



ORIGINAL ARTICLE OPEN ACCESS

Bioinformatics Analysis of the Expression and Prognostic Significance of Transcription Factor YY1 in Gastric Cancer

Wenliang Chen¹ | Huanhuan Wang² | Ntiak Achi³ | Jinjin Hao³ | Rui Gong³ | Qiang Zhao³

¹Department of General Surgery, The 2nd Affiliated Hospital of Shanxi Medical University, Taiyuan, Shanxi, China | ²Department of General Surgery, Jincheng People's Hospital, Jincheng, Shanxi, China | ³Graduate Department of Shanxi Medical University, Taiyuan, Shanxi, China

Correspondence: Wenliang Chen (wlchen@sxmu.edu.cn)

Received: 27 October 2024 | **Revised:** 26 February 2025 | **Accepted:** 5 March 2025

Funding: This study was supported by the Natural Science Foundation of Shanxi Province, China (201801D121320).

Keywords: bioinformatics analysis | gastric cancer | immune cell infiltration | Yin Yang 1

ABSTRACT

Background: Emerging evidence indicates that the transcription factor Yin Yang 1 (YY1) plays a critical role in the carcinogenesis and progression of various human malignancies. YY1 is highly expressed in gastric cancer (GC), raising interest in its role in GC.

Aims: This study aims to analyze the role of YY1 in gastric cancer, investigate its effect on the tumor microenvironment, and assess its potential as a prognostic marker.

Methods and Results: Transcriptomic data and clinical information from GC patients were obtained from the TCGA and UCSC databases. YY1 expression was analyzed using the R “limma” package. Gene ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses were performed with the online tool clusterProfiler. The relationship between YY1 expression levels and the tumor microenvironment was examined in different risk groups of GC patients. Additionally, YY1-positive staining in 26 clinical GC samples was measured using ImageJ software. Co-expression analysis was used to identify prognostic genes associated with YY1, and a prognostic risk model was built and optimized.

Results showed that YY1 was significantly overexpressed in 415 GC ($p < 0.001$) and was associated with poorer survival outcomes ($p = 0.043$). GO and KEGG showed that YY1 was involved in key biological processes of the disease. Higher YY1 expression was correlated with lower stromal and immune cell content in the tumor microenvironment. Immunohistochemical staining confirmed YY1 overexpression in GC tissues compared to normal tissues ($p = 0.0293$). Positive correlations were observed between YY1 and the genes MTA1, TTL15, HNRNPU, WDR20, and PPP4R3A. The prognostic model, which included genes significantly associated with YY1 (risk score AUC = 0.690), predicted patient survival better than other clinical variables.

Conclusion: These findings suggest that YY1 plays an important role in the development of GC. Targeting the YY1 pathway may be a potential treatment strategy for GC.

1 | Background

Gastric cancer (GC) is the fifth most prevalent cancer and the fourth leading cause of cancer-related mortality globally [1]. Clinically, the prognosis of GC patients is mainly assessed using the TNM staging system and postoperative pathological

grading. The TNM system evaluates tumor size and extent (T), lymph node involvement (N), and the presence of distant metastasis (M) [2]. Postoperative pathological grading further evaluates how differentiated and aggressive the tumor is, helping to predict patient outcomes and guide treatment decisions. However, due to the inherent heterogeneity of tumors, even

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Cancer Reports* published by Wiley Periodicals LLC.

patients with the same TNM stage and pathological grade may have very different survival outcomes. This highlights the complexity of GC and the need for better prognostic biomarkers and personalized treatment strategies to improve patient care and outcomes [3].

Yin Yang 1 (YY1), located on the q32 region of chromosome 14, is a widely expressed zinc finger protein that regulates gene expression [4]. YY1, also known by other names, such as *UCRBP*, *CF1*, δ , and *NF-E1*, is a member of the GLI-Krüppel family of transcription factors characterized by their zinc finger motifs [5, 6]. These motifs allow YY1 to bind DNA and regulate many target genes involved in cell growth, differentiation, and apoptosis [7–9].

Recent studies have shown that YY1 plays an important role in GC progression. YY1 promotes GC by increasing cell proliferation, invasion, and metastasis. Specifically, it binds directly to the promoter regions of *CCDC43* and *ADRM1*, causing their overexpression and enhancing GC proliferation and metastasis [10]. Additionally, YY1 regulates *ATP6V1A*, a subunit of vacuolar H⁺-ATPase, which contributes to GC invasion and metastasis [11]. YY1 also interacts with *SRF* and *ING5*, regulating the transcription of *ING5* in GC [12]. Research has also focused on YY1's role in immune regulation. YY1 can help tumor cells escape immune response by regulating *LAG-3* expression in CD8⁺ T cells within tumors [13]. *circ-ATAD1* promotes GC cell progression by targeting the miR-140-3p/YY1/*PCIF1* signaling axis [13]. Beyond this, YY1 is also associated with immune cell infiltration in GC. Overexpression of YY1 reduces immune cell infiltration in GC tumors, suggesting a mechanism of immune evasion [14]. Clinically, elevated YY1 expression is associated with poor prognosis in GC patients, emphasizing its potential as a prognostic biomarker [10, 15].

This study aims to investigate the differences in survival among GC patients with similar clinical and pathological characteristics. It also seeks to develop a prognostic risk model based on YY1-related gene signatures to predict overall survival. Finally, it evaluates YY1 and its associated genes as potential biomarkers and therapeutic targets in GC, offering insights into their roles in GC progression and patient outcomes.

2 | Materials and Methods

2.1 | Data Acquisition and Pretreatment

Transcriptome data of GC patients were obtained from the Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/>, Project Name: Stomach Adenocarcinoma, Project ID: TCGA-STAD). HTSeq-Counts data for TCGA-STAD were downloaded and processed using a Perl script to construct mRNA expression matrices and convert transcriptome data identifiers. Clinical information for these patients was also retrieved from the University of California, Santa Cruz (UCSC) database (<https://genome.ucsc.edu/>).

In addition, the study cohort included 26 inpatients diagnosed with GC at the Department of General Surgery, the Second

Hospital of Shanxi Medical University (Taiyuan, Shanxi, China), between May 2018 and April 2019. All patients underwent surgical treatment, and the diagnosis of GC was confirmed through postoperative pathological examination. None of the patients received radiotherapy, chemotherapy, or other anticancer treatments before surgery. Ethical approval for this study was obtained from the Ethics Committee of the Second Hospital of Shanxi Medical University (no. 2019YX294). Written informed consent was provided by all participants following the principles outlined in the Declaration of Helsinki.

2.2 | Analysis of YY1 Expression in GC

In the R programming environment (R Studio version 2023.6.2.0, <https://posit.co/download/rstudio-desktop/>), the “limma” package was used to process TCGA data and extract the YY1 expression matrix [16]. Differential expression analysis was done with the “ggplot2” and “ggpubr” packages to create plots showing YY1 expression levels in GC. The YY1 expression matrix was then combined with survival data from GC patients. Prognostic survival analysis of the YY1 gene was carried out using R packages like “survival,” “limma,” and “survminer.” Kaplan–Meier survival curves were created with the “ggpubr” package to visualize and analyze the survival outcomes based on YY1 expression levels in GC patients.

2.3 | Enrichment Analysis of Function and Pathway

Samples were stratified into high and low groups, and differential expression analysis was performed using the “limma” package. Genes were identified as differentially expressed genes (DEGs) based on the criteria of $p < 0.05$, $FDR < 0.05$, and $|\log FC| > 1$. Enrichment analysis of the DEGs was conducted with the “clusterProfiler,” “org.Hs.eg.db,” “enrichplot,” and “dplyr” packages under thresholds of $p < 0.05$ and $FDR < 0.05$ [17]. The results were visualized with bubble and bar plots generated by the “ggplot2” package. Gene set enrichment analysis (GSEA) was also carried out using the “limma,” “org.Hs.eg.db,” “clusterProfiler,” and “enrichplot” packages, along with the Molecular Signatures Database (MSigDB) (c2.cp.kegg.v7.4.symbols.gmt data set, <https://www.gsea-msigdb.org/gsea/msigdb/index.jsp>). Significant enrichment was defined as $p < 0.05$ and corrected $FDR < 0.05$. The GSEA results were visualized with the “ggplot2” package to display the enriched pathways graphically.

2.4 | Analysis of Tumor Microenvironment (TME) and Immune Cell Infiltration

TME scores were calculated from TCGA data using the “limma” and “estimate” packages [18]. Samples were divided into high and low YY1 expression groups, and TME differences were visualized using the “reshape2,” “ggpubr,” and “ggplot2” packages. Statistical significance was represented with asterisks (* for $p < 0.05$, ** for $p < 0.01$, and *** for $p < 0.001$). Gene expression data for 22 immune cell types were obtained, and the CIBERSORT algorithm was applied to the TCGA-STAD

expression matrix data using the “e1071” and “preprocessCore” packages to measure immune cell abundances. This provided detailed information on immune cell infiltration in each sample. Differential analysis of immune infiltration was performed using the “limma,” “vioplot,” and “ggExtra” packages, with a $p < 0.05$ threshold to identify significant differences. Asterisks were used to indicate significance levels. Box plots and correlation graphs showing immune infiltration profiles were created with the “ggplot2” package to clearly display and analyze the immune landscape.

2.5 | Immunohistochemical Experiments

The two-step immunohistochemical detection system was used for staining, which included DAB color development, ethanol gradient dehydration, xylene transparency, and sealing with neutral gum. Following these steps, microscopic examination and image capture were conducted. Cells with brown nuclear particles were identified as positive for YY1 expression, and ImageJ software was used to measure YY1-positive staining. The integrated optical density (IOD) of YY1-positive staining in GC and adjacent normal tissues was analyzed with GraphPad Prism 8 to compare expression levels.

2.6 | The Acquisition of Co-Expressed Genes

Co-expressed genes associated with YY1 were performed using the “limma” package, with a correlation coefficient threshold (corFilter) set at 0.6 and a significance threshold (pFilter) of 0.001. The “ggExtra,” “ggplot2,” and “ggpubr” packages were subsequently used to visualize these correlations. For each gene pair, correlation coefficients and p values were calculated to assess their strength and significance. Scatter plots were created to visually show the relationships between YY1 and its co-expressed genes, allowing for detailed analysis of their interactions.

2.7 | Construction and Evaluation of Prognostic Model

Prognosis-related differential genes were identified using univariate Cox regression analysis with a significance threshold of $p < 0.05$. A LASSO proportional hazards regression model was built using multivariate Cox regression analysis to calculate risk coefficients for the selected genes. The model was optimized to improve prediction accuracy. Using the median risk score as a cutoff, patients were divided into high-risk and low-risk groups. Kaplan–Meier survival curves were created with the “Survival” package and the “survminer” package (version 3.8-3, <https://cran.r-project.org/web/packages/survival/index.html>) to compare survival between the two groups. The “pheatmap” package was used to generate risk curves, survival state plots, and heatmaps showing patient risk profiles. Independent prognostic analysis was done with the “survival” package, and a forest plot was made with the “ggplot2” package to show hazard ratios and confidence intervals for

each factor. The “survivalROC” package was used to create and show receiver operating characteristic (ROC) curves, which evaluated the model's ability to distinguish risk groups based on gene expression.

2.8 | Statistical Analysis

Statistical analysis was performed using R software (version 4.3.1) in the R Studio environment (version 2023.6.2.0). A p value < 0.05 was used to determine statistical significance for all tests in this study.

3 | Results

3.1 | High Expression of YY1 as a Poor Prognostic Factor in GC

Analysis of the YY1 expression matrix from the TCGA database showed that YY1 expression was significantly higher in GC tissues ($n = 415$) compared to normal tissues ($n = 35$) ($p < 0.001$) (Figure 1A,B). Based on the median YY1 expression value in GC, patients were grouped into high and low expression categories. Kaplan–Meier survival analysis demonstrated a significant association between YY1 expression and patient survival ($p = 0.043$) (Figure 1C). Specifically, patients with low YY1 expression had a much higher 5-year survival rate than those with high YY1 expression, suggesting that YY1 could be a useful prognostic marker for GC.

3.2 | Enrichment Analysis of Biological Functions and Pathways

To explore the biological functions and pathways associated with YY1 in GC, enrichment analyses were performed. Gene ontology enrichment analysis showed that YY1 was related to pathways involving the immunoglobulin complex, ion channel complex, immune response mediator production, channel activity, and passive transmembrane transporter activity among DEGs (Figure 2A,B). Similarly, the Kyoto Encyclopedia of Genes and Genome analysis indicated pathways such as cholinergic synapse, GABAergic synapse, and dopaminergic synapse (Figure 2C,D).

To refine the analysis further, GSEA was performed using up-regulated and downregulated DEGs. The results showed that pathways enriched in downregulated DEGs were mainly related to olfactory transduction, while upregulated DEGs were linked to the cell cycle, focal adhesion, and Wnt signaling pathway (Figure 2E,F). High expression of these DEGs was associated with processes like ferric iron binding, chylomicron assembly, and neutrophil-mediated transduction. In contrast, low expression of these DEGs was related to Type I interferon receptor binding, chemical stimuli detection, and olfactory receptor activity. These findings highlight the important roles of YY1 in GC, suggesting it is involved in key biological processes and pathways that may influence disease progression and guide treatment strategies.

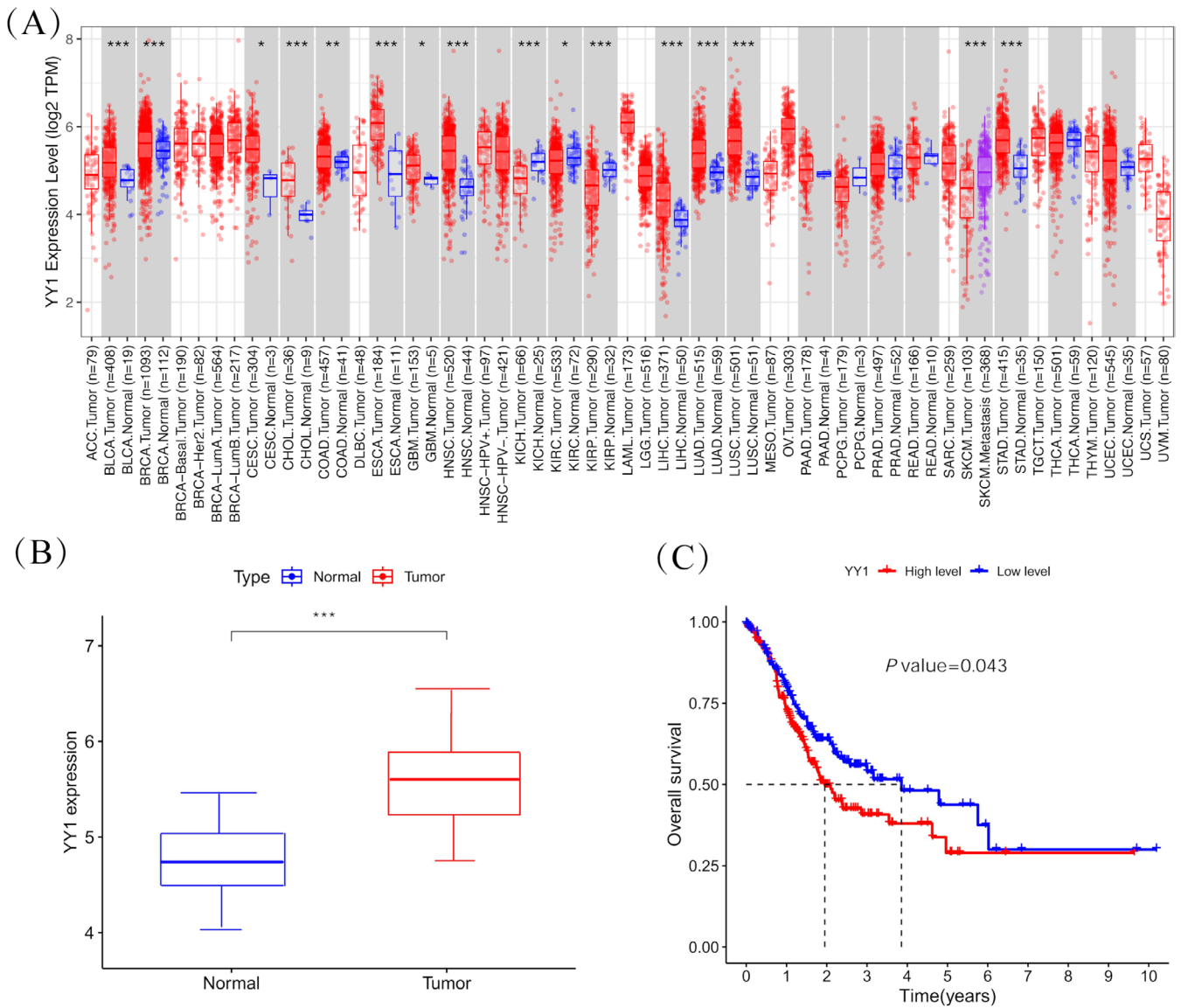


FIGURE 1 | YY1 expression in TCGA gastric cancer. (A) YY1 expression in pan-carcinoma. (B) The expression level of YY1 is higher in STAD Tumors ($n=415$) compared to STAD Normal samples ($n=35$). (C) Prognostic value of YY1 expression in patients with GC. According to the median value of YY1 expression in GC, patients with gastric cancer were stratified into a high-expression group and a low-expression group. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

3.3 | TME and Immune Cell Infiltration Analysis

YY1 expression levels were divided into high and low groups based on the median value. Significant differences were found in stromal score ($p < 0.05$), immune infiltration score ($p < 0.001$), and ESTIMATE score ($p < 0.001$) between these groups (Figure 3A). These scores were higher in tissues with low YY1 expression, indicating that higher YY1 expression is linked to lower stromal cell and immune cell content within the TME (Figure 3B). Moreover, the analysis showed clear differences in immune cell profiles between the GC and normal tissue samples (Figure 3C). Correlation analysis between YY1 expression and specific immune cell types revealed significant negative correlations with natural killer (NK) cells ($r = -0.140963$, $p = 0.028344$), memory B cells ($r = -0.155232$, $p = 0.015648$), plasma cells ($r = -0.193884$, $p = 0.002490$), and

CD8+ T cells ($r = -0.193292$, $p = 0.002529$) (Figure 3D). In contrast, YY1 expression was positively correlated with naive B cells ($r = 0.162505$, $p = 0.011353$), memory CD4+ T cells ($r = 0.203071$, $p = 0.001494$), and macrophage M0 ($r = 0.178314$, $p = 0.005405$) (Figure 3D). These results suggest that YY1 expression may affect the immune cell composition in the GC TME, potentially influencing disease progression and immune response.

3.4 | YY1 Protein Expression in GC Tissue

Immunohistochemical staining showed that YY1 expression was mainly localized in the nucleus, with some presence in the cytoplasm. Imaging was done using Motic DSAssistant software at 4 \times , 10 \times , and 40 \times magnifications. The areas with YY1-positive

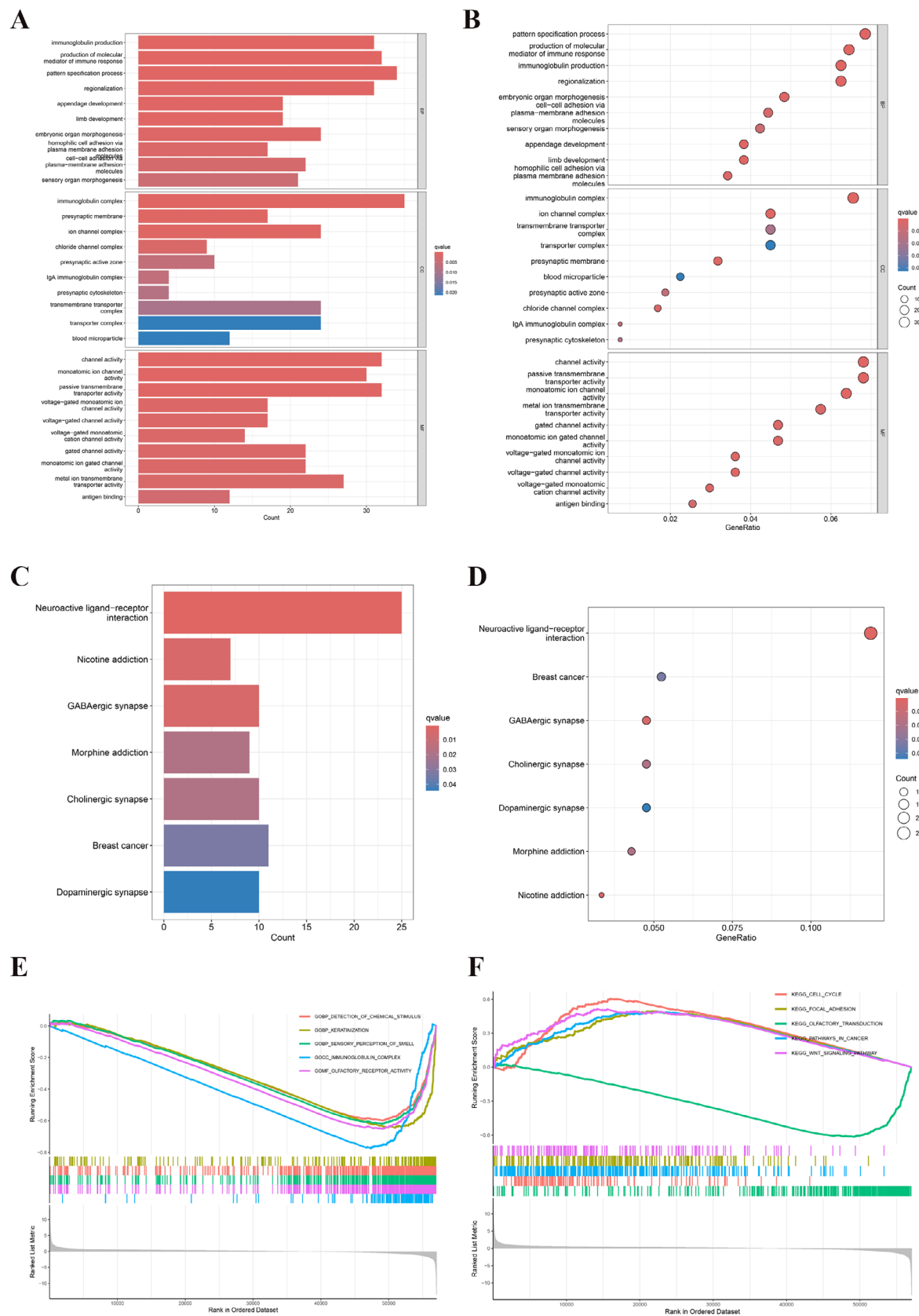


FIGURE 2 | Functions enrichment analysis of YY1. (A) GO enrichment analysis of upregulated DEGs. (B) GO enrichment analysis of downregulated DEGs. (C) KEGG pathway analysis of upregulated DEGs. (D) KEGG pathway analysis of downregulated DEGs. (E) GSEA enrichment analysis of GOs. (F) GSEA enrichment analysis of KEGG pathways.

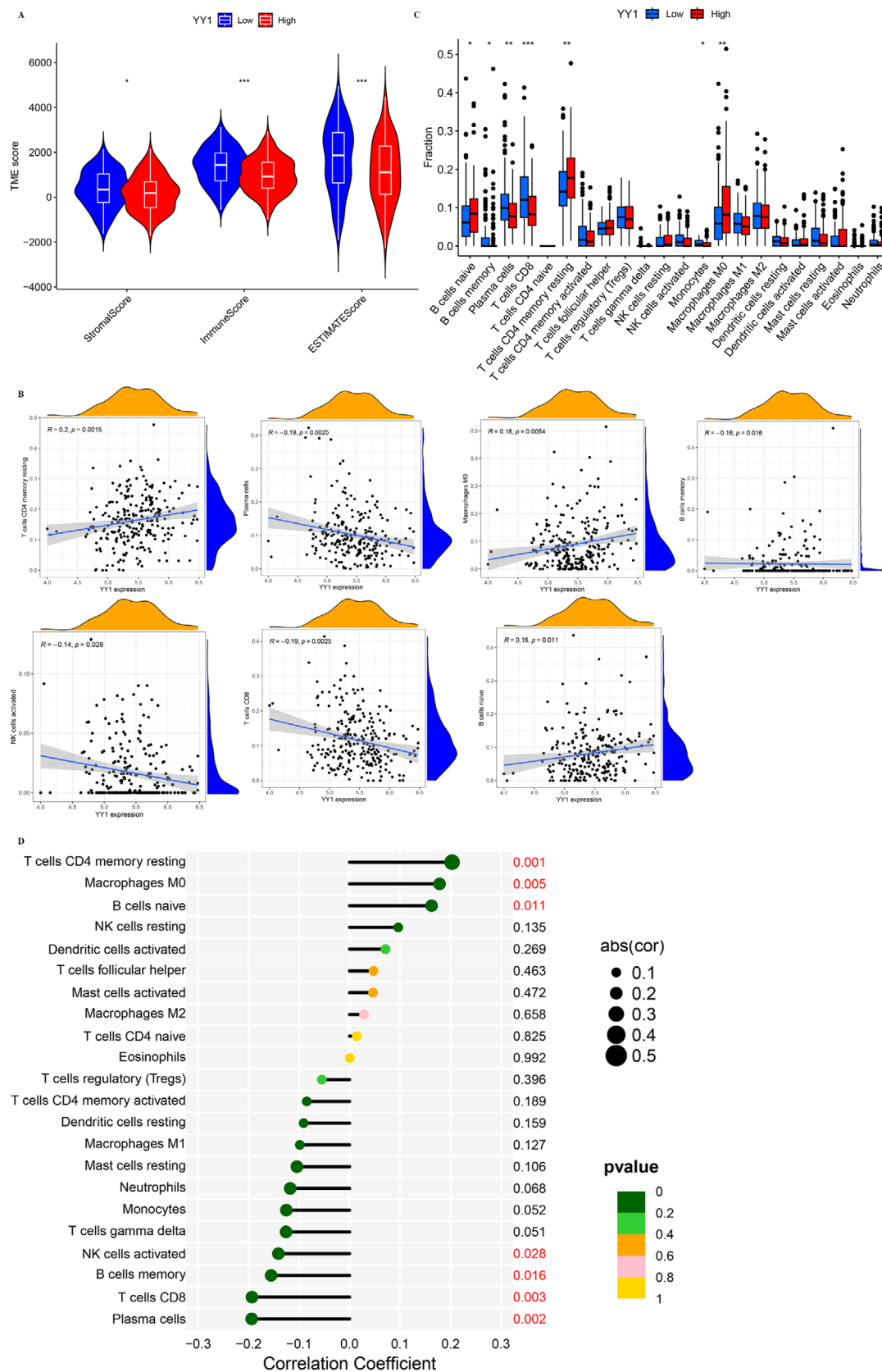


FIGURE 3 | Analysis of tumor microenvironment and immune cell infiltration among patients in different risk groups. (A) Tumor microenvironment score in two groups. (B) Correlation analysis of immune cell infiltration and YY1 expression level. (C) Differential analysis of immune cell infiltration in two groups. (D) Correlation analysis of immune cell population and YY1 expression level. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

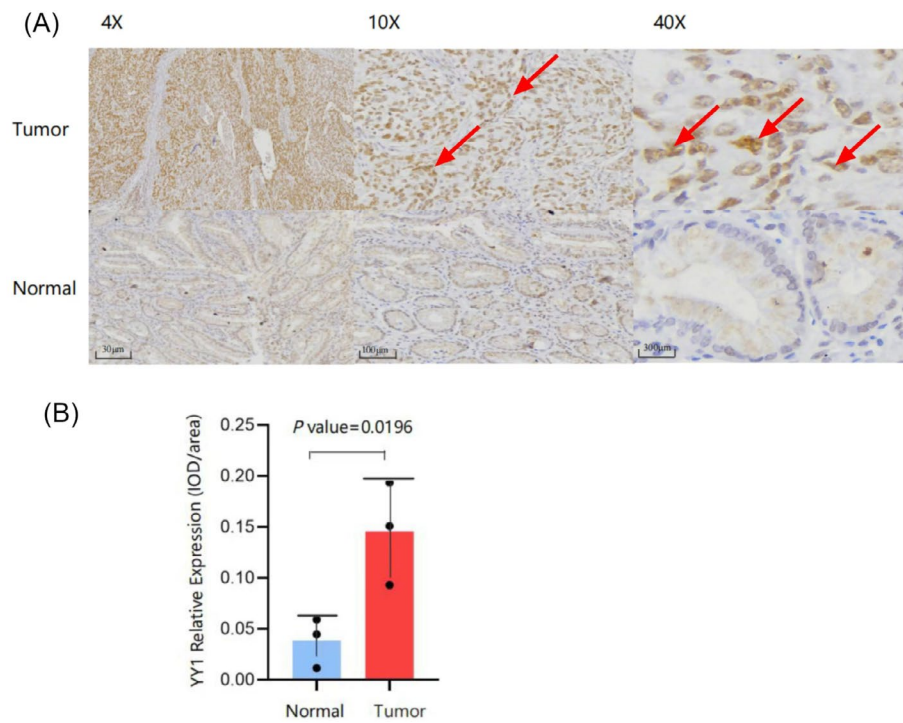


FIGURE 4 | Immunohistochemistry staining of YY1 in GC tissues. (A) Positive expression of YY1 in the nucleus (red arrow). 4x, 10x, and 40x representing the magnification of the objective lens. (B) The immunostaining score of YY1 in gastric cancer and adjacent normal tissues.

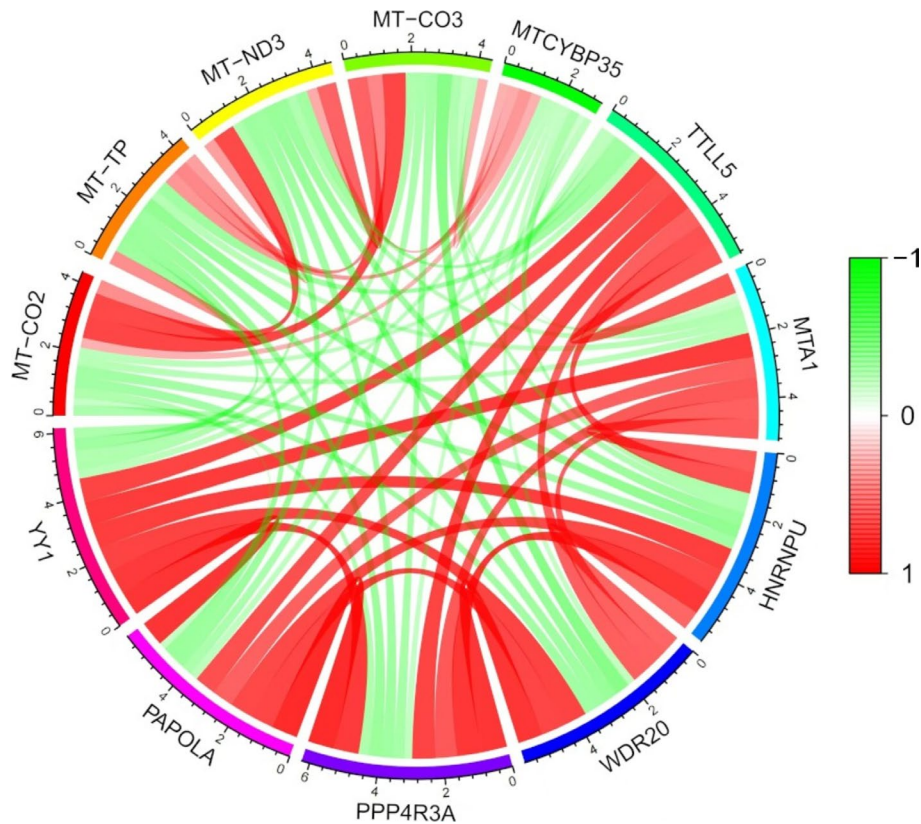


FIGURE 5 | Co-expression analysis of YY1 with other related genes.

cells in GC and normal tissues were quantified using ImageJ software (Figure 4A). Semi-quantitative staining scores showed significantly higher YY1-positive staining in GC tissues compared to

normal tissues ($p=0.0293$) (Figure 4B). These results suggest that YY1 protein overexpression is specific to GC tissue, supporting its potential as a biomarker for GC pathogenesis and diagnosis.

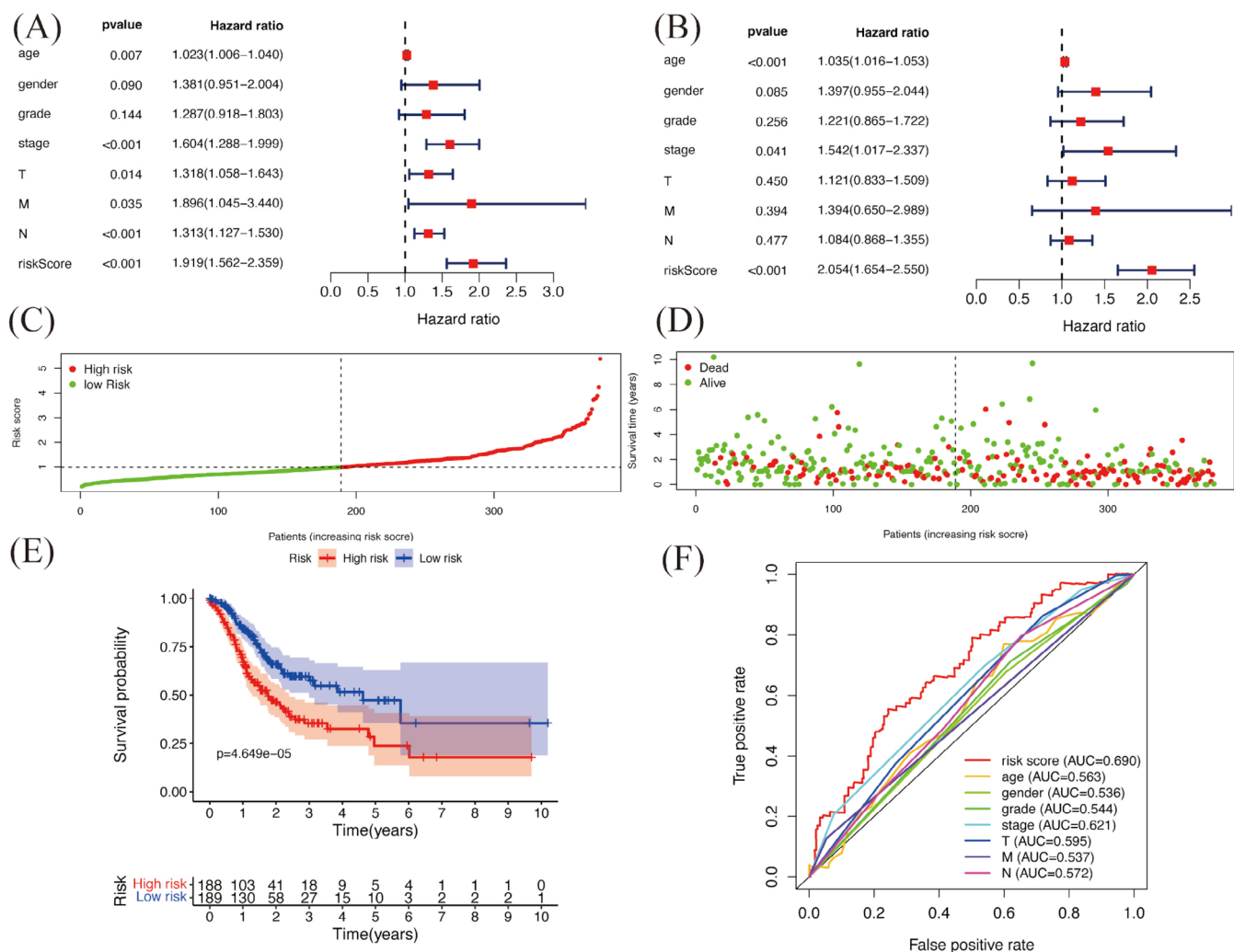


FIGURE 6 | Construction and evaluation of the prognostic model. (A) Univariate independent prognostic factor analysis. (B) Multivariate independent prognostic factor analysis. (C) Prognostic model risk curve. (D) Patients' risk score with survival time. (E) Kaplan-Meier survival curve. (F) ROC curve for validating the prognostic model.

3.5 | Co-Expression Analysis

Co-expression analysis showed positive correlations between YY1 and genes such as TTL15, MTA1, HNRNPU, WDR20, and PPP4R3A. On the other hand, negative correlations were found between YY1 and genes like MT-CO2, MT-TP, MT-ND3, MT-CO3, and MTCYBP35 (Figure 5). These results suggest that YY1 may interact with and regulate these genes in GC. These associations offer useful insights into the role of YY1 and its involvement in disease progression.

3.6 | Construction and Evaluation of the Prognostic Model

To create a strong prognostic model, univariate and multivariate Cox regression analyses were carried out. The model was defined by the following formula:

Univariate analysis found several factors linked to survival, such as age, clinical stage, TNM classification, and risk score (Figure 6A). Multivariate analysis confirmed that age, clinical stage, and risk score were independent predictors of prognosis (Figure 6B), with *p* values of age and clinical stage less than 0.001. These results show that the model is effective for predicting patient outcomes.

Based on the risk score, patients were divided into high-risk and low-risk groups using the median score (Figure 6C,D). Kaplan-Meier survival analysis showed a significant difference in survival between the two groups (*p*=4.649e-05; Figure 6E). The overall survival rates for the whole group were 67.5% at 1 year, 35.4% at 3 years, and 23.7% at 5 years. In contrast, the low-risk group had better survival rates: 84.4% at 1 year, 57.4% at 3 years, and 47.3% at 5 years. The results confirm that lower risk scores

$$\text{Risk score} = -0.42535 \times \text{ALG11} + 0.910403 \times \text{KBTBD2} - 0.37175 \times \text{POLR1A} - 0.4042 \times \text{PPOSER1} - 0.37019 \times \text{RRBM15} + 0.023839 \times \text{YY1}$$

are linked to longer survival, supporting the model's prognostic value.

To check the model's accuracy, we performed ROC curve analysis. The area under the curve (AUC) for the risk score was 0.690, higher than other clinical variables, showing its better predictive power (Figure 6F). These findings highlight the model's potential to help guide treatment for GC patients.

4 | Discussion

GC is characterized by significant molecular and phenotypic heterogeneity [19, 20], with tumor tissues consisting of diverse clonal subpopulations, each of which exhibits distinct phenotypes and genetic profiles. This intratumoral molecular diversity makes it harder to understand the disease and its treatment. Emerging evidence suggests that such heterogeneity contributes to the limited efficacy of chemotherapy and immunotherapy, as well as the poor prognosis commonly associated with GC [21, 22]. Identifying distinct patient subpopulations with prognostic and/or predictive biomarkers is therefore important for creating targeted treatments and improving precision oncology.

YY1 is a multifunctional protein that acts as a transcription factor, RNA-binding protein, and chromatin regulator. It plays a key role in processes like tumor cell growth, invasion, epithelial-mesenchymal transition, DNA damage response, and carcinogenesis [23]. Our bioinformatics analysis has revealed that YY1 is highly expressed in GC tissues, which was confirmed by immunohistochemistry. Elevated YY1 expression correlates with poorer prognosis in GC patients, making it a potential prognostic biomarker [24]. Further analysis indicated that YY1 overexpression activates key pathways, such as the cell cycle and Wnt signaling, both of which are involved in carcinogenesis. This may help explain the poor prognosis associated with YY1 overexpression. Other evidence suggests that silencing YY1 in GC cells inhibits multiple important signaling pathways, including JNK/MAPK, ER, Wnt/ β -catenin, ERK/MAPK, and HIF-1 α [25]. Furthermore, the GSEA results also show how YY1 promotes GC proliferation and metastasis by regulating the CCDC43-ADRM1 axis [10].

CD8+ T cell-mediated cellular immunity plays a critical role in eliminating tumor cells, and the effectiveness of immunotherapy largely depends on the activation of these cells. However, cancer cells can avoid immune surveillance by suppressing CD8+ T cell activity, enabling their survival through immune escape mechanisms [26, 27]. In analyzing the TME and immune cell infiltration, we showed a complicated relationship between YY1 expression and the immune landscape in GC. Specifically, YY1 expression negatively correlated with immune cell populations such as memory B cells, NK cells, CD8+ T cells, and plasma cells. High YY1 expression was associated with reduced content of both the extracellular matrix and immune cells, suggesting an immunosuppressive effect. Notably, high YY1 expression in tumor tissues was often linked to a decrease in immune cell numbers, potentially contributing to immune evasion and poor prognosis [28, 29]. Merenstein et al. further elucidated a potential molecular mechanism underlying this phenomenon. Their findings revealed that YY1 overexpression regulates the

transcription of LAG-3 in CD8+ T cells, leading to their inactivation and depletion. This loss of CD8+ T cell functionality facilitates immune escape and supports tumor progression [13]. Interestingly, YY1 expression was positively correlated with naive B cells, memory CD4+ T cells, and M0 macrophages, suggesting a more complicated role of YY1 in modulating the immune microenvironment. In some cases, CD4+ T cell subtypes and M0 macrophages can promote tumor growth, which needs further study [30–32].

Co-expression analysis revealed interesting relationships between YY1 and other genes in GC. Specifically, YY1 was positively correlated with TTL15, MTA1, HNRNPU, WDR20, and PPP4R3A, which may suggest potential functional interactions or shared regulatory mechanisms. A particularly significant finding was the positive correlation between YY1 and MTA1, a gene linked to metastasis. This suggests that YY1 may contribute to tumor progression and metastasis [33]. Furthermore, YY1 expression was negatively correlated with mitochondrial genes such as MT-CO2, MT-TP, MT-ND3, MT-CO3, and MTCYBP35. This suggests that YY1 may be linked to mitochondrial function or energy metabolism in cancer cells, providing new insights into the metabolic mechanisms behind YY1's in GC [34].

The development of a six-gene prognostic model that includes YY1 expression represents a significant advancement in the risk stratification of GC patients. The model effectively distinguishes between high-risk and low-risk groups with different survival outcomes, showing its potential for clinical treatment. Compared to traditional clinical variables, this model exhibits better predictive accuracy, as shown by its higher AUC values. Multivariate analyses also confirm the independent prognostic value of the risk scores, with age and cancer stage as additional factors, highlighting the strength of our model. By combining molecular biomarkers with clinical parameters, this approach supports the growing focus on personalized cancer treatment.

A literature search using keywords like YY1, risk scores, prognostic models, multivariate analysis, and GC did not find studies that include all these elements together. This study appears to be the first to combine a risk score based on YY1 expression with clinical parameters like age and GC stage in a prognostic model. These results have important clinical implications. Since YY1 is overexpressed and linked to poor clinical outcomes, strategies to inhibit YY1 activity or expression could offer promising treatments. Potential approaches might include small molecule inhibitors, antisense oligonucleotides, or other innovative therapies [35, 36]. Furthermore, the immunosuppressive effect linked to high YY1 expression suggests that targeting YY1 could enhance the efficacy of immunotherapy. Combining YY1 suppression with immune checkpoint inhibitors or other immunotherapeutic strategies may yield synergistic benefits [37, 38].

4.1 | Limitations

This study primarily focuses on the role of YY1 in GC tumorigenesis through bioinformatics approaches. While many genes and pathways are linked to YY1 overexpression, the specific

ones driving GC cell growth and invasion are still unknown. Future studies using in vitro and in vivo models are needed to clarify these mechanisms. Also, the effect of YY1 on immune cell infiltration in the GC microenvironment requires further study. Research shows that gene sets regulated by YY1 are mainly linked to intestinal-type GC and some cases of diffuse-type GC, highlighting the need to study these two subtypes further [21].

5 | Conclusion

In this study, we created a new prognostic risk model that combines YY1 expression with clinical variables like age and tumor stage to predict clinical outcomes in GC. Our results show that YY1 overexpression is closely related to key aspects of tumor biology, such as immune evasion, metabolic weaknesses, and pathway-specific alterations. What makes our research new is the combination of YY1 expression, risk scores, and multivariate analysis in the context of GC, which has not been done before. These results lay the groundwork for further studies on YY1 as a prognostic marker and a therapeutic target in GC, possibly leading to better outcomes for this disease.

Author Contributions

Wenliang Chen: conceptualization (lead), data curation (equal), formal analysis (equal), funding acquisition (lead), investigation (lead), methodology (lead), project administration (lead), resources (lead), software (equal), supervision (lead), validation (lead), visualization (equal), writing – original draft (lead), writing – review and editing (equal). **Huanhuan Wang:** conceptualization (supporting), data curation (equal), formal analysis (equal), methodology (lead), resources (equal), software (lead), validation (lead), writing – original draft (equal), writing – review and editing (equal). **Ntiak Achi:** data curation (equal), formal analysis (supporting), methodology (supporting), resources (equal), validation (equal), visualization (supporting), writing – original draft (equal), writing – review and editing (equal). **Jinjin Hao:** data curation (supporting), formal analysis (equal), software (supporting), visualization (equal), writing – original draft (equal), writing – review and editing (equal). **Qiang Zhao:** investigation (equal), methodology (supporting), resources (supporting), software (equal), writing – original draft (equal), writing – review and editing (equal). **Rui Gong:** data curation (supporting), resources (supporting), software (equal), validation (equal), writing – original draft (equal), writing – review and editing (equal).

Acknowledgments

This study was supported by the Natural Science Foundation of Shanxi Province, China (Grant no. 201801D121320).

Ethics Statement

This study followed the Declaration of Helsinki and was approved by the Ethics Committee of the Second Hospital of Shanxi Medical University (approval no. 2019YX294).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available in the UCSC database at (<https://genome.ucsc.edu/>).

References

1. J. Ferlay, M. Colombet, I. Soerjomataram, et al., “Estimating the Global Cancer Incidence and Mortality in 2018: GLOBOCAN Sources and Methods,” *International Journal of Cancer* 144, no. 8 (2019): 1941–1953, <https://doi.org/10.1002/ijc.31937>.
2. Y. Hirata, A. Noorani, S. Song, L. Wang, and J. A. Ajani, “Early Stage Gastric Adenocarcinoma: Clinical and Molecular Landscapes,” *Nature Reviews Clinical Oncology* 20, no. 7 (2023): 453–469, <https://doi.org/10.1038/s41571-023-00767-w>.
3. Y. Zeng and R. U. Jin, “Molecular Pathogenesis, Targeted Therapies, and Future Perspectives for Gastric Cancer,” *Seminars in Cancer Biology* 86 (2022): 566–582, <https://doi.org/10.1016/j.semcancer.2021.12.004>.
4. S. Gordon, G. Akopyan, H. Garban, and B. Bonavida, “Transcription Factor YY1: Structure, Function, and Therapeutic Implications in Cancer Biology,” *Oncogene* 25, no. 8 (2006): 1125–1142, <https://doi.org/10.1038/sj.onc.1209080>.
5. N. Agarwal and D. Theodorescu, “The Role of Transcription Factor YY1 in the Biology of Cancer,” *Critical Reviews in Oncogenesis* 22, no. 1–2 (2017): 13–21, <https://doi.org/10.1615/CritRevOncog.2017021071>.
6. Z. Shao, W. Yang, X. Meng, et al., “The Role of Transcription Factor Yin Yang-1 in Colorectal Cancer,” *Cancer Medicine* 12, no. 10 (2023): 11177–11190, <https://doi.org/10.1002/cam4.5745>.
7. A. A. Cho and B. Bonavida, “Targeting the Overexpressed YY1 in Cancer Inhibits EMT and Metastasis,” *Critical Reviews in Oncogenesis* 22, no. 1–2 (2017): 49–61, <https://doi.org/10.1615/CritRevOncog.2017024073>.
8. L. M. Khachigian, “The Yin and Yang of YY1 in Tumor Growth and Suppression,” *International Journal of Cancer* 143, no. 3 (2018): 460–465, <https://doi.org/10.1002/ijc.31255>.
9. P. Panneerandian, D. B. Rao, and K. Ganesan, “Calcium Channel Blockers Lercanidipine and Amlodipine Inhibit YY1/ERK/TGF- β Mediated Transcription and Sensitize the Gastric Cancer Cells to Doxorubicin,” *Toxicology In Vitro* 74 (2021): 105152, <https://doi.org/10.1016/j.tiv.2021.105152>.
10. J. Wang, X. Wu, W. Dai, et al., “The CCDC43-ADRM1 Axis Regulated by YY1, Promotes Proliferation and Metastasis of Gastric Cancer,” *Cancer Letters* 482 (2020): 90–101, <https://doi.org/10.1016/j.canlet.2020.03.026>.
11. P. Wang, L. Wang, J. Sha, et al., “Expression and Transcriptional Regulation of Human ATP6V1A Gene in Gastric Cancers,” *Scientific Reports* 7, no. 1 (2017): 3015, <https://doi.org/10.1038/s41598-017-03021-3>.
12. H. C. Zheng, H. Xue, X. Wu, H. L. Xu, E. H. Zhao, and Z. G. Cui, “Transcriptional Regulation of ING5 and Its Suppressive Effects on Gastric Cancer,” *Frontiers in Oncology* 12 (2022): 918954, <https://doi.org/10.3389/fonc.2022.918954>.
13. A. Merenstein, L. Obeidat, A. Zaravinos, and B. Bonavida, “The Role of YY1 in the Regulation of LAG-3 Expression in CD8 T Cells and Immune Evasion in Cancer: Therapeutic Implications,” *Cancers* 17, no. 1 (2024): 19, <https://doi.org/10.3390/cancers17010019>.
14. L. Zhang, X. Chang, T. Zhai, et al., “A Novel Circular RNA, Circ-ATAD1, Contributes to Gastric Cancer Cell Progression by Targeting miR-140-3p/YY1/PCIF1 Signaling Axis,” *Biochemical and Biophysical Research Communications* 525, no. 4 (2020): 841–849, <https://doi.org/10.1016/j.bbrc.2020.02.100>.
15. Z. H. Geng, J. X. Du, Y. D. Chen, et al., “YY1: A Key Regulator Inhibits Gastric Cancer Ferroptosis and Mediating Apatinib-Resistance,” *Cancer Cell International* 24, no. 1 (2024): 71, <https://doi.org/10.1186/s12935-024-03262-z>.
16. G. K. Smith, “Limma: Linear Models for Microarray Data,” in *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, ed. R. Gentleman, V. Carey, S. Dudoit, R. Irizarry, and W. Huber (Springer, 2005), 397–420.

17. G. Yu, L. G. Wang, Y. Han, and Q. Y. He, "clusterProfiler: An R Package for Comparing Biological Themes Among Gene Clusters," *OMICS* 16, no. 5 (2012): 284–287, <https://doi.org/10.1089/omi.2011.0118>.
18. K. Yoshihara, M. Shahmoradgol, E. Martinez, et al., "Inferring Tumour Purity and Stromal and Immune Cell Admixture From Expression Data," *Nature Communications* 4 (2013): 2612, <https://doi.org/10.1038/ncomms3612>.
19. S. W. T. Ho and P. Tan, "Dissection of Gastric Cancer Heterogeneity for Precision Oncology," *Cancer Science* 110, no. 11 (2019): 3405–3414, <https://doi.org/10.1111/cas.14191>.
20. B. Wang, B. Song, Y. Li, Q. Zhao, and B. Tan, "Mapping Spatial Heterogeneity in Gastric Cancer Microenvironment," *Biomedicine & Pharmacotherapy* 172 (2024): 116317, <https://doi.org/10.1016/j.biopha.2024.116317>.
21. Cancer Genome Atlas Research Network, "Comprehensive Molecular Characterization of Gastric Adenocarcinoma," *Nature* 513, no. 7517 (2014): 202–209, <https://doi.org/10.1038/nature13480>.
22. R. Cristescu, J. Lee, M. Nebozhyn, et al., "Molecular Analysis of Gastric Cancer Identifies Subtypes Associated With Distinct Clinical Outcomes," *Nature Medicine* 21, no. 5 (2015): 449–456, <https://doi.org/10.1038/nm.3850>.
23. G. U. Martinez-Ruiz, A. Morales-Sanchez, and A. F. Pacheco-Hernandez, "Roles Played by YY1 in Embryonic, Adult and Cancer Stem Cells," *Stem Cell Reviews and Reports* 17, no. 5 (2021): 1590–1606, <https://doi.org/10.1007/s12015-021-10151-9>.
24. W. Kang, J. H. Tong, A. W. Chan, et al., "Yin Yang 1 Contributes to Gastric Carcinogenesis and Its Nuclear Expression Correlates With Shorter Survival in Patients With Early Stage Gastric Adenocarcinoma," *Journal of Translational Medicine* 12, no. 1 (2014): 80, <https://doi.org/10.1186/1479-5876-12-80>.
25. D. Bhaskar Rao, P. Panneerandian, K. Balakrishnan, and K. Ganesan, "YY1 Regulated Transcription-Based Stratification of Gastric Tumors and Identification of Potential Therapeutic Candidates," *Journal of Cell Communication and Signaling* 15, no. 2 (2021): 251–267, <https://doi.org/10.1007/s12079-021-00608-4>.
26. X. Chen, Q. Lu, H. Zhou, et al., "A Membrane-Associated MHC-I Inhibitory Axis for Cancer Immune Evasion," *Cell* 186, no. 18 (2023): 3903–3920.e21, <https://doi.org/10.1016/j.cell.2023.07.016>.
27. C. Fu and A. Jiang, "Dendritic Cells and CD8 T Cell Immunity in Tumor Microenvironment," *Frontiers in Immunology* 9 (2018): 3059, <https://doi.org/10.3389/fimmu.2018.03059>.
28. C. Luo, X. Chen, Y. Bai, et al., "Upregulation of Yin-Yang-1 Associates With Proliferation and Glutamine Metabolism in Esophageal Carcinoma," *International Journal of Genomics* 2022 (2022): 9305081, <https://doi.org/10.1155/2022/9305081>.
29. I. T. S. Meliala, R. Hosea, V. Kasim, and S. Wu, "The Biological Implications of Yin Yang 1 in the Hallmarks of Cancer," *Theranostics* 10, no. 9 (2020): 4183–4200, <https://doi.org/10.7150/thno.43481>.
30. S. Chen, K. Lu, Y. Hou, et al., "YY1 Complex in M2 Macrophage Promotes Prostate Cancer Progression by Upregulating IL-6," *Journal for Immunotherapy of Cancer* 11, no. 4 (2023): e006020, <https://doi.org/10.1136/jitc-2022-006020>.
31. E. Hays and B. Bonavida, "YY1 Regulates Cancer Cell Immune Resistance by Modulating PD-L1 Expression," *Drug Resistance Updates* 43 (2019): 10–28, <https://doi.org/10.1016/j.drug.2019.04.001>.
32. A. Dillen, I. Bui, M. Jung, S. Agioti, A. Zaravinos, and B. Bonavida, "Regulation of PD-L1 Expression by YY1 in Cancer: Therapeutic Efficacy of Targeting YY1," *Cancers (Basel)* 16, no. 6 (2024): 1237, <https://doi.org/10.3390/cancers16061237>.
33. Y. L. Yao and W. M. Yang, "The Metastasis-Associated Proteins 1 and 2 Form Distinct Protein Complexes With Histone Deacetylase Activity," *Journal of Biological Chemistry* 278, no. 43 (2003): 42560–42568, <https://doi.org/10.1074/jbc.M302955200>.
34. E. Errichiello, A. Balsamo, M. Cerni, and T. Venesio, "Mitochondrial Variants in MT-CO2 and D-Loop Instability Are Involved in MUTYH-Associated Polyposis," *Journal of Molecular Medicine (Berlin, Germany)* 93, no. 11 (2015): 1271–1281, <https://doi.org/10.1007/s00109-015-1312-0>.
35. R. Hosea, S. Hillary, S. Wu, and V. Kasim, "Targeting Transcription Factor YY1 for Cancer Treatment: Current Strategies and Future Directions," *Cancers (Basel)* 15, no. 13 (2023): 3506, <https://doi.org/10.3390/cancers15133506>.
36. D. Kwiatkowska, E. Mazur, and A. Reich, "YY1 Is a Key Player in Melanoma Immunotherapy/Targeted Treatment Resistance," *Frontiers in Oncology* 12 (2022): 856963, <https://doi.org/10.3389/fonc.2022.856963>.
37. W. J. Liu, L. Wang, F. M. Zhou, et al., "Elevated NOX4 Promotes Tumorigenesis and Acquired EGFR-TKIs Resistance via Enhancing IL-8/PD-L1 Signaling in NSCLC," *Drug Resistance Updates* 70 (2023): 100987, <https://doi.org/10.1016/j.drug.2023.100987>.
38. Z. Qiu, L. Zhao, J. Z. Shen, et al., "Transcription Elongation Machinery Is a Druggable Dependency and Potentiates Immunotherapy in Glioblastoma Stem Cells," *Cancer Discovery* 12, no. 2 (2022): 502–521, <https://doi.org/10.1158/2159-8290.CD-20-1848>.