Society of Medical Education & Research

2024, Volume 4, Issue 1, Page No: 24-36
ISSN: 3108-4834

Archive of International Journal of Cancer and Allied Science

Bioinformatics Analysis of Gene Modules and Key Genes for the Early Diagnosis of Gastric Cancer

Ling Xu[1], Jinyan Yang[1], Yu Zhang[1], Xudong Liu[2,3], Zhe Liu[4], Feipeng Sun[1], Ying Ma[1], Lei Wang[1], Feng Xing[1*]

[1]College of Life Sciences, Xinyang Normal University, Xinyang 464000, China.
[2]School of Medicine, Chongqing University, Chongqing 400044, China.
[3]Chongqing Key Laboratory of Translational Research for Cancer Metastasis and Individualized Treatment, Chongqing University Cancer Hospital, Chongqing 400030, China.
[4]Department of Computer Science, City University of Hong Kong, Hong Kong, China.

***E-mail** ✉ xfengr@163.com

Abstract

Gastric cancer (GC) remains one of the most prevalent and fatal cancers worldwide, with high rates of both incidence and mortality. Unfortunately, most patients are diagnosed at advanced stages, often through routine screenings, missing the best time for intervention. In this research, weighted gene co-expression network analysis (WGCNA) was used to identify key gene modules and hub genes associated with GC. Using the "limma" package in R, differentially expressed genes (DEGs) were examined from TCGA's GC dataset, resulting in the identification of 4892 DEGs. To better understand their biological roles, Gene ontology (GO) enrichment and KEGG pathway analysis were performed, which revealed that the DEGs were strongly involved in processes such as extracellular matrix organization, DNA replication, the cell cycle, and the p53 signaling pathway. WGCNA was also applied to identify gene modules associated with clinical characteristics in both GC and normal tissue samples. Six distinct gene modules were identified, with two showing significant association with GC. Hub genes in these modules were determined through survival and expression analyses. In addition, one-way ANOVA was used to examine how these hub genes were expressed across different stages of GC compared to normal tissues. This analysis revealed 19 genes with significant differential expression and positive prognosis implications. These hub genes showed significant differences in expression levels between normal and different GC stages, highlighting their potential as biomarkers for early detection of GC and as valuable tools to aid in diagnosis at earlier stages.

Keywords: Early diagnosis, Gastric cancer, Differentially expressed genes, Bioinformatics, Weighted gene co-expression network analysis

Introduction

Gastric cancer (GC), a type of cancer that originates from the stomach's epithelial cells or its surface glands, continues to be one of the most common and lethal cancers worldwide [1]. According to the 2020 GLOBOCAN report, GC is responsible for being the fifth most prevalent cause of cancer diagnoses and the fourth leading cause of cancer-related deaths globally [2]. This disease arises from multiple factors, such as diet, *Helicobacter pylori* infections, obesity, and genetics, all of which contribute to its development [3, 4]. GC is associated with high rates of metastasis, incidence, and mortality. However, early detection is often missed, leading to low rates of survival and the inability to perform radical surgical treatments. Furthermore, GC incidence and mortality are rising, particularly among younger populations [5-7]. Since early-stage GC tends to be asymptomatic, a majority of patients are diagnosed only at later stages of the disease, which results in more limited treatment options. Around 70% of cases are

Access this article online https://smerpub.com/

Received: 05 February 2024; **Accepted:** 04 May 2024

How to cite this article: Xu L, Yang J, Zhang Y, Liu X, Liu Z, Sun F, et al. Bioinformatics Analysis of Gene Modules and Key Genes for the Early Diagnosis of Gastric Cancer. Arch Int J Cancer Allied Sci. 2024;4(1):24-36. https://doi.org/10.51847/2gUBplgIMV

discovered after they have advanced beyond the point of early intervention [6]. Therefore, early diagnosis and treatment are crucial, which calls for the identification of biomarkers that can improve early detection and patient outcomes.

Weighted gene co-expression network analysis (WGCNA) is a powerful method for analyzing gene expression data across different samples. This technique groups genes with similar expression profiles into modules and links them with clinical traits, allowing the identification of important genes within those modules [8-10]. WGCNA has been successfully used in various cancer studies to pinpoint key genes involved in the progression and prognosis of cancers, including breast cancer, colorectal cancer, and renal cell carcinoma, aiding in the development of new diagnostic and therapeutic strategies [11-13].

In this research, we combined WGCNA with other analytical tools to examine clinical and RNA sequencing data from gastric cancer patients available in the TCGA database. We aimed to identify critical genes that are linked to clinical factors such as disease status (tumor_normal), gender, and pathologic features (T stage, stage, vital status, and initial weight). The results identified several potential hub genes that could serve as biomarkers or therapeutic targets for early-stage GC, providing a theoretical framework for improving diagnostic and therapeutic practices for this challenging disease.

Materials and Methods

Data acquisition and processing

RNA sequencing profiles and associated clinical data were retrieved from the TCGA repository via the UCSC Xena platform (https://tcga.xenahubs.net), which is hosted by the University of California, Santa Cruz (http://xena.ucsc.edu/) [14]. This dataset comprised 380 gastric cancer (GC) tissue samples alongside 37 normal tissue samples. To evaluate the overall data distribution and structure, we employed principal component analysis (PCA) as an initial screening method [15-19]. Outliers and inconsistent samples identified through PCA were removed to enhance the robustness and accuracy of subsequent analyses **(Figure 1)**.

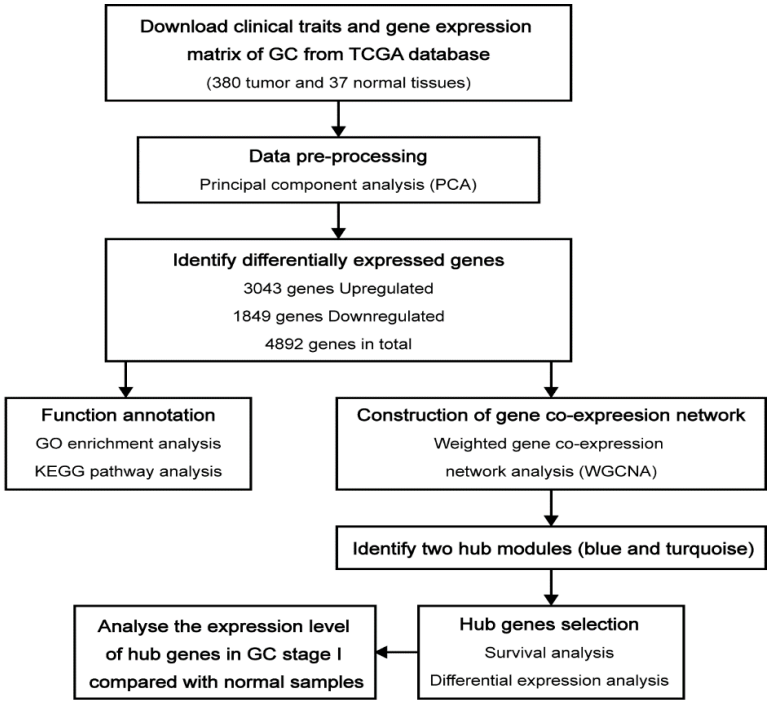


Figure 1. Workflow of searching hub genes in GC.

Identification of clinically significant modules

The construction of a gene co-expression network from the differentially expressed genes (DEGs) was carried out using the "WGCNA" R package (version 1.70-3) [8, 10]. The process began with the application of the flashClust tool in R, which was used to perform

hierarchical clustering on the samples and to identify any potential outliers. Next, the Pearson correlation coefficient (PCC) was computed for each gene pair to evaluate the level of co-expression. To ensure that the resulting network adhered to a scale-free topology, an optimal soft threshold power (β) was determined using the pickSoftThreshold function. Once β was selected, a weighted adjacency matrix (AM) was generated using a power transformation, where the connection strength (a_{mn}) between gene m and gene n was calculated by raising the absolute PCC (c_{mn}) to the β power, as outlined in Eq 1:

$$a_{mn} = |c_{mn}|^{\beta} \quad (1)$$

In this context, a_{mn} represents the degree of adjacency between two genes, c_{mn} is the Pearson correlation coefficient between them, and β denotes the soft thresholding power. The resulting AM was then converted into a topological overlap matrix (TOM), which was derived using the adjacency function as shown in Eq 2. In this formulation, l_{mn} is defined as the cumulative sum of the products of adjacency values between shared neighboring genes of m and n, while km and k_n represent the total connectivity of gene m and gene n with all other nodes in the network. The TOM was computed using the following equation:

$$TOM_{mn} = \frac{l_{mn}+a_{mn}}{\min(k_m+k_n)+1-a_{mn}} \quad (2)$$

To group genes with comparable expression characteristics into distinct modules, average linkage hierarchical clustering was carried out using a dissimilarity metric derived from the topological overlap matrix, calculated as (1 − TOM). A minimum module size was defined, allowing no fewer than 30 genes per module, while a merging threshold of 0.25 was applied to combine modules with highly similar expression patterns. The module eigengene (ME), defined as the leading principal component of each module, was used to summarize the overall expression trend within that module. To identify modules most strongly associated with the clinical phenotype, Pearson's correlation coefficient was computed between the ME values and the corresponding phenotype data, enabling the detection of modules with the highest degree of relevance to the clinical traits under investigation.

Hub gene selection

To pinpoint biologically meaningful hub genes, two core metrics were applied: gene significance (GS), indicating the strength of association between genes and clinical traits, and module membership (MM), representing the alignment between a gene and the rest of its module. Only genes exceeding specific GS and MM thresholds were considered for further investigation. To explore their biological relevance, analyses were carried out using GEPIA (http://gepia.cancer-pku.cn) [20], where both survival impact and expression variation were assessed. These evaluations focused on whether the selected genes influenced survival outcomes in gastric cancer (GC) patients and if their expression levels differed between early-stage tumors and normal tissues. Significance thresholds were set at $P < 0.05$ for survival analysis and $P < 0.01$ for evaluating expression differences.

Evaluation of the diagnostic significance of hub genes

A total of 241 TCGA-derived samples were utilized to examine the diagnostic potential of the hub genes. This analysis involved comparing gene expression in normal tissues with that across various GC pathological stages, taking into account clinical phenotype data. One-way ANOVA, performed using GraphPad Prism 8 software [21], was used to determine whether expression patterns of the hub genes varied significantly across disease stages.

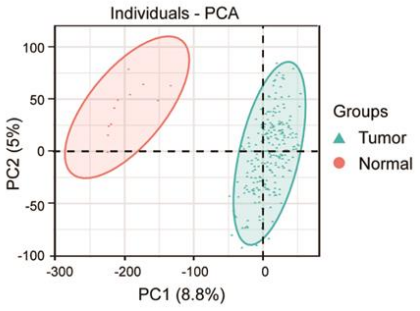
Results and Discussion

Pre-processing of data and DEG selection

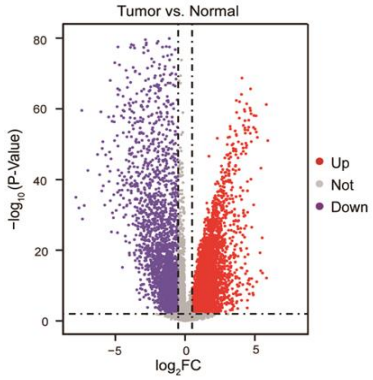
Following the exclusion of 50 GC tumor samples and 26 normal samples that could not be properly classified, the remaining samples were classified into GC tumor and normal groups using principal component analysis (PCA). The analysis showed that PC1 and PC2 explained 8.8% and 5% of the total variance, respectively **(Figures 2a and 2b)**. A total of 4,892 genes were identified as differentially expressed between the GC and normal samples, with 3,043 genes being upregulated and 1,849 downregulated **(Figures 2c and 2d)**.

	Sample numbers	
	Tumor	Normal
Before processing	380	37
After processing	230	11

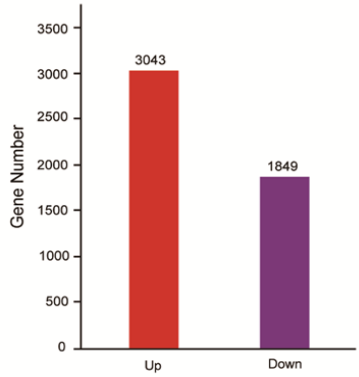
a)



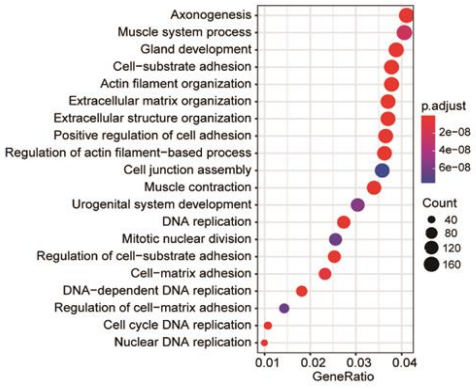
b)



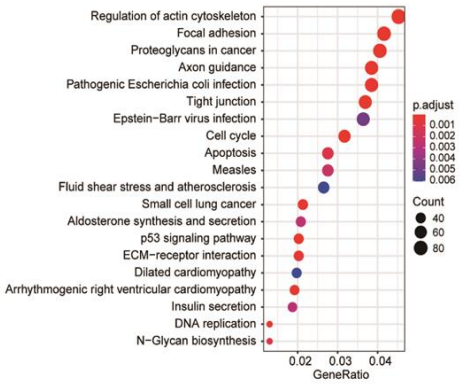
c)



d)



e)



f)

Figure 2. DEGs identification and functional enrichment analysis in GC; (a) number of samples included in the study, (b) principal component analysis (PCA) results, (c) volcano plot representation: upregulated genes are marked with red dots, downregulated genes with purple dots, and genes with no significant change are in gray, (d) total count of differentially expressed genes (DEGs), (e) gene ontology (GO) enrichment analysis, (f) KEGG pathway enrichment analysis; the color of each dot represents the significance level of the enrichment, while the size corresponds to the number of genes involved in each enriched term, with the top 20 results listed based on the P-value

Gene ontology (GO) enrichment and KEGG pathway analysis of DEGs

To explore the potential biological roles of the differentially expressed genes (DEGs) in gastric cancer (GC), we performed GO and KEGG pathway enrichment analyses. The GO analysis identified several biological processes consistently observed in prior research, including extracellular matrix organization, DNA replication, and the cell cycle, particularly DNA replication and nuclear DNA replication **(Figure 2e)** [22, 23]. KEGG pathway analysis revealed several enriched pathways such as focal adhesion, proteoglycans in cancer, the impact of *E. coli* infection, apoptosis, cell cycle regulation, small cell lung cancer, p53 signaling, and ECM-receptor interactions **(Figure 2f)** [23, 24].

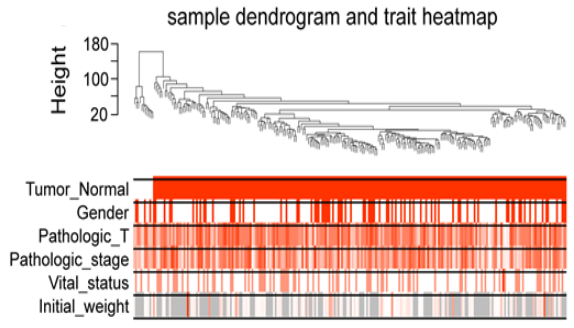
WGCNA and identification of key modules

For the construction of a co-expression network, WGCNA was applied to the tumor and normal samples. The clustering of these samples resulted in distinct

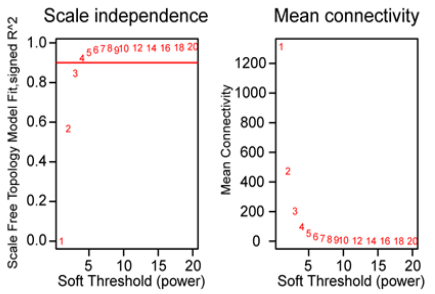
branches, with no outliers identified **(Figure 3a)**. Using a scale-free topology criterion of 0.9, a soft threshold (β) of 4 was selected **(Figure 3b)**. Six different gene modules were identified: blue, green, yellow, brown, turquoise, and grey. The grey module, consisting of genes that did not fit well into any of the other groups, was excluded from further analysis **(Figures 3c and 3d)**. Correlation analysis between these modules and clinical phenotypes revealed that the turquoise module (cor = 0.93, P = 8e-107) had a strong positive correlation with GC, while the blue module (cor = -0.62, P = 6e-27) exhibited a negative correlation. These findings indicate that the turquoise module plays a significant role in GC development, while the blue module may offer protective effects against the disease. As a result, the turquoise and blue modules were selected for subsequent investigation.

Hub gene identification and validation

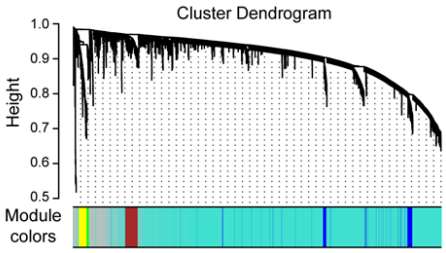
We examined the relationship between module membership (MM) and gene significance (GS) for both the blue and turquoise modules, revealing a strong positive correlation **(Figures 3e and 3f)**. To identify hub genes, thresholds of |MM| > 0.8 and |GS| > 0.6 were set for the blue module, and |MM| > 0.8 and |GS| > 0.8 for the turquoise module. This process identified 89 genes in the blue module and 216 genes in the turquoise module. Survival analysis and comparison of gene expression levels in GC tissues versus normal tissues led to the identification of 19 hub genes (ASF1B, DPT, ZBTB16, WISP2, PRIMA1, EPCAM, PDZD4, ATP1A2, FAM83H, ABCA9, C8orf46, MAMDC2, TCEAL2, CEP55, LIMS2, LMOD1, PLP1, TMEM100, ADHFE1). These hub genes were found to be significantly associated with GC prognosis, and expression differences were observed between GC and normal tissues **(Figures 4 and 5)**.



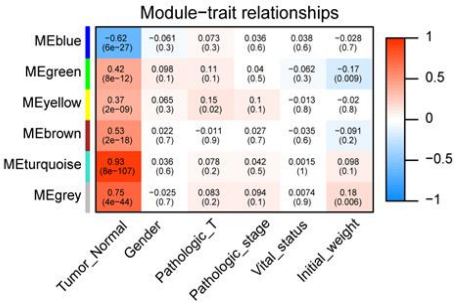
a)



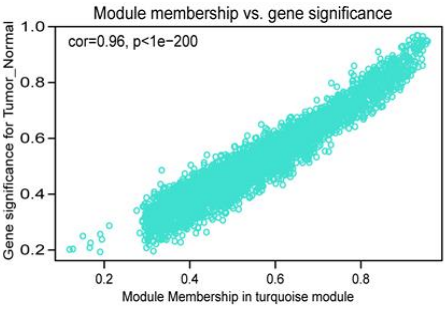
b)



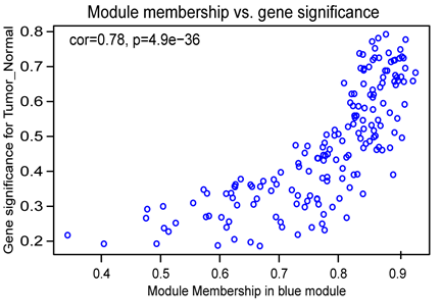
c)



d)



e)



f)

Figure 3. WGCNA analysis of DEGs in gastric

cancer (GC); (a) hierarchical clustering of GC and normal tissue samples based on clinical data, (b) analysis of the scale-free topology fit index and average gene connectivity across various soft threshold values (β), (c) dendrogram representing the clustering of DEGs based on the dissimilarity measure (1-TOM), (d) a heatmap showing the relationship between module eigengenes (ME) and different clinical traits in GC, (e, f) scatter plots illustrating the association between module membership (MM) and gene significance (GS) for genes in the turquoise (e) and blue (f) modules.

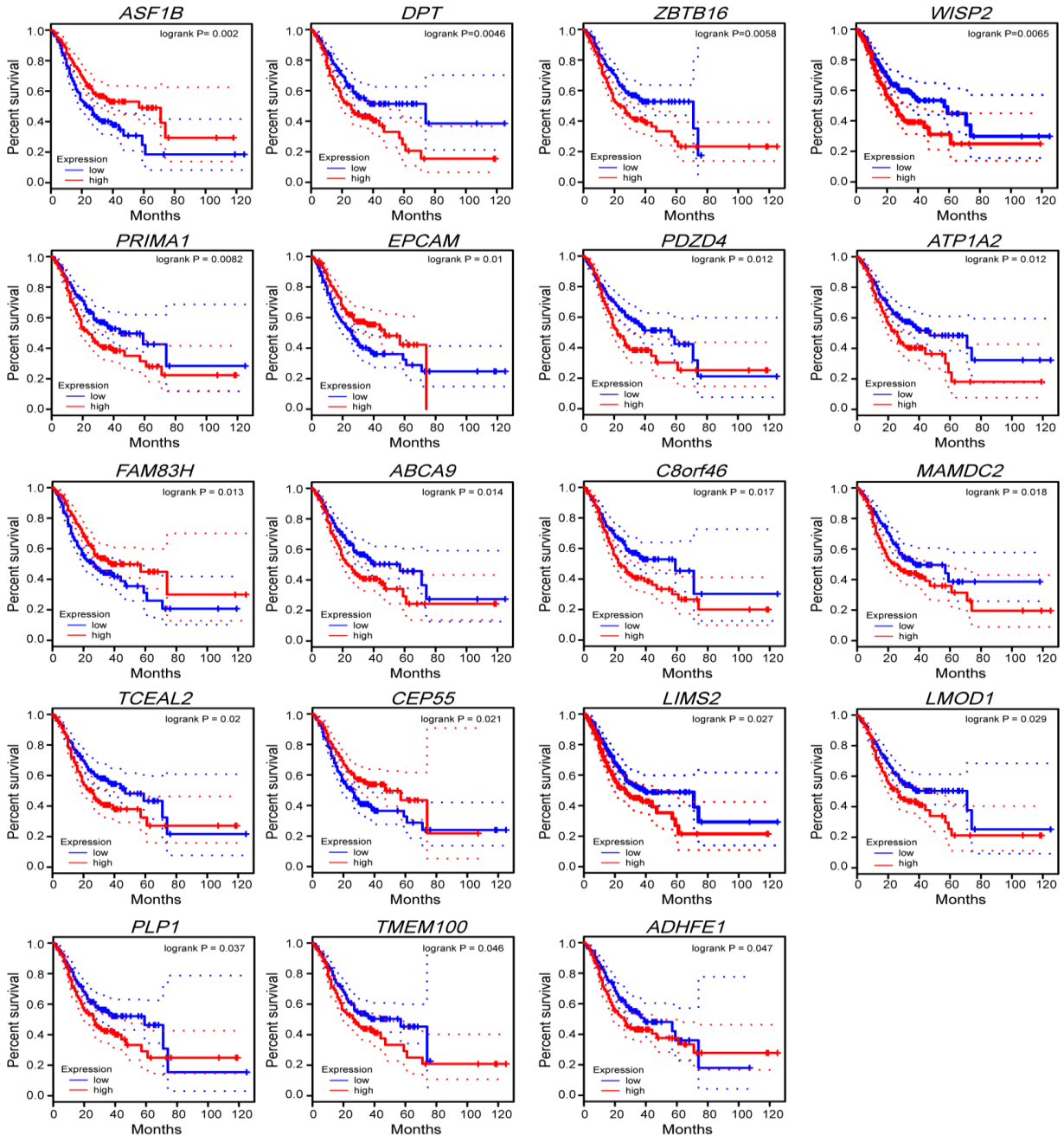


Figure 4. Survival analysis of hub genes in gastric cancer (GC)

This analysis examines the survival outcomes associated with 19 hub genes, showing a significant correlation with patient survival. The red line represents the high-expression group of each gene, while the blue line

corresponds to the low-expression group. A P-value threshold of < 0.05 was considered statistically significant.

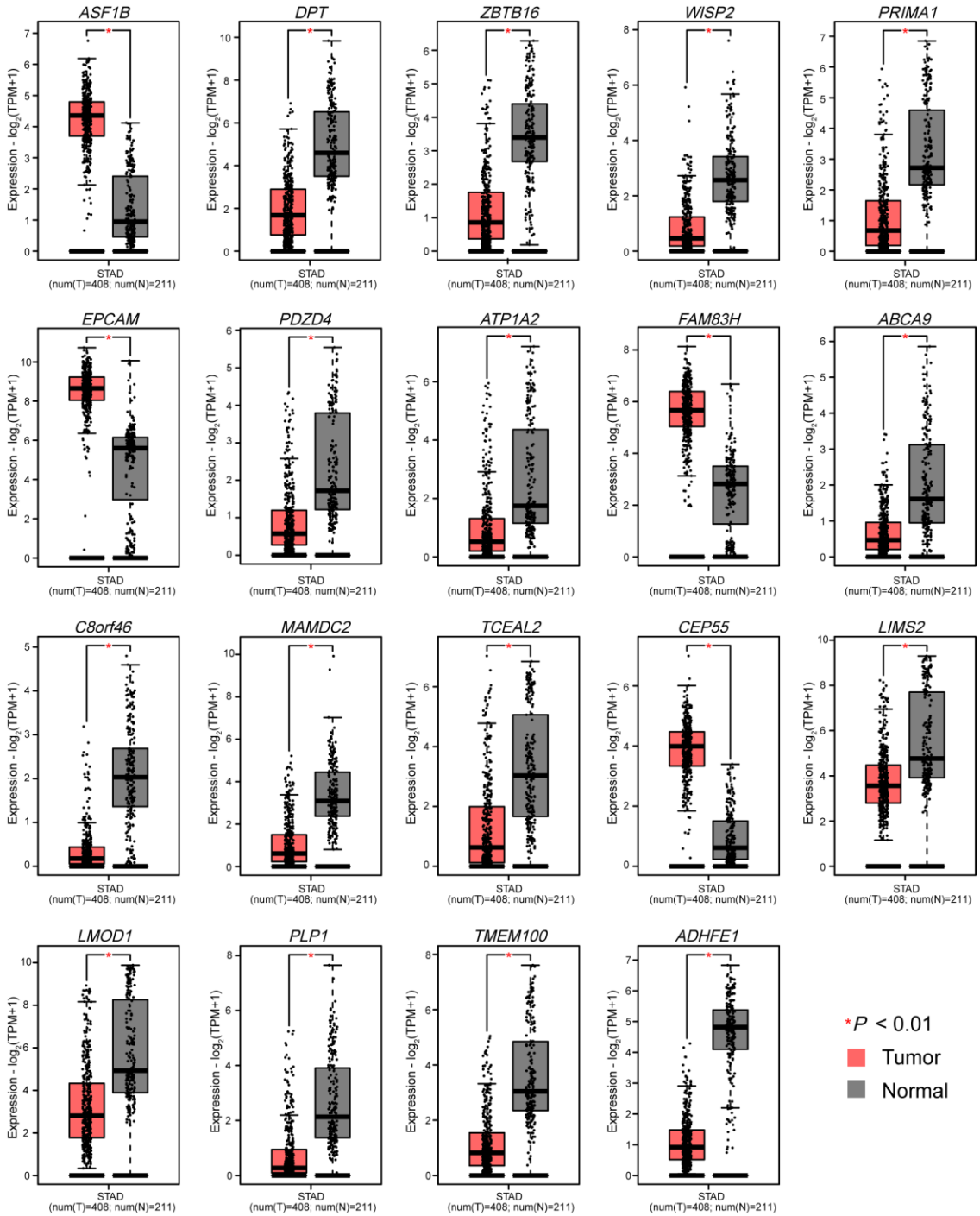


Figure 5. Comparison of hub gene expression between GC and normal tissues

The expression profiles of hub genes were evaluated using the GEPIA database, focusing on their levels in both GC and normal tissues. A significance level of *P < 0.01 was applied to determine statistically relevant differences.

Early detection markers for gastric cancer (GC)

Upon assessing the expression levels of the 19 identified hub genes at different stages of GC, we found that these genes exhibited notable differences between normal gastric tissues and those affected by early-stage GC. Specifically, genes such as ASF1B, FAM83H, EPCAM, and CEP55 had higher expression levels in early GC tissues compared to normal gastric tissue, whereas the expression of other genes was notably reduced in the early stages of cancer. These observations suggest that these hub genes may serve as crucial markers for the early identification of GC **(Figure 6)** [25]. ASF1B and CEP55, both previously implicated in cancer development, mitosis, and cytokinesis, were barely detectable in normal gastric tissues but showed elevated levels in tumor tissues, even at early stages. Conversely, genes such as PDZD4, ABCA9, ATP1A2, C8orf46, and TCEAL2 demonstrated high expression in normal gastric tissues but almost negligible expression in early GC **(Table 1; Figure 6)**. The combination of these 19 hub genes—particularly ASF1B, CEP55, PDZD4, ATP1A2, ABCA9, C8orf46, and TCEAL2—offers substantial promise as biomarkers for the early detection of GC.

Table 1. Functional roles of the 19 hub genes

Gene	Function	Source PMIDs#
ASF1B	Involved in the progression of various cancers.	35362843; 21179005
DPT	A non-collagenous component of the extracellular matrix that modulates tumor cell growth and invasiveness.	30391671; 25149533; 21796630
ZBTB16	Functions as a transcriptional repressor, inhibiting the proliferation and spread of cancer cells.	10688654; 24359566; 29358655; 24339862; 32517789; 30431129
WISP2	Exhibits bidirectional effects on tumor cell regulation.	34385183; 30808397; 32711570
PRIMA1	Binds to AChE, anchoring it to neural cell membranes.	11804574
EPCAM	Directly regulates the cell cycle and proliferation, and enhances the expression of proto-oncogene c-myc and cyclins A/E.	15195135
PDZD4	A novel gene containing a PDZ domain contributes to tumor cell proliferation.	15077175
ATP1A2	Catalytic subunit of an enzyme that hydrolyzes ATP, facilitating sodium and potassium ion exchange across the plasma membrane.	33880529
FAM83H	Controls the migration of epithelial cells.	23902688
ABCA9	A transporter involved in monocyte differentiation and lipid transport regulation	12150964
C8orf46	Participates in neurogenesis.	32558188
MAMDC2	Suppresses tumor cell activity.	32707597
TCEAL2	Functions as a tumor suppressor.	33061644
CEP55	Critical for mitotic exit and cytokinesis.	16198290; 17853893
LIMS2	Regulates the migration and spreading of tumor cells.	16959213
LMOD1	Enhances the migration of tumor cells.	35488236
PLP1	Essential for the formation and maintenance of myelin structure.	30094605
TMEM100	Inhibits metastasis of tumor cells.	34687431; 31188741
ADHFE1	Associated with tumor cell proliferation and embryonic development.	16959974; 23517143; 29202474; 24886599

#PMID: PubMed unique identifier

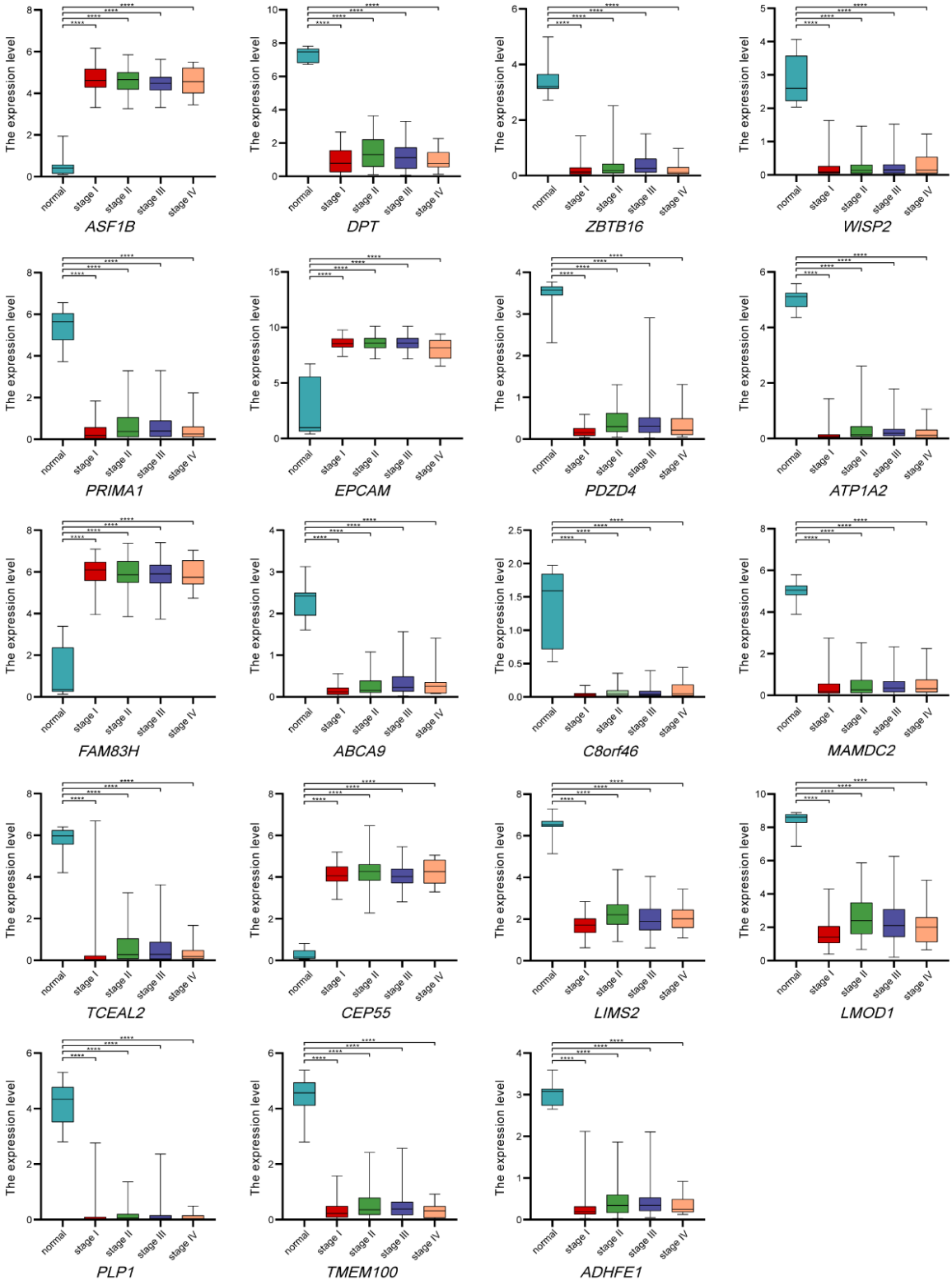


Figure 6. Association between hub genes expression and tumor stage; **** was considered as P < 0.0001.

Gastric cancer (GC), a leading cause of cancer-related deaths, remains challenging to treat, particularly in its advanced stages, with many therapeutic issues still unresolved [26]. Thus, the identification of new

biomarkers and hub genes for early-stage GC is critical. In our study, we identified 4892 differentially expressed genes (DEGs). The results of gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses indicated that these DEGs were predominantly involved in processes such as extracellular matrix (ECM) organization, DNA replication, the cell cycle, and the p53 signaling pathway. The cell cycle, frequently disrupted in cancer, represents a potential target for therapeutic intervention in GC [27, 28]. A deeper understanding of how the cell cycle pathway contributes to GC's initiation and progression may provide new insights into treatment strategies. DNA replication, crucial for cell division, when altered by replication errors or mutations, can lead to genomic instability, chromosomal abnormalities, and cancer progression [29]. The ECM plays a significant role in cancer metastasis and the development of tumors [30], and the p53 signaling pathway is essential for maintaining genomic integrity by regulating cell cycle progression in response to DNA damage [31]. Consequently, we hypothesize that these DEGs could be involved in GC progression and impact prognosis via the p53 pathway, contributing to the poor survival outcomes associated with GC.

Through weighted gene co-expression network analysis (WGCNA), we identified six distinct gene co-expression modules, with the blue and turquoise modules showing a strong correlation with GC (tumor_normal). Following further analysis, we identified 19 hub genes through survival and differential expression analysis, which suggests these genes may play key roles in the diagnosis and treatment of GC.

Early cancer detection has immense clinical value, and we aimed to address this challenge by evaluating the expression patterns of the 19 hub genes across different stages of GC and in normal tissue. Our findings demonstrated that ASF1B, EPCAM, FAM83H, and CEP55 exhibited elevated expression levels in early GC compared to normal gastric tissues, while the remaining 14 hub genes had lower expression in early-stage GC. This pattern suggests that these genes may be valuable for early-stage GC diagnosis.

Several of the identified hub genes have been linked to other cancers. For instance, ASF1B has been associated with increased metastasis and poor prognosis in breast cancer [32]. DPT, involved in cell adhesion and invasiveness, is crucial for the progression of oral squamous cell carcinoma [33]. Additionally, genes like ZBTB16, MAMDC2, TCEAL2, and TMEM100 have been shown to inhibit tumor proliferation and metastasis in various cancers, including gallbladder, breast, renal, and non-small cell lung cancers [34-37]. PDZD4 and ADHFE1 have been implicated in the proliferation of synovial sarcoma and colorectal cancer cells, respectively [38, 39]. LIMS2 and LMOD1 have been shown to promote GC cell migration [40, 41]. WISP2, exhibiting bidirectional effects, influences the proliferation of tumor cells in different cancers, including esophageal and ovarian cancer [42, 43]. The genes LIMS2, LMOD1, TCEAL2, TMEM100, and ZBTB16 have also been highlighted for their involvement in GC development [40, 41, 44-46]. This supports the idea that these hub genes could serve as key biomarkers for early GC detection.

The early identification, diagnosis, and treatment of GC are essential for improving survival rates and enhancing the quality of life for patients. Our study provides valuable insights into the potential use of these hub genes as biomarkers and therapeutic targets for early GC. Their identification holds significant promise for advancing clinical approaches in GC diagnosis, treatment, and prognosis. Nevertheless, this study has limitations, notably the absence of additional experimental validation to further explore the functional roles of these hub genes in GC.

Conclusion

Through WGCNA analysis, we identified six distinct co-expression network modules and proceeded to isolate genes from the significant modules. Utilizing the GEPIA database, we validated 19 hub genes, which were further confirmed as key biomarkers for early GC. These 19 hub genes present a promising tool for early detection and treatment of GC, offering the potential to lower mortality rates among GC patients and providing a solid foundation for improving prediction, diagnosis, and therapeutic strategies for early-stage GC.

Acknowledgments: None

Conflict of Interest: None

Financial Support: This research was funded by the Nanhu Scholars Program for Young Scholars of XYNU.

Ethics Statement: None

References

1. Karimi P, Islami F, Anandasabapathy S, Freedman ND, Kamangar F. Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention. Cancer Epidemiol Biomarkers Prev. 2014;23(5):700-13. Available from: https://pubmed.ncbi.nlm.nih.gov/24618998
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209-49. Available from: https://pubmed.ncbi.nlm.nih.gov/33538338
3. Forman D, Burley VJ. Gastric cancer: global pattern of the disease and an overview of environmental risk factors. Best Pract Res Clin Gastroenterol. 2006;20(4):633-49. Available from: https://pubmed.ncbi.nlm.nih.gov/16997150
4. Yusefi AR, Lankarani KB, Bastani P, Radinmanesh M, Kavosi Z. Risk factors for gastric cancer: a systematic review. Asian Pac J Cancer Prev. 2018;19(3):591-603. Available from: https://pubmed.ncbi.nlm.nih.gov/29579788
5. Wu H, Wang W, Tong S, Wu C. Nucleostemin regulates proliferation and migration of gastric cancer and correlates with its malignancy. Int J Clin Exp Med. 2017;10(11):15800. Available from: https://pubmed.ncbi.nlm.nih.gov/29369305
6. Song Z, Wu Y, Yang J, Yang D, Fang X. Progress in the treatment of advanced gastric cancer. Tumour Biol. 2017;39(7):1010428317714626. Available from: https://pubmed.ncbi.nlm.nih.gov/28671042
7. Sun Z, Wang Q, Yu X, Ou C, Yao L, Liu K, et al. Risk factors associated with splenic hilar lymph node metastasis in patients with advanced gastric cancer in northwest China. Int J Clin Exp Med. 2015;8(11):21358-64. Available from: https://pubmed.ncbi.nlm.nih.gov/26885077
8. Yang L, Ma TJ, Zhang YB, Wang H, An RH. Construction and analysis of lncRNA-miRNA-mRNA ceRNA network identify an eight-gene signature as a potential prognostic factor in kidney renal papillary cell carcinoma (KIRP). Altern Ther Health Med. 2022;28(6):42-51. Available from: https://pubmed.ncbi.nlm.nih.gov/35648698
9. Van Dam S, Võsa U, Van der Graaf A, Franke L, de Magalhães JP. Gene co-expression analysis for functional classification and gene-disease predictions. Brief Bioinform. 2018;19(4):575-92. Available from: https://pubmed.ncbi.nlm.nih.gov/28077403
10. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol. 2005;4:Article17. Available from: https://pubmed.ncbi.nlm.nih.gov/16646834 doi:10.2202/1544-6115.1128
11. Jin H, Huang X, Shao K, Li G, Wang J, Yang H, et al. Integrated bioinformatics analysis to identify 15 hub genes in breast cancer. Oncol Lett. 2019;18(2):1023-34. Available from: https://pubmed.ncbi.nlm.nih.gov/31423162
12. Yuan L, Chen L, Qian K, Qian G, Wu CL, Wang X, et al. Co-expression network analysis identified six hub genes in association with progression and prognosis in human clear cell renal cell carcinoma (ccRCC). Genom Data. 2017;14:132-40. Available from: https://pubmed.ncbi.nlm.nih.gov/29159069
13. Zhang Y, Luo J, Liu Z, Liu X, Ma Y, Zhang B, et al. Identification of hub genes in colorectal cancer based on weighted gene co-expression network analysis and clinical data from the cancer genome atlas. Biosci Rep. 2021;41(7):BSR20211280. Available from: https://pubmed.ncbi.nlm.nih.gov/34308980
14. Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, et al. Visualizing and interpreting cancer genomics data via the Xena platform. Nat Biotechnol. 2020;38(6):675-8. Available from: https://pubmed.ncbi.nlm.nih.gov/32444850
15. Ma S, Dai Y. Principal component analysis based methods in bioinformatics studies. Brief Bioinform. 2011;12(6):714-22. Available from: https://pubmed.ncbi.nlm.nih.gov/21242203
16. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47. Available from: https://pubmed.ncbi.nlm.nih.gov/25605792
17. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: a tool for the unification of biology. The gene ontology consortium. Nat Genet. 2000;25(1):25-9. Available from: https://pubmed.ncbi.nlm.nih.gov/10802651
18. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on

genomes, pathways, diseases, and drugs. Nucleic Acids Res. 2017;28(1):27-30. Available from: https://pubmed.ncbi.nlm.nih.gov/27899662
19. Yu G, Wang LG, Han Y, He QY. Cluster Profiler: an R package for comparing biological themes among gene clusters. Omics. 2012;16(5):284-7. Available from: https://pubmed.ncbi.nlm.nih.gov/22455463
20. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res. 2017;45(W1):W98-102. Available from: https://pubmed.ncbi.nlm.nih.gov/28407145
21. Mishra P, Singh U, Pandey CM, Mishra P, Pandey G. Application of student's t-test, analysis of variance, and covariance. Ann Card Anaesth. 2019;22(4):407-11. Available from: https://pubmed.ncbi.nlm.nih.gov/31621677
22. Li Y, Gu J, Xu F, Zhu Q, Ge D, Lu C. Transcriptomic and functional network features of lung squamous cell carcinoma through integrative analysis of GEO and TCGA data. Sci Rep. 2018;8(1):15834. Available from: https://pubmed.ncbi.nlm.nih.gov/30367091
23. Yang D, He Y, Wu B, Deng Y, Wang N, Li M, et al. Integrated bioinformatics analysis for the screening of hub genes and therapeutic drugs in ovarian cancer. J Ovarian Res. 2020;13(1):10. Available from: https://pubmed.ncbi.nlm.nih.gov/31987036
24. Dai H, Guo L, Lin M, Cheng Z, Li J, Tang J, et al. Comprehensive analysis and identification of key genes and signaling pathways in the occurrence and metastasis of cutaneous melanoma. Peer J. 2020;8:e10265. Available from: https://pubmed.ncbi.nlm.nih.gov/33240619
25. Chen G, Luo D, Zhong N, Li D, Zheng J, Liao H, et al. GPC2 Is a potential diagnostic, immunological, and prognostic biomarker in pan-cancer. Front Immunol. 2022;13:857308. Available from: https://pubmed.ncbi.nlm.nih.gov/35345673
26. Chen W, Zheng R, Zeng H, Zhang S, He J. Annual report on the status of cancer in China, 2011. Chin J Cancer Res. 2015;27(1):2-12. Available from: https://pubmed.ncbi.nlm.nih.gov/25717220
27. Aarts M, Linardopoulos S, Turner NC. Tumour selective targeting of cell cycle kinases for cancer treatment. Curr Opin Pharmacol. 2013;13(4):529-35. Available from: https://pubmed.ncbi.nlm.nih.gov/23597425
28. Liu L, Michowski W, Kolodziejczyk A, Sicinski P. The cell cycle in stem cell proliferation, pluripotency, and differentiation. Nat Cell Biol. 2019;21(9):1060-7. Available from: https://pubmed.ncbi.nlm.nih.gov/31481793
29. Ekundayo B, Bleichert F. Origins of DNA replication. PLoS Genet. 2019;15(9):e1008320. Available from: https://pubmed.ncbi.nlm.nih.gov/31513569
30. Levental KR, Yu H, Kass L, Lakins JN, Egeblad M, Erler JT, et al. Matrix crosslinking forces tumor progression by enhancing integrin signaling. Cell. 2009;139(5):891-906. Available from: https://pubmed.ncbi.nlm.nih.gov/19931152
31. Lukin DJ, Carvajal LA, Liu WJ, Resnick-Silverman L, Manfredi JJ. p53 Promotes cell survival due to the reversibility of its cell-cycle checkpoints. Mol Cancer Res. 2015;13(1):16-28. Available from: https://pubmed.ncbi.nlm.nih.gov/25158956
32. Corpet A, De Koning L, Toedling J, Savignoni A, Berger F, Lemaître C, et al. Asf1b, the necessary Asf1 isoform for proliferation, is predictive of outcome in breast cancer. Embo J. 2011;30(3):480-93. Available from: https://pubmed.ncbi.nlm.nih.gov/21179005
33. Yamatoji M, Kasamatsu A, Kouzu Y, Koike H, Sakamoto Y, Ogawara K, et al. Dermatopontin: a potential predictor for metastasis of human oral cancer. Int J Cancer. 2012;130(12):2903-11. Available from: https://pubmed.ncbi.nlm.nih.gov/21796630
34. Lee H, Park BC, Soon Kang J, Cheon Y, Lee S, Jae Maeng P. MAM domain containing 2 is a potential breast cancer biomarker that exhibits tumour-suppressive activity. Cell Prolif. 2020;53(9):e12883. Available from: https://pubmed.ncbi.nlm.nih.gov/32707597
35. Shen H, Zhan M, Zhang Y, Huang S, Xu S, Huang X, et al. PLZF inhibits the proliferation and metastasis of gallbladder cancer by regulating IFIT2. Cell Death Dis. 2018;9(2):71. Available from: https://pubmed.ncbi.nlm.nih.gov/29358655
36. Wang Y, Ha M, Li M, Zhang L, Chen Y. Histone deacetylase 6-mediated downregulation of TMEM100 expedites the development and progression of non-small cell lung cancer. Hum Cell.

2022;35(1):271-85. Available from: https://pubmed.ncbi.nlm.nih.gov/34687431
37. Zhou Y, Zhang Y, Li W, Xu J, He X, Li X, et al. TCEAL2, as a tumor suppressor in renal cell carcinoma, is associated with the good prognosis of patients. Cancer Manag Res. 2020;12:9589-97. Available from: https://pubmed.ncbi.nlm.nih.gov/33061644
38. Nagayama S, Iiizumi M, Katagiri T, Toguchida J, Nakamura Y. Identification of PDZK4, a novel human gene with PDZ domains, that is up-regulated in synovial sarcomas. Oncogene. 2004;23(32):5551-7. Available from: https://pubmed.ncbi.nlm.nih.gov/15077175
39. Tae CH, Ryu KJ, Kim SH, Kim HC, Chun HK, Min BH, et al. Alcohol dehydrogenase, iron-containing, 1 promoter hypermethylation associated with colorectal cancer differentiation. BMC Cancer. 2013;13:142. Available from: https://pubmed.ncbi.nlm.nih.gov/23517143
40. Kim SK, Jang HR, Kim JH, Noh SM, Song KS, Kim MR, et al. The epigenetic silencing of LIMS2 in gastric cancer and its inhibitory effect on cell migration. Biochem Biophys Res Commun. 2006;349(3):1032-40. Available from: https://pubmed.ncbi.nlm.nih.gov/16959213
41. Tan Y, Chen Q, Pan S, An W, Xu H, Xing Y, et al. LMOD1, an oncogene associated with the Lauren classification, regulates the metastasis of gastric cancer cells through the FAK-AKT/mTOR pathway. BMC Cancer. 2022;22(1):474. Available from: https://pubmed.ncbi.nlm.nih.gov/35488236
42. Chai DM, Qin YZ, Wu SW, Ma L, Tan YY, Yong X, et al. WISP2 exhibits its potential anti-tumor activity by targeting ERK and E-cadherin pathways in esophageal cancer cells. J Exp Clin Cancer Res. 2019;38(1):102. Available from: https://pubmed.ncbi.nlm.nih.gov/30808397
43. Shi ZQ, Chen ZY, Han Y, Zhu HY, Lyu MD, Zhang H, et al. WISP2 promotes cell proliferation by targeting ERK and YAP in ovarian cancer cells. J Ovarian Res. 2020;13(1):85. Available from: https://pubmed.ncbi.nlm.nih.gov/32711570
44. Wang JB, Jin Y, Wu P, Liu Y, Zhao WJ, Chen JF, et al. Tumor suppressor PLZF regulated by lncRNA ANRIL suppresses proliferation and epithelial-mesenchymal transformation of gastric cancer cells. Oncol Rep. 2019;41(2):1007-18. Available from: https://pubmed.ncbi.nlm.nih.gov/30431129
45. Yu B, Dai W, Pang L, Sang Q, Li F, Yu J, et al. The dynamic alteration of transcriptional regulation by crucial TFs during tumorigenesis of gastric cancer. Mol Med. 2022;28(1):41. Available from: https://pubmed.ncbi.nlm.nih.gov/35421923
46. Zhuang J, Huang Y, Zheng W, Yang S, Zhu G, Wang J, et al. TMEM100 expression suppresses metastasis and enhances sensitivity to chemotherapy in gastric cancer. Biol Chem. 2020;401(2):285-96. Available from: https://pubmed.ncbi.nlm.nih.gov/31188741