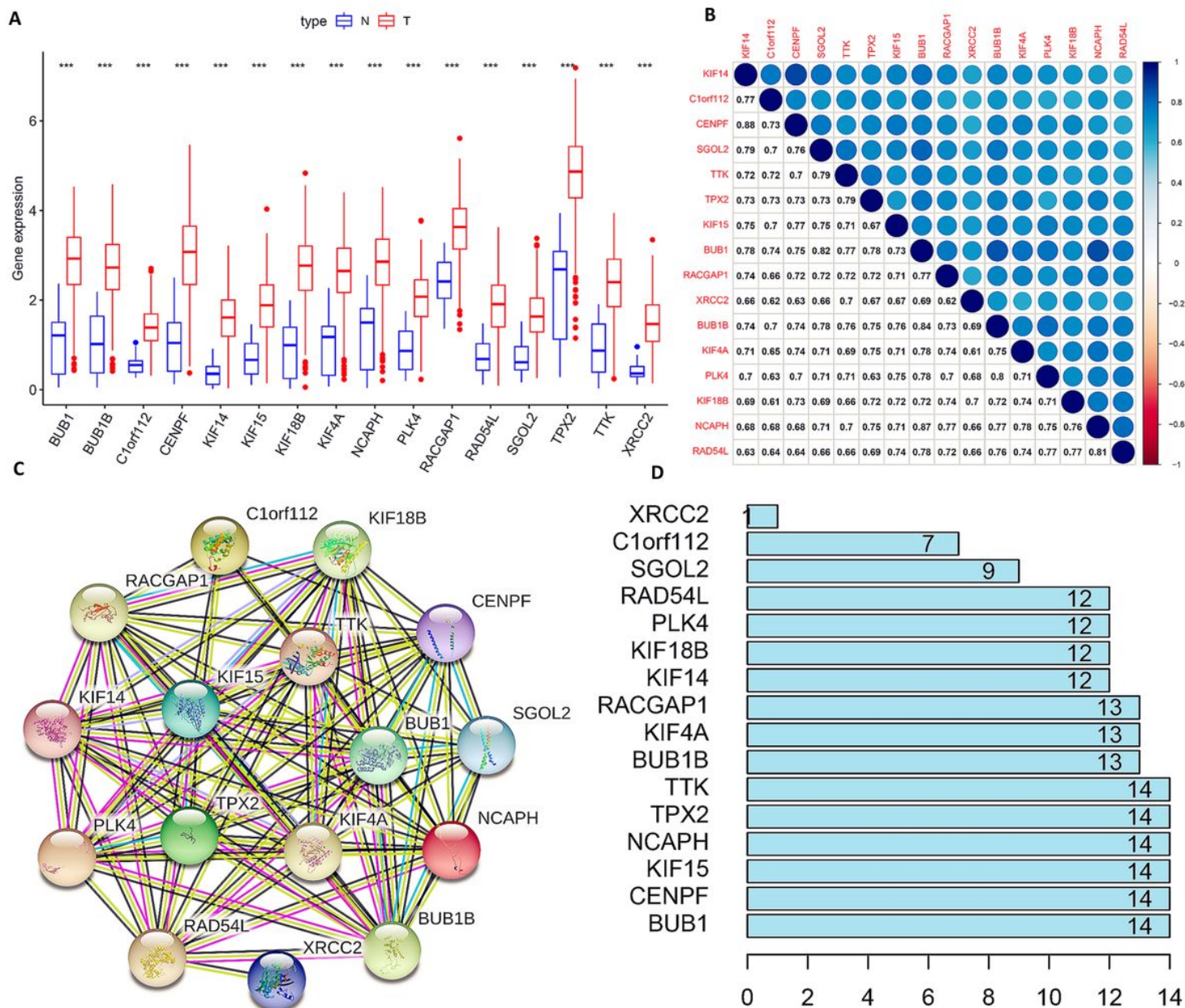


Figure 2

(A) Expression of the key genes in gastric tumor tissues and normal gastric tissues. The red histogram indicates gastric tumor tissue. The blue histogram indicates normal gastric tissue. *** $P < 0.001$. (B) Correlation of key genes. The number in the box represents the size of the correlation, and the larger the number, the higher the correlation. (C) Protein-protein interaction networks of key genes. The color of the lines represents the source of the evidence. (D) Connection nodes of key genes. The number represents the number of connection nodes.

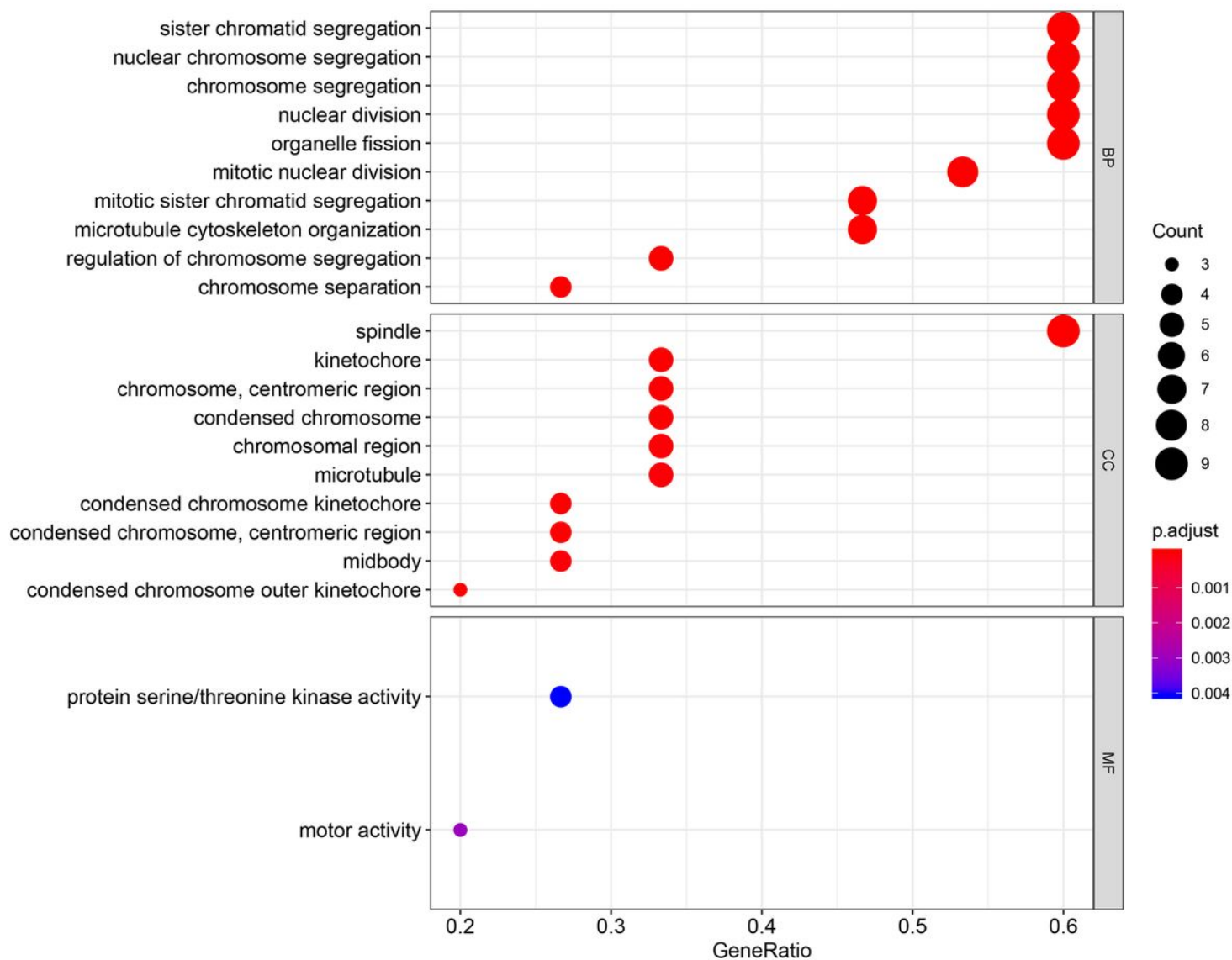


Figure 3

Significantly enriched GO biological process terms. BP indicates biological process, CC indicates cell component and MF indicates molecular function. The larger the circle, the more number of enrichment, and the redder the color, the more significant the difference.

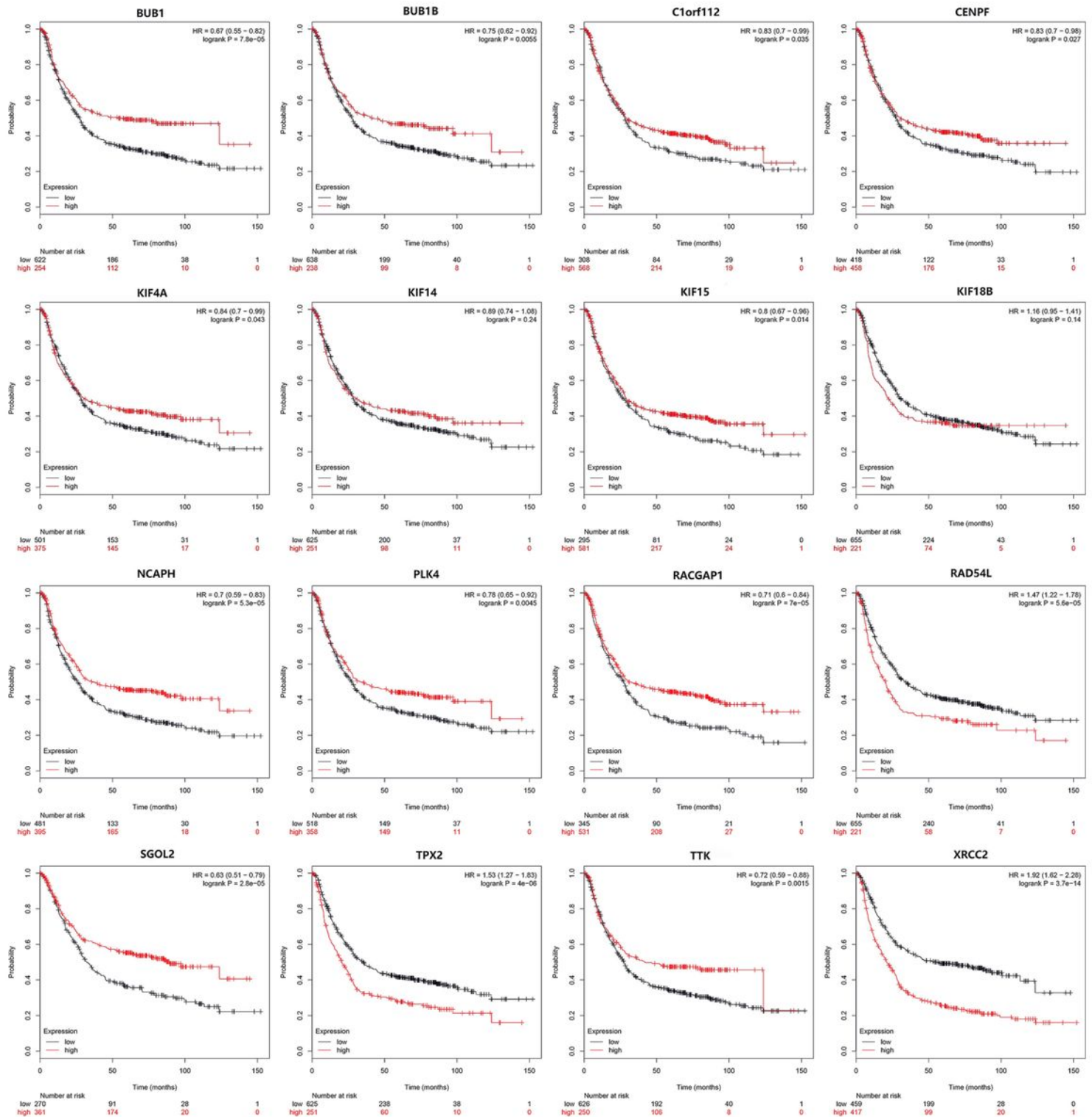


Figure 4

Survival curves of key genes using the online tool Kaplan-Meier Plotter. The red indicates high expression and black indicates low expression. A log-rank $P < 0.05$ was considered statistically significant.

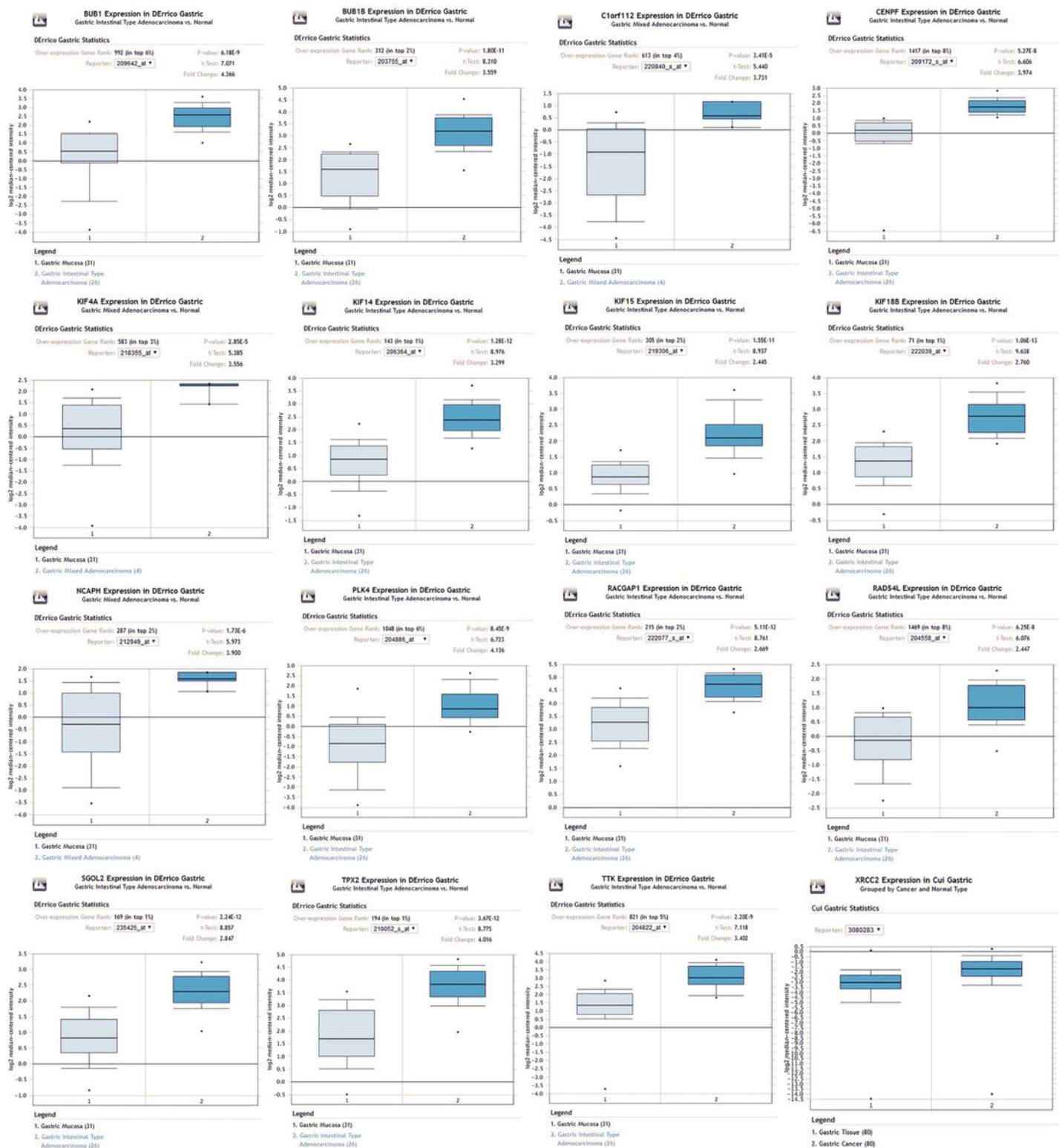


Figure 5

Expression of key genes in gastric cancer based on the Oncomine online database. The key genes were highly expressed in gastric cancer tissues.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTable1.doc](#)