

Identification of BANF1 as a novel prognostic biomarker in gastric cancer and validation via *in-vitro* and *in-vivo* experiments

Yuanmin Xu^{1,*}, Xu Wang^{1,*}, Weiwei Yuan², Ling Zhang³, Wei Chen¹, Kongwang Hu^{1,4}

¹Department of General Surgery, First Affiliated Hospital of Anhui Medical University, Hefei 230022, China

²Department of General Surgery, Anhui Public Health Clinical Center, Hefei 230032, China

³Department of Immunology, School of Basic Medical Sciences, Anhui Medical University, Hefei 230032, China

⁴Department of General Surgery, Fuyang Affiliated Hospital of Anhui Medical University, Fuyang 236000, China

*Equal contribution

Correspondence to: Wei Chen, Kongwang Hu; **email:** chenwei366@ahmu.edu.cn; hukongwang69@163.com, <https://orcid.org/0009-0007-4781-3948>

Keywords: gastric cancer, barrier to autointegration factor 1, overall survival, TCGA, GEO

Received: September 18, 2023 **Accepted:** December 14, 2023 **Published:** January 22, 2024

Copyright: © 2024 Xu et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/) (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Gastric cancer (GC) is a widespread malignancy characterized by a notably high incidence rate and an unfavorable prognosis. We conducted a meticulous analysis of GC high-throughput sequencing data downloaded from the Gene Expression Omnibus (GEO) repository to pinpoint distinctive genes associated with GC. Our investigation successfully identified three signature genes implicated in GC, with a specific focus on the barrier to autointegration factor 1 (BANF1), which exhibits elevated expression across various cancer types, including GC. Bioinformatic analysis has highlighted BANF1 as a prognostic indicator for patients with GC, with direct implications for immune cell infiltration. To gain a more comprehensive understanding of the significance of BANF1 in GC, we performed a series of *in vitro* experiments to confirm its high expression in GC tissues and cellular components. Intriguingly, the induction of BANF1 knockdown resulted in a marked attenuation of proliferation, migratory capacity, and invasive potential in GC cells. Moreover, our *in vivo* experiments using nude mouse models revealed a notable impediment in tumor growth following BANF1 knockdown. These insights underscore the feasibility of BANF1 as a novel therapeutic target for GC.

INTRODUCTION

GC poses a significant global health challenge within the gastrointestinal tract. In 2020 alone, the global incidence of GC surpassed 1,089,103 cases, leading to the unfortunate demise of over 760,000 individuals [1]. *Helicobacter pylori* infection stands out as a major contributing factor to GC [2]. Diagnostic strategies, such as computed tomography, endoscopy, and histopathological examination, form the primary tools for GC detection. Surgical intervention remains pivotal for early-stage GC, while unresectable cases necessitate a multimodal approach combining symptomatic management, chemotherapy, and targeted therapies to

enhance both quality of life and survival periods [3, 4]. The intrinsic heterogeneity of GC, as classified by The Cancer Genome Atlas (TCGA) into four subtypes, poses a formidable challenge, emphasizing the need for personalized treatment strategies [5]. Molecular biology investigations utilizing high-throughput sequencing hold promise in uncovering novel therapeutic targets for a more effective GC treatment landscape.

Barrier to autointegration factor 1 (BANF1), predominantly situated in the inner nuclear membrane, plays a crucial role in fundamental biological processes [6]. In the context of DNA repair, BANF1 emerges as a central figure in resolving DNA double-stranded breaks,

safeguarding cells from demise caused by the exposure of double-stranded DNA to the cytoplasm during nuclear rupture [7]. Its influence on the cell cycle, intricately linked to cell mitosis initiation and cessation in phosphorylated and dephosphorylated states, further underscores its significance [8]. The relevance of BANF1 to human health is exemplified in premature aging syndromes, as indicated by previous studies [9]. Recent research has linked elevated BANF1 expression to breast cancer, correlating with lymph node positive and pathological staging, and to differentiation in esophageal tumor and influencing the proliferation, migration, and invasion of cervical cancer [10–12]. These findings suggest the potential of BANF1 as a biomarker and therapeutic target in various malignant neoplasms, including GC.

Our study aimed to comprehensively analyze transcriptome data from GC samples obtained from GEO and TCGA repositories. We sought to elucidate the pivotal role of BANF1 in GC pathogenesis through bioinformatic analyses, assessing its expression pattern, diagnostic utility, prognostic implications, and relevance within the immune milieu. *In vitro* experiments were conducted to validate BANF1 expression in gastric cells and understand its impact on their proliferative, migratory, and invasive properties. The investigation in nude mice aimed to assess the effect of BANF1 on tumorigenesis. This study was driven by the dual objectives of identifying prospective GC biomarkers and empirically validating BANF1 as a therapeutic target for this malignancy.

RESULTS

Differentially expressed genes (DEGs) in GC

We obtained gene expression data for a cohort of 162 GC samples from the GSE54129 and GSE118916

datasets. The principal component analysis (PCA) plots depict gene expression in the GSE54129 and GSE118916 datasets before and after removal of batch effect. Supplementary Figure 1A illustrates the PCA results before batch-effect removal, where the two datasets were initially separated. Subsequently, the PCA plot after batch-effect removal demonstrated an intersection of gene expression data from the two datasets (Supplementary Figure 1B), enabling further analysis. By examining the expression patterns across tumor and normal samples, we successfully identified 157 DEGs characterized by $|\text{Log}_2\text{FC}| > 2$ and $p < 0.01$, as shown in Figure 1A. Among these, 93 genes were downregulated, while 64 were upregulated, visually depicted in Figure 1B, 1C.

Identification of hub genes through weighted correlation network analysis (WGCNA)

To uncover pivotal genes of significance within GC samples, we conducted a comprehensive WGCNA. The optimal soft threshold was selected as 7, as depicted in Figure 2A, led to the establishment of gene clustering results marked by a threshold ensuring at least 60 genes per module, resulting in 11 distinct modules (Figure 2B). These modules underwent consolidation with a clipping height of 0.25, ultimately amalgamating into eight unique modules (Figure 2C). Subsequently, we computed the correlation coefficients between each gene module and samples from both normal and GC contexts. Upon rigorous assessment of correlation coefficients and associated p -values, we observed that the black module exhibited a noteworthy correlation with clinical traits, characterized by a correlation coefficient of 0.78 and $p < 0.001$ (Figure 2D). Figure 2E contains 444 data points, visually represents the relationship between gene significance (GS) and module membership (MM). The horizontal axis signifies MM, while the vertical axis represents GS.

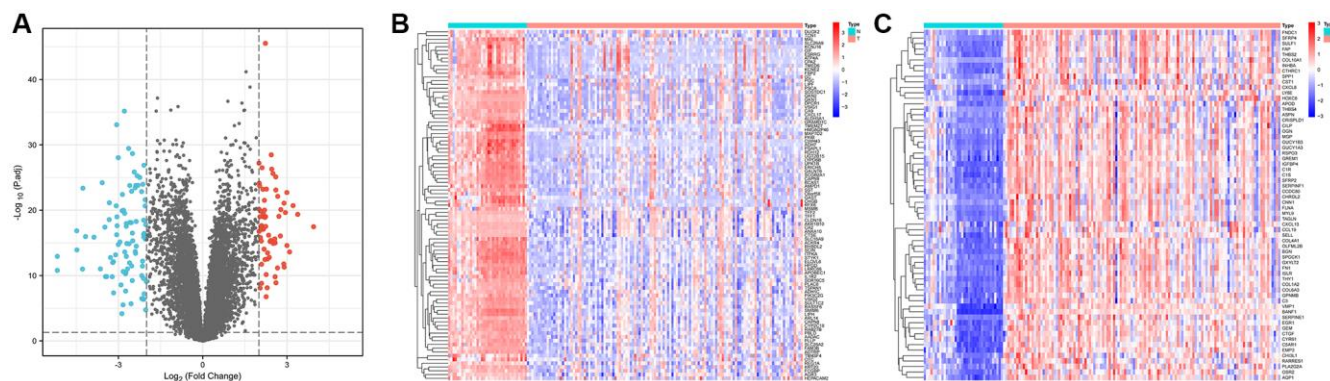


Figure 1. Identification of differentially expressed genes in gastric cancer and paracancerous tissues. (A) Volcano plot of 157 differentially expressed genes ($|\text{Log}_2\text{FC}| > 2$, $p < 0.01$). (B) Heatmap of 93 down-regulated genes expressed in gastric cancer tissues. (C) Heatmap of 64 expressed up-regulated genes in gastric cancer tissues.

Applying stringent criteria (MM >0.8 and GS >0.5), we identified a cadre of 213 hub genes.

Machine learning screens for disease characterizing genes

To delineate the key genes linked to GC, we conducted a rigorous analysis employing least absolute shrinkage and selection operator (LASSO), support vector machine-recursive feature elimination (SVM-RFE), and random forest (RF) methodologies. Within the LASSO algorithm framework, we carefully selected optimal lambda parameters and executed a 10-fold cross-validation procedure, ultimately pinpointing a cohort of 12 genes, as illustrated in Figure 3A, 3B. The SVM-RFE algorithm identified the optimal subset of genes by pinpointing points corresponding to the lowest error rate, resulting in the selection of 28 genes, as demonstrated in Figure 3C. Furthermore, the RF algorithm provided a comprehensive assessment of each gene, ranking them based on their relative importance levels, as depicted in Figure 3D, 3E. Converging the

outcomes derived from WGCNA, LASSO, SVM-RFE, and RF algorithms, we conducted an intersection analysis using a Venn diagram. This meticulous process led to the identification of three distinct genes – BANF1, ADH7, and TMEM27 – collectively characterized as hallmark genes associated with GC (Figure 3F).

Diagnostic evaluation of three characterizing genes

The expression levels and diagnostic efficacy of three key genes, namely BANF1, ADH7, and TMEM27, were thoroughly examined across diverse datasets. This assessment utilized the Wilcoxon test and receiver operating characteristic (ROC) analysis. Notably, BANF1 exhibited significant upregulation in the GSE54129-GSE118916 dataset, GSE65801, and TCGA datasets, while ADH7 and TMEM27 showed down-regulation (Figure 4A, 4C, 4E; $p < 0.05$). In the GSE54129-GSE118916 dataset, the area under the curve (AUC) values for BANF1, ADH7, and TMEM27 were 0.996, 0.952, and 0.949 respectively (Figure 4B).

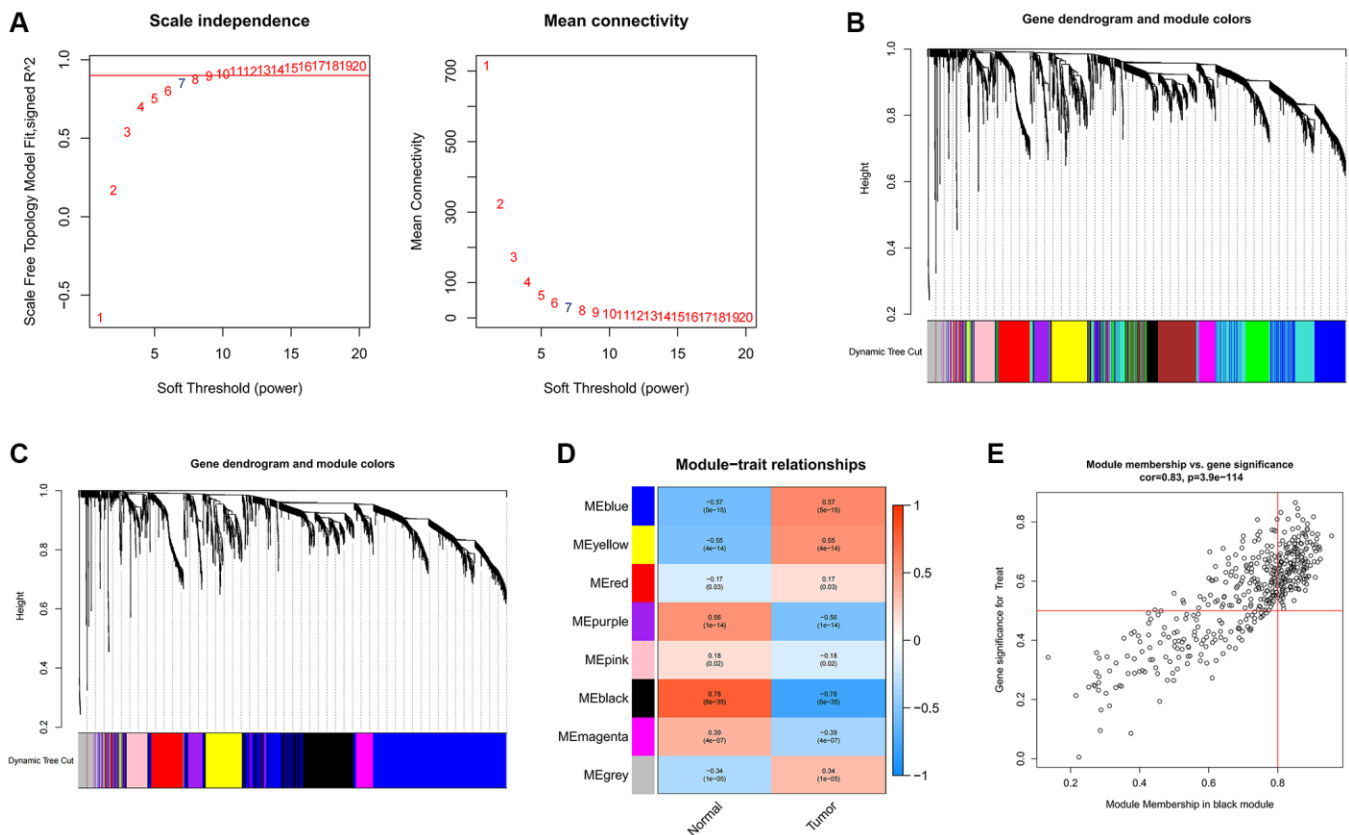


Figure 2. Identification of hub genes in gastric cancer using the WGCNA algorithm. (A) The soft threshold is determined by function. The left panel shows the relationship between the scale-free network evaluation metric R^2 and the soft threshold, and the right panel shows the relationship between average connectivity and the soft threshold. (B) Dendrogram of gene clustering and different colored modules. (C) Gene clustering dendrogram obtained by merging similar modules. (D) Correlations between merged modules and clinical traits, correlation coefficients and p-values are shown in the corresponding modules of different colors. (E) Scatterplot of MM and GS in the black module.

For the GSE65801 dataset, the AUC values for BANF1, ADH7, and TMEM27 were 0.737, 0.902, and 0.897 respectively (Figure 4D). In the TCGA dataset, AUC values of 0.878, 0.854, and 0.617 were assigned to BANF1, ADH7, and TMEM27, respectively (Figure 4F).

Prognostic evaluation of three characterizing genes

To assess the prognostic implications of BANF1, ADH7, and TMEM27, we utilized the Kaplan–Meier plotter database to construct survival curves. Compared with patients with low BANF1 expression, GC patients with high BANF1 expression level had significantly lower overall survival (OS), progression-free survival (FP), and post-progression survival (PPS) than GC patients with low BANF1 expression. (Figure 5A–5C, $p < 0.001$). Patients in the ADH7 high-expression cohort were distinctly associated with significantly lower OS and PPS compared to those in the low-expression group (Figure 5D, 5E, $p < 0.01$). However, there was no significant difference between the expression of ADH7 and FP in GC patients (Figure 5F, $p > 0.05$). Intriguingly, TMEM27 expression levels showed no substantive disparities in OS, FP, or PPS between the high- and low-expression groups (Figure 5G–5I, $p > 0.05$).

Expression of BANF1 in different cells and its relationship with tumor immunity

We obtained single-cell RNA sequencing (scRNA-seq) datasets, specifically GSE167297 and GSE134520, from Tumor Immune Single Cell Center (TISCH) database. Using the t-distributed stochastic neighbor embedding (t-SNE) algorithm, we conducted a comprehensive clustering analysis, resulting in 12 distinct cell clusters within the GSE167297 dataset. In the GSE134520 dataset, a total of 14 cell clusters were delineated. Subsequent cell annotations were anchored in the expression patterns of diverse cellular feature genes. Notably, the uniform manifold approximation and projection (UMAP) plot revealed the ubiquity of BANF1 expression across all cell types, as shown in Figure 6A, 6B. Gene set enrichment analysis (GSEA) showed that BANF1 expression was associated with the cell cycle, Huntington's disease, olfactory transduction, spliceosomes, and systemic lupus erythematosus (Supplementary Figure 2). Subsequent correlation analysis between BANF1 expression data and immune infiltration data revealed a noteworthy negative correlation. Specifically, BANF1 expression exhibited an inverse relationship with the presence of various immune cell types, including Tcm, Mast cells, B cells,

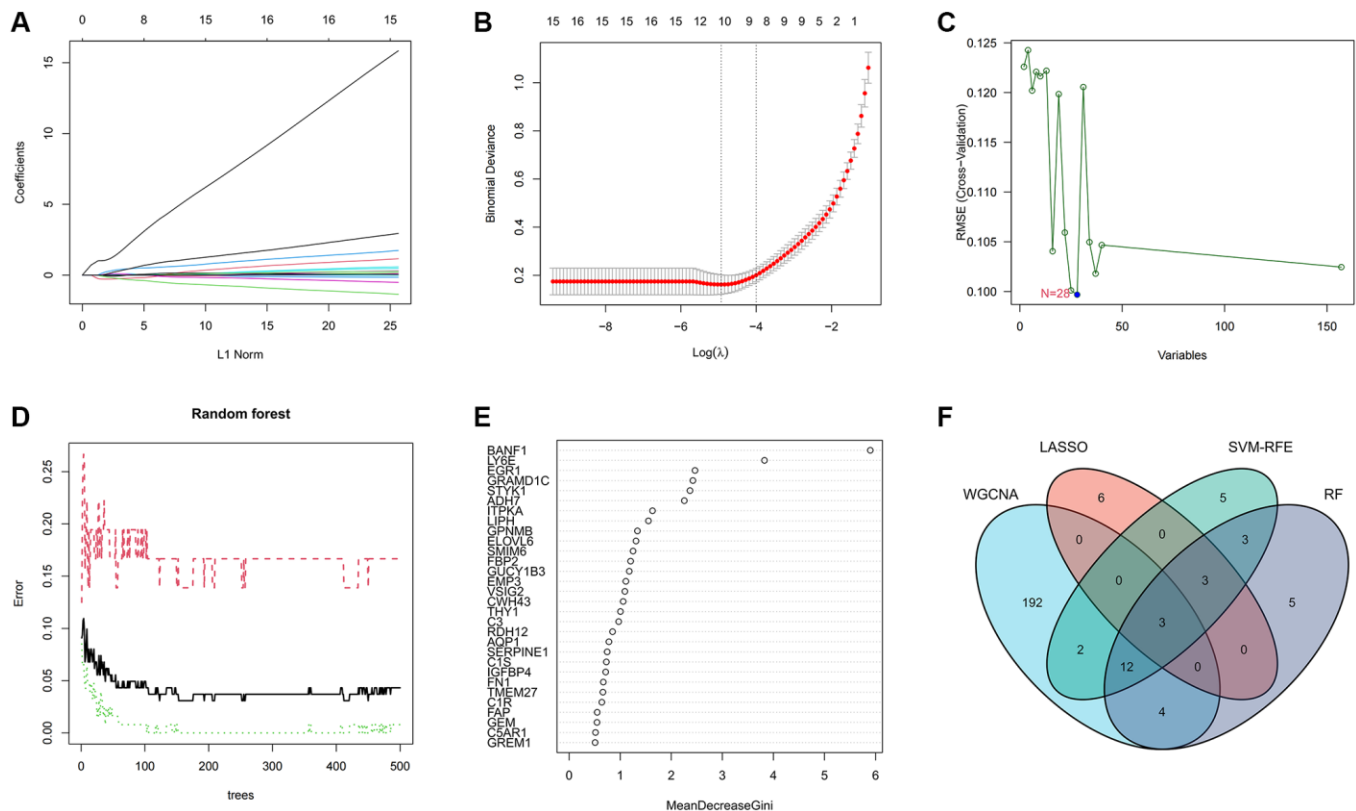


Figure 3. Machine learning screens for disease characterizing genes. (A, B) LASSO regression analysis screening variables. (C) Cross-validation error rate graph based on SVM-RFE. (D, E) Genes were scored using a random forest algorithm to rank genes by importance algorithm. (F) VENN graph to obtain the intersection of key genes screened by the 4 methods.

Tem, pDC, TFH, Eosinophils, CD8+ T cells, T helper cells, T cells, and NK cells, within GC tissues (Figure 6C; $p < 0.05$). In the realm of tumor microenvironment (TME) analysis, we ascertained that the immune score, stromal score, and Estimation of Stromal and Immune cells in Malignant Tumor tissues using Expression data (ESTIMATE) score in the BANF1 high-expression cohort markedly lagged those in the BANF1 low-expression group, as exemplified in Figure 6D ($p < 0.05$). Moreover, the investigation of disparities in immune infiltration between the high and low BANF1 expression groups underscored a reduction in the abundance of various immune cell populations, including B cells, Eosinophils, Mast cells, pDC, T helper cells, Tcm, Tem, and TFH, within the high BANF1 expression group, as elucidated in Figure 6E ($p < 0.05$). In addition, we investigated the effect of

BANF1 on immunotherapy. The results showed that BANF1 was significantly associated with immunotherapy efficacy in the IMvigor210 cohort (Supplementary Figure 3, $p < 0.01$).

BANF1 expression is elevated in GC cell lines and tissues

We conducted a thorough analysis of BANF1 expression across all cancer types within the TCGA database. The results revealed a significant upregulation of BANF1 expression in various cancers, totaling 15, including GC, when analyzing unpaired samples, as depicted in Figure 7A. A parallel analysis of paired samples showed a similar pattern, with BANF1 expression significantly increased in 13 cancers, once again including GC, as shown in Figure 7B. Examining

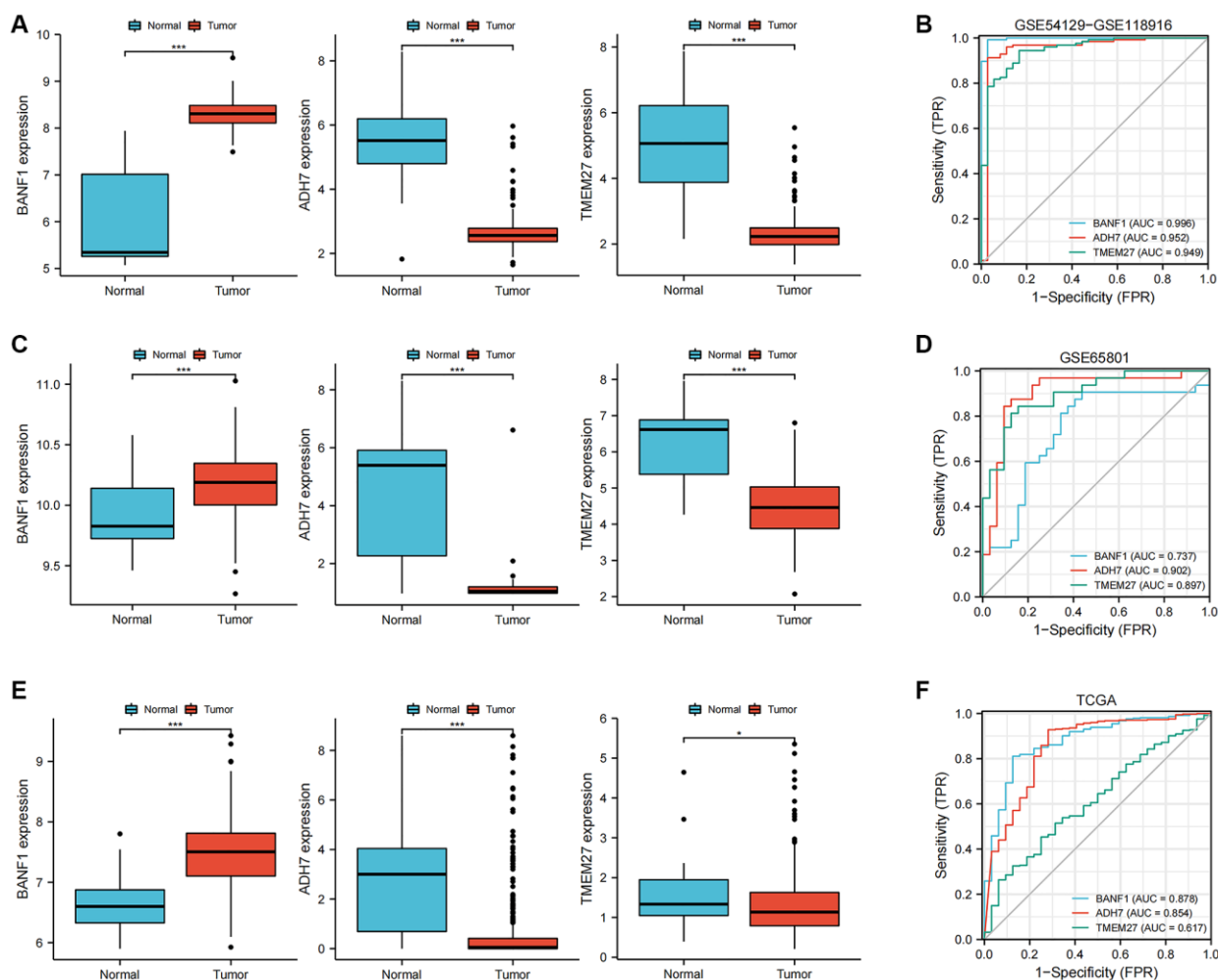


Figure 4. Expression and diagnostic value of characterized genes. (A) The box plots showed the expression of three GC characteristic genes (BANF1, ADH7, TMEM27) in GSE54129-GSE118916 dataset. (B) ROC curves were used to evaluate the diagnostic efficacy of three GC characteristic genes (BANF1, ADH7, TMEM27) in GSE54129-GSE118916 dataset. (C) The box plots showed the expression of three GC characteristic genes (BANF1, ADH7, TMEM27) in GSE65801 dataset. (D) ROC curves were used to evaluate the diagnostic efficacy of three GC characteristic genes (BANF1, ADH7, TMEM27) in GSE65801 dataset. (E) The box plots showed the expression of three GC characteristic genes (BANF1, ADH7, TMEM27) in TCGA dataset. (F) ROC curves were used to evaluate the diagnostic efficacy of three GC characteristic genes (BANF1, ADH7, TMEM27) in TCGA dataset. * $p < 0.05$, *** $p < 0.001$.

mRNA and protein expression levels, our investigation revealed a marked increase in BANF1 within cancer tissues compared to para-cancerous tissues, as demonstrated in Figure 7C ($n = 23$; $p < 0.001$) and Figure 7D ($n = 10$; $p < 0.001$). Furthermore, both the mRNA and protein expression levels of BANF1 were conspicuously elevated in various GC cell lines, namely AGS, HGC-27, MKN-45, MGC-803, and BGC-823, in comparison to a Gastric mucosal epithelial cell line (GES-1), as illustrated in Figure 7E, 7F ($p < 0.05$). The expression level of BANF1 was the highest in MKN-45

and BGC-823 cell lines. Immunofluorescence staining on MKN-45 and BGC-823 cell lines delineated the predominant localization of BANF1 protein within the nucleus, with limited presence in the cytoplasm, as illustrated in Figure 7G.

Knockdown of BANF1 inhibits proliferation, migration, and invasion of GC cells

Silencing BANF1 remarkably impedes the proliferation, migration, and invasion of GC cells. To examine the

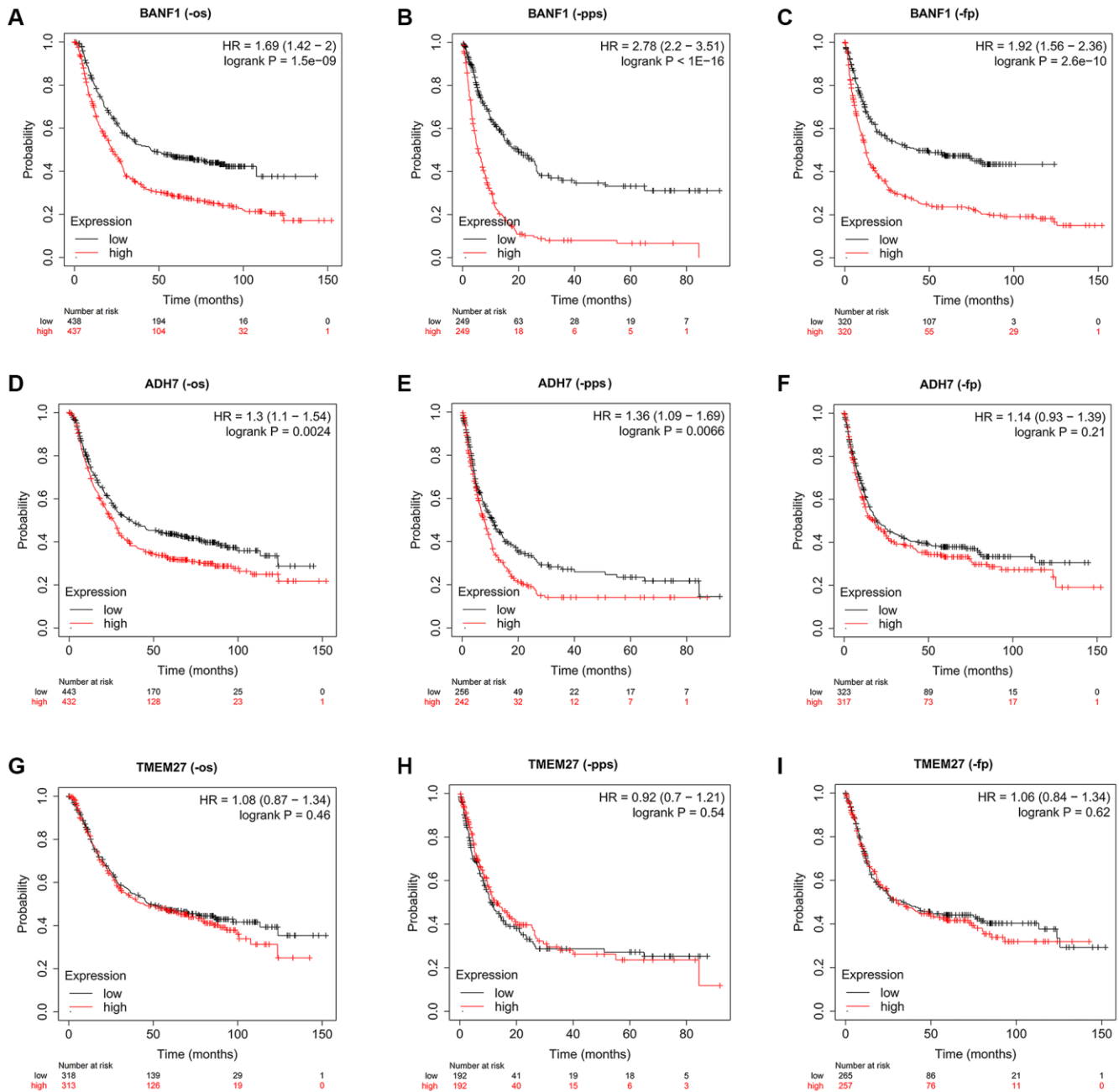


Figure 5. The prognostic value of three characteristic genes. (A–C) OS, PPS, FP km survival curves between high and low expression groups of BANF1. (D–F) OS, PPS, FP km survival curves between high and low expression groups of ADH7. (G–I) OS, PPS, FP km survival curves between high and low expression groups of TMEM27.

effect of BANF1 on the phenotypic attributes of these cells, we used lentiviral vectors to establish stable cell lines with downregulated BANF1 expression. This intervention resulted in a substantial reduction in both

the mRNA ($p < 0.001$) and protein levels of BANF1 in MKN-45 and BGC-823 cell lines, as shown in Figure 8A, 8C. Furthermore, a cell viability assay decisively underscored the diminished proliferative capacity of

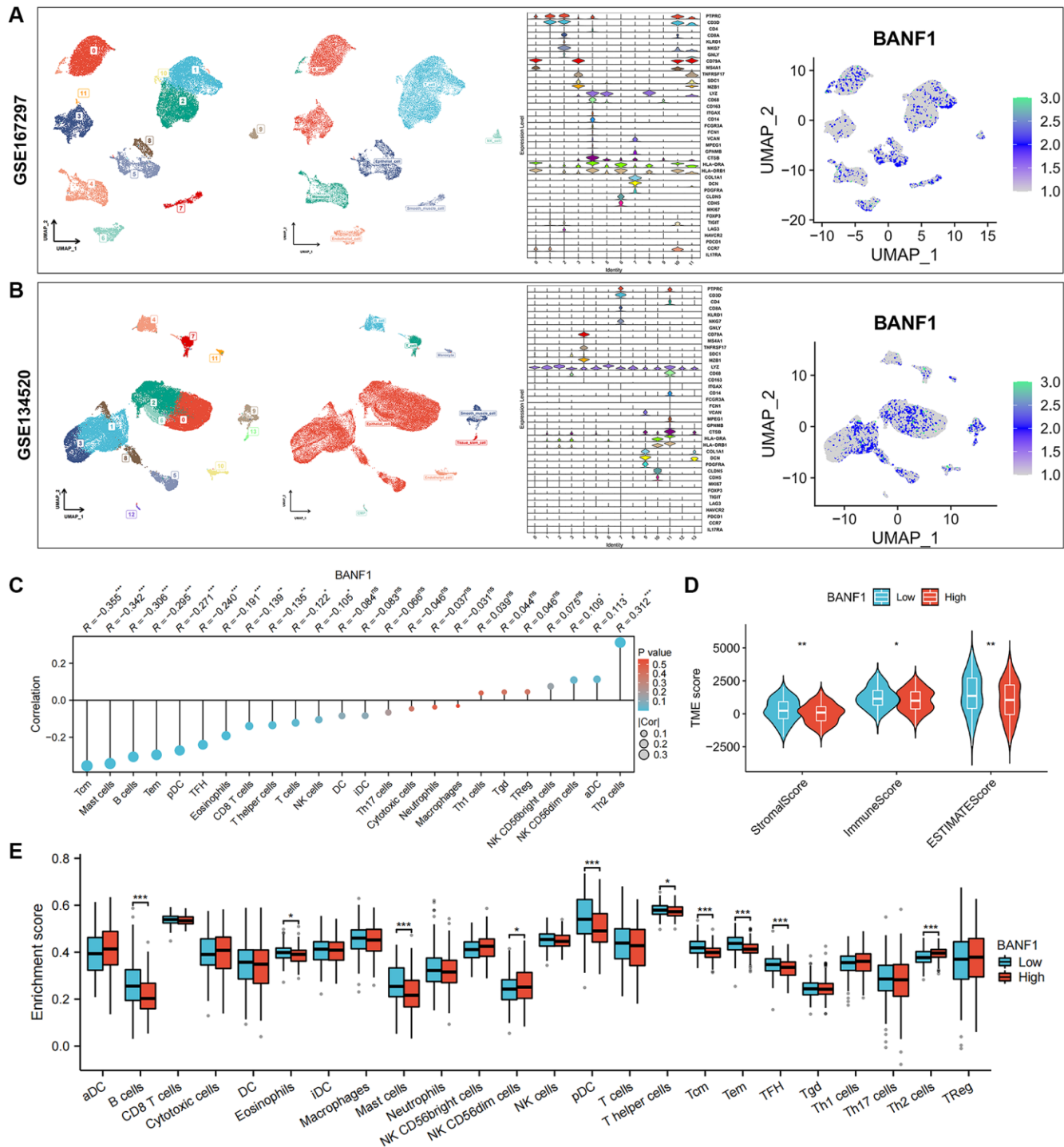


Figure 6. Single-cell expression of BANF1 in gastric cancer and its immune correlation. (A, B) Validation of BANF1 expression in different cell types in two gastric cancer single cell datasets (GSE167297, GSE134520). From left to right, cell clustering plot, cell annotation plot, violin plot of cellular signature genes expressed in different cell clusters, and BANF1 expression plot in different cells. (C) Lollipop plot of BANF1 expression correlating with immune cells. (D) Stroma score, immune score, ESTIMATE score violin plots between high and low BANF1 expression groups. (E) Box line plot showing immune infiltration between high and low BANF1 expression groups. ns means no statistical difference, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

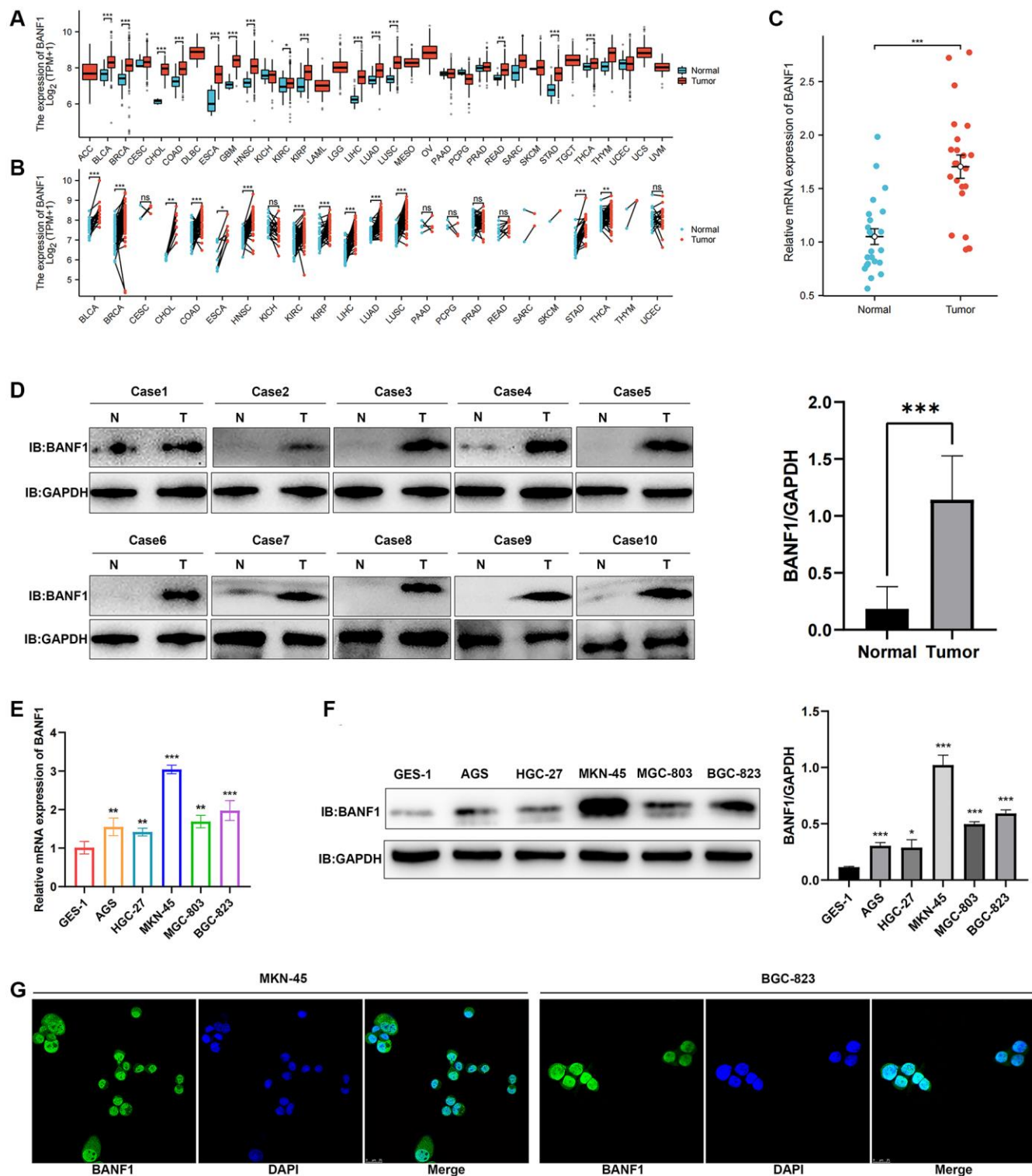


Figure 7. Expression and localization of BANF1. (A) BANF1 expression among pan-cancer unpaired samples in the TCGA database. (B) BANF1 expression among paired samples of pan-cancer in the TCGA database. (C) RT-qPCR detection of BANF1 mRNA expression in 23 pairs of GC tissues and paired paracancerous tissues. (D) WB detection of BANF1 protein expression in 10 pairs of GC tissues, and the difference of gray values between gastric cancer tissues ($n = 10$) and adjacent tissues ($n = 10$) were compared. (E) RT-qPCR was used to detect the expression of BANF1 mRNA in GC cell lines (AGS, HGC-27, MKN-45, MGC-803, BGC-823) and gastric mucosal epithelial cell line (GES-1), respectively. (F) The expression of BANF1 protein in 6 cell lines was detected by WB, and the gray values of gastric cancer cell lines (AGS, HGC-27, MKN-45, MGC-803, and BGC-823) and gastric epithelial cells (GES-1) were compared. (G) Immunofluorescence staining showed the expression and localization of BANF1 protein in MKN-45 and BGC-823 cells. All experiments were repeated at least three times. ns means no statistical difference, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

MKN-45 and BGC-823 cells after BANF1 knockdown in contrast to their blank control counterparts, as shown in Figure 8B, 8D ($p < 0.001$). The plate clone formation assay also showed that the knockdown of BANF1 inhibited the proliferation of GC cells (Figure 8E, 8F; $p < 0.001$). Elaborating on the migratory attributes, wound healing assays unveiled a significant attenuation in the migration capability of GC cells following BANF1 knockdown, an observation vividly captured in Figure 8G, 8H ($p < 0.001$). Transwell assays further reinforced these findings, affirming that the knockdown of BANF1 resulted in a marked impediment to the migration and invasion capacities of GC cells, as shown in Figure 8I, 8J ($p < 0.001$).

Knockdown of BANF1 significantly inhibited the growth of subcutaneous tumors in nude mice

To elucidate the influence of BANF1 expression on tumorigenesis and tumor growth, we performed *in vivo* experiments using a cell line-derived xenograft (CDX) model. In this model, both the control and experimental groups received injections of MKN-45 cells treated with sh-NC and sh-BANF1, respectively, enabling a meticulous assessment of subcutaneous tumor development. The results distinctly indicated a noteworthy suppression of tumor growth after BANF1 knockdown, with the knockdown group showing substantially reduced tumor volume and weight in

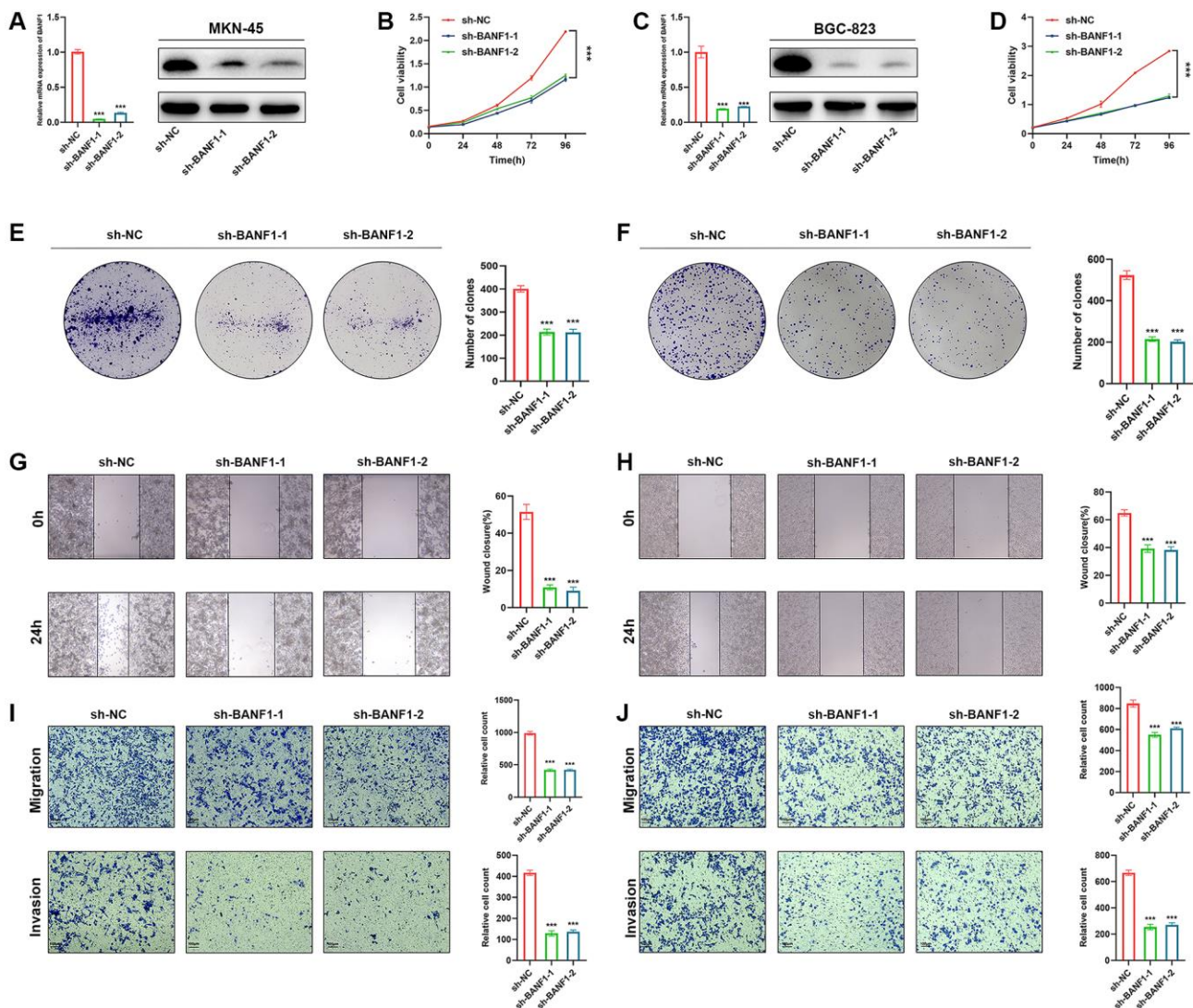


Figure 8. Knockdown of BANF1 inhibits the proliferation, migration and invasion ability of GC cells in vitro. (A) Knockdown efficiency of MKN-45 cell line was detected using RT-qPCR and WB. (B) CCK8 assay was used to detect the effect of BANF1 knockdown on the viability of MKN-45 cell line. (C) Knockdown efficiency of BGC-823 cell line was detected using RT-qPCR and WB. (D) CCK8 assay was used to detect the effect of BANF1 knockdown on the viability of BGC-823 cell line. (E, F) The effect of BANF1 knockdown on the proliferative capacity of MKN-45 and BGC-823 cells was detected by plate clone formation assay. (G, H) Effects of BANF1 knockdown on the migration ability of MKN-45 and BGC-823 cells were detected by wound healing assay. (I, J) Effect of BANF1 knockdown on the migration and invasion of MKN-45 and BGC-823 cells by Transwell assay. All experiments were repeated at least three times. *** $p < 0.001$.

comparison to the control group, as depicted in Figure 9A–9D ($n = 6$). Furthermore, immunohistochemical staining of mouse tumors provided deeper insights into the influence of BANF1 on tumor proliferative capacity. The results unequivocally revealed a conspicuous reduction in Ki67 expression within the BANF1 knockdown group as opposed to that in the control group, as shown in Figure 9E. Subsequently, the terminal deoxynucleotidyl transferase (TdT)-mediated dUTP nick-end labeling (TUNEL) assay was deployed to examine the impact of BANF1 on apoptosis in GC cells, where the findings showcased a markedly heightened level of tumor apoptosis after BANF1 knockdown relative to the control group, as depicted in Figure 9F.

DISCUSSION

The incidence and fatality rates of GC remain high, often leading to delayed surgical interventions due to late-stage diagnoses, resulting in unfavorable prognoses [13]. The current therapeutic landscape for GC emphasizes personalized and precision-oriented

approaches, including surgery, chemotherapy, radiotherapy, targeted therapy, and immunotherapy, all aimed at improving patient outcomes [14]. GC is a highly heterogeneous disease, leading to varying treatment responses among patients [15]. The molecular subtype-based classification has paved the way for individualized therapeutic strategies, offering a promising avenue for identifying novel therapeutic targets [16].

Various analytical techniques, such as WGCNA, LASSO, SVM-RFE, and RF, have been widely employed to explore disease markers and potential therapeutic targets. In a study by Zhang et al. [17], the WGCNA and RF algorithms were used to identify marker genes in acute pancreatitis. Feng et al. [18] utilized the LASSO algorithm to identify key genes for constructing a diagnostic model for coronary artery disease. A recent study [19] employed LASSO and SVM methods to identify metastatic biomarkers in colorectal cancer. These studies utilized either one or both machine learning algorithms, and the results validated the effectiveness and reliability of these

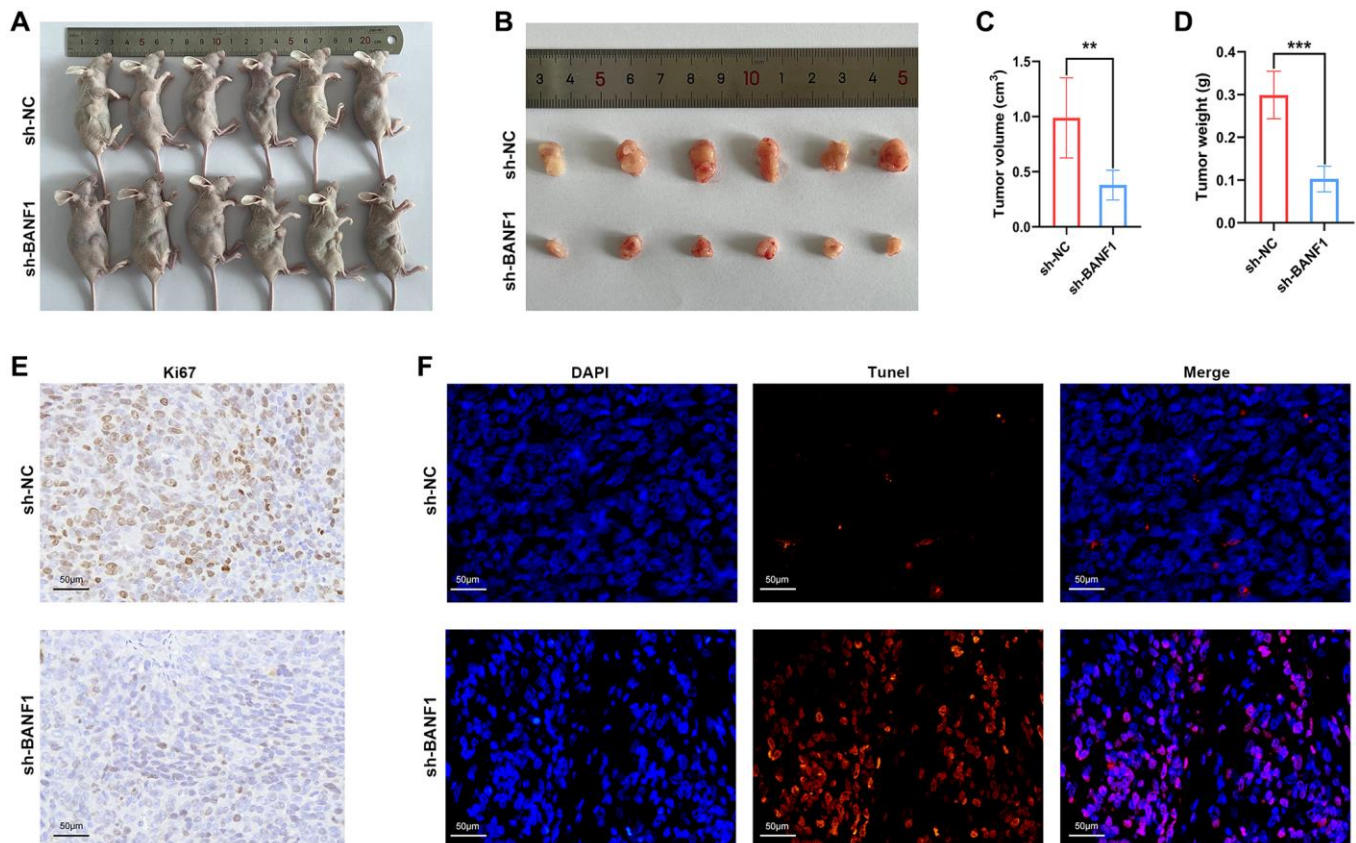


Figure 9. Knockdown of BANF1 inhibited tumor growth in in vivo experiments. (A) Photographs of nude mice injected with BANF1 knockdown MKN-45 ($n = 6$) and blank control MKN-45 cells ($n = 6$). (B) Photographs of subcutaneous tumors of nude mice in the knockdown group ($n = 6$) and control group ($n = 6$). (C, D) Comparison of subcutaneous tumor volume and weight of nude mice in knockdown and control groups. (E) Immunohistochemical staining of Ki67 in subcutaneous tumors of nude mice. (F) TUNEL staining of subcutaneous tumors in nude mice. All experiments were repeated at least three times. $**p < 0.01$, $***p < 0.001$.

approaches. To identify new ways to treat GC, we jointly applied these four algorithms in an exhaustive exploration of GC transcriptome data derived from the GEO database. Ultimately, we identified three genes (BANF1, ADH7, and TMEM27). The KM plotter database, known for its extensive data and reliable results, has been widely used for studying the correlation between gene expression and the prognosis of various cancers, mainly breast, ovarian, lung and stomach cancers [20–23]. Survival analysis of the three characterized genes revealed a significant correlation between elevated BANF1 expression and poor prognosis (including OS, FP, and PPS) in patients with GC. Our study, based on TCGA database, revealed substantial overexpression of BANF1 in more than a dozen different cancer types, including GC. Zhang et al. [10] showed that BANF1 expression is elevated in breast cancer and is correlated with lymph node metastasis. Jin et al. [11] showed that BANF1 expression is elevated in esophageal cancer. We also verified the high expression of BANF1 in GC tissues and cells, which is consistent with the results of previous bioassay analyses. Mao et al. [12] showed that inhibition of BANF1 in cervical cancer inhibited the proliferation, migration, and invasion of cervical cancer cells, and Ren et al. [24] showed that inhibition of BANF1 inhibited the proliferation and migration of esophageal cancer cells. Recent clinical studies have shown that BANF1 expression is elevated in patients with pancreatic cancer and short survival times [25]. There are very few studies on the mechanism of action of BANF1 in cancer, and most of the studies have only included expression- and cellular-level studies. Our study demonstrated that BANF1 inhibition significantly suppressed the proliferation, migration, and invasion of GC cells. In addition, we performed animal experiments and showed that BANF1 knockdown significantly inhibited the growth of subcutaneous tumors in nude mice, further demonstrating the pro-tumorigenic role of BANF1 in GC. However, the exact molecular mechanism of BANF1 involvement in cancer remains unclear. We analyzed BANF1 expression and immune infiltration in GC and found that BANF1 expression negatively correlated with immune infiltration. Immune infiltration in the TME plays a critical role in tumor development and affects patient prognosis; an effective immune response can inhibit tumor progression [26]. This may also be one of the reasons for the poor prognosis in patients with high BANF1 expression.

Immunotherapy is a highly effective approach to cancer treatment. However, its clinical efficacy varies due to individual differences and the development of drug resistance. Cyclic GMP-AMP synthase (cGAS), a key player in the stimulator of the interferon genes (STING)

pathway, has gained widespread attention for its critical role in promoting antitumor immune responses [27–30]. The cGAS-STING pathway enhances the cytotoxicity of immune effectors, including T and natural killer cells, by initiating downstream signaling events, particularly type I interferons. This activation significantly improves the efficacy of immunotherapy [31]. The exposure of large amounts of nuclear DNA to the cytoplasm and its interaction with cGAS effectively trigger an immune response. BANF1, vital for maintaining the structural integrity of the nuclear membrane, competitively binds to cGAS [32]. Consequently, BANF1 exerts a regulatory effect on the cGAS-STING pathway. Recent studies have shown that the upregulation of BANF1 results in reduced expression of cGAS proteins, hindering the innate immune response through the cGAS-STING cascade. Conversely, inhibiting BANF1 upregulates cGAS protein expression, thereby triggering an antiviral immune response [33]. We investigated the effect of BANF1 on immunotherapy efficacy (Supplementary Figure 3) and showed that there was a significant difference in the expression of BANF1 in the response and non-response groups in the iMvigor210 cohort. This suggests that BANF1 may be an important target for modulating the effects of immunotherapy. A recent study exploring the role of BANF1 in tumor immunity revealed that the knockdown of BANF1 increased immune cell infiltration into the tumor micro-environment. This resulted in a significant inhibition of melanoma and colon tumor growth in immunocompetent mice. Moreover, BANF1 knockdown enhanced the therapeutic effect of anti-PD-1 [34]. Despite these findings, the precise role of BANF1 in GC immunotherapy remains unknown. Conventional therapies have demonstrated limited efficacy in treating middle- and advanced-stage GC, prompting the exploration of immunotherapy as a viable alternative [35]. Clinical observations underscore the potential of immune checkpoint inhibitors as stand-alone or combination therapy options, leading to improved prognostic outcomes in selected patient cohorts [36, 37]. Our analysis, based on the TCGA database, indicated a negative correlation between elevated BANF1 expression in GC tumor tissues and immune cell infiltration within these tissues. This suggests that increased BANF1 expression may inhibit anti-tumor immune responses within GC. Combining these findings with existing literature, targeting BANF1 has emerged as a promising avenue to enhance GC responsiveness to immunotherapeutic interventions.

Our investigation has encountered several limitations. First, we have not fully elucidated the precise molecular mechanisms that govern the inhibition of GC progression following BANF1 knockdown. Second, the

limited availability of clinical specimens has impeded our ability to conduct a comprehensive analysis of the relationship between BANF1 expression and clinical parameters. Lastly, the effectiveness and safety of therapeutic interventions targeting BANF1 in the clinical milieu await rigorous empirical verification. In summary, despite these limitations, our study presents a promising avenue for the prospective treatment of GC through the strategic targeting of BANF1.

CONCLUSION

Our study used WGCNA in conjunction with three distinct machine-learning algorithms to identify a triad of signature genes implicated in the pathogenesis of GC. These genes—BANF1, ADH7, and TMEM27—emerged as crucial players in this disease. Notably, BANF1 exhibits heightened expression levels in GC and is closely associated with an unfavorable prognosis in individuals affected by this malignancy. Additionally, our study provides evidence of the impact of BANF1 on the proliferation and migratory capacity of GC cells. This evidence is drawn from both *in vivo* and *in vitro* experimental paradigms. However, it is important to emphasize that the mechanistic intricacies underlying the role of BANF1 in the genesis and progression of GC require further comprehensive elucidation in future studies.

MATERIALS AND METHODS

GC transcriptome data and tissue samples

Three high-throughput sequencing datasets about GC (GSE54129, GSE118916, and GSE65801) were acquired from the GEO repository (<https://www.ncbi.nlm.nih.gov/geo/>), while cancer transcriptome data from the TCGA database were procured using the GDC download tool. The data were meticulously organized using the R software (v4.1.2). Twenty-three paired fresh GC tissues and corresponding paracancerous tissues were surgically excised and collected under the auspices of the Medical Ethics Committee of the First Affiliated Hospital of Anhui Medical University. The samples were preserved in liquid nitrogen for subsequent use.

Identify differentially expressed genes

Transcriptome data encompassing a cohort of 162 GC samples were meticulously curated through a process involving the integration and harmonization of batch effects within the GSE54129 and GSE118916 datasets using the sva R package. PCA was performed, and the results were visualized using the ggplot2 package. To elucidate the intricate disparities in molecular

expression patterns between tumor and normal tissues, we subjected the transcriptome data to rigorous analysis using the Limma package. Differential gene expression profiling was performed, adhering to stringent criteria, with a focus on genes exhibiting a $|\log_2(\text{fold change (FC)})| > 2$ and a false discovery rate (FDR) < 0.01 . Visualization of the DEGs within GC tissues was accomplished by generating heat maps using the pheatmap R package.

Weighted correlation network analysis

WGCNA is a pivotal biological approach used to explore gene interrelationships across multiple samples [38–40]. This method discerns clusters of genes that exhibit high inter-correlations, thus forming cohesive modules. Moreover, we aimed to identify biomarkers and prospective therapeutic targets by scrutinizing the correlation between these gene modules and the disease phenotype. Gene expression data were meticulously processed using the WGCNA R package, culminating in the creation of a heat map illustrating the interplay between gene modules and clinical information. Among the diverse spectrum of gene modules identified, paramount attention was paid to the one exhibiting the highest correlation with GC, designating it as the pivotal module. Subsequently, hub genes were methodically sieved in accordance with stringent criteria, whereby they were mandated to possess a GS exceeding 0.5, and MM surpassing the threshold of 0.8.

Identification of hub genes related to GC

Machine learning, with its capacity to adeptly navigate intricate genomic data and discern more precise associations with diseases, offers an efficient avenue for exploring disease-related genes. In this study, we harnessed a trio of machine learning algorithms, namely, LASSO regression analysis, Random Forest (RF), and Support Vector Machine Recursive Feature Elimination (SVM-RFE) [41–43], to conduct an additional layer of scrutiny of the previously identified DEGs. The ensuing step involved the convergence of the key genes identified through WGCNA with a set of differential genes gleaned through the application of the three machine-learning algorithms. This amalgamation yielded a distinctive ensemble of genes that are characteristic of GC.

Evaluation of characterization genes

To assess the predictive performance of the three identified genes in diagnosing GC, we conducted an ROC analysis of the transcriptomic data. This analysis was performed using the pROC package, with the AUC serving as the benchmark for evaluating diagnostic

efficacy. Subsequently, we investigated the prognostic relevance of these three genes in the context of GC using the Kaplan–Meier plotter database (<https://kmplot.com/analysis>). Patients were stratified into high- and low-expression groups based on median gene expression values and survival curves, encompassing OS, FP, and PPS, which were constructed. Differences in survival outcomes between the two groups were rigorously scrutinized using the log-rank test for statistical comparison.

Single-cell RNA sequencing data analysis and immunocorrelation analysis

Retrieve the scRNA-seq datasets pertinent to GC from the TISCH (database <http://tisch.comp-genomics.org/>). Employing the “FindClusters” function with a resolution parameter set at 0.2, we ascertained cellular clusters within the datasets. Subsequently, we visualized the data by employing t-SNE, we visualized the data. To identify the DEGs within each of these clusters, we leveraged the FindAllMarkers algorithm and selected the top 10 genes with the most significant differences as markers for each cell cluster. Cell annotation was performed manually using these markers. For data dimensionality reduction analysis, we opted for the UMap method and visualized the expression of BANF1 across various cell types. Furthermore, in the TCGA GC samples, we computed immune infiltration using 24 immune cell markers and the single sample gene set enrichment analysis (ssGSEA) algorithm in the GSVA R package. The Spearman method was employed to analyze the correlation between BANF1 expression and the degree of immune infiltration, and the results were depicted as a Laplace plot using the ggplot2 R package. We categorized the samples into high and low BANF1 expression groups based on the median BANF1 expression value, and compared immune infiltration between these groups using the Wilcoxon rank-sum test. To gauge the impact of BANF1 expression on the immune microenvironment, we calculated the immune, stromal, and ESTIMATE scores for GC samples using the ESTIMATE R package. Higher scores indicated a greater abundance of these components within the tumor microenvironment. Differences in scores between the groups characterized by high and low BANF1 expression were assessed using the Wilcoxon rank-sum test. The renal cell carcinoma and melanoma immunotherapy cohorts were obtained from the GEO website, the urothelial carcinoma dataset was downloaded from the IMvigor210 (<http://research-pub.gene.com/IMvigor210CoreBiologies>), and the gastric cancer immunotherapy cohort (PRJEB25780) was obtained from Tumor Immune Dysfunction and Exclusion (TIDE). The patients were divided into response group and non-response group, and the

difference of BANF1 expression between the two groups was compared.

Gene set enrichment analysis

To identify the enriched pathways associated with BANF1 expression, GSEA was performed on the BANF1 high and low expression groups using GSEA software (v4.2.1) and gene set (C2. Cp.kegg. V7.4. Symbols GMT). The top five pathways were selected based on the $p < 0.05$ criterion.

Pan-cancer expression analysis of BANF1

Data retrieval and curation involved the acquisition of information from TCGA database (<https://portal.gdc.cancer.gov/>), encompassing transcript data expressed in transcripts per kilobase million (TPM) format for 33 tumor types. Groups consisting of fewer than three samples or those displaying zero variance were excluded from subsequent analyses. To discern disparities in BANF1 expression, the Wilcoxon rank-sum test was used to compare the expression between the cancer and paracancer groups.

RNA extraction and real-time quantitative polymerase reaction

In the investigation of GC tissues and adjacent non-tumor tissues, total RNA extraction was performed utilizing the HiPure Universal RNA Kit (Magen, Shanghai, China), following the manufacturer’s guidelines. Subsequently, the concentration and purity of the RNA samples were assessed employing a Nanodrop 2000 (Thermo Fisher Scientific, USA). The transcribed cDNA was generated from the extracted RNA using the PrimeScript RT kit (Vazyme, Nanjing, China). The quantification of BANF1 mRNA expression was conducted through the $2^{-\Delta\Delta CT}$ method, utilizing GAPDH as an internal control.

Western blotting

Tissues and cells underwent lysis using a phenyl-methanesulfonyl fluoride (PMSF)-containing lysis buffer (Beyotime, Shanghai, China) to extract proteins. After electrophoresis and membrane transfer, the proteins were sealed with 5% skim milk for 1.5 hours, and the membrane was incubated with anti-BANF1 (1:1000, Abcam, UK) or anti-GAPDH (1:50000, Proteintech, China) at 4°C overnight. After washing with Tris-buffered saline containing 0.1% Tween 20 (TBST), the membranes were immersed in anti-rabbit/mouse secondary antibodies (1:10000, Proteintech, China) for 2 hours at room temperature. Finally, membrane visualization and imaging were

achieved using an ECL chemiluminescence system (Tanon, Shanghai, China).

Immunohistochemistry

Immunohistochemical staining was used to detect ki67 expression in subcutaneous tumors from nude mice. Samples were embedded in paraffin to prepare paraffin sections, which were dewaxed, hydrated, and antigenically repaired according to standard procedures. Sections were titrated with anti-ki67 antibody (1:1000, CST, USA) and incubated at 4°C overnight. After washing, enzyme-labeled sheep anti-rabbit IgG polymer was added dropwise and incubated at 37°C for 20 min. Incubate with freshly prepared diaminobenzidine chromogenic solution and finally with hematoxylin staining solution.

Immunofluorescence

The assessment of BANF1 protein expression and subcellular localization was conducted in MKN-45 and BGC-823 cells. Following fixation with 4% paraformaldehyde, cells were subjected to overnight incubation with the BANF1 primary antibody at 4°C. Subsequent to thorough washing, Fluorescent secondary antibody was applied and incubated at room temperature for 1 hour. The cellular nuclei were counterstained with 4',6-diamidino-2-phenylindole (DAPI) for a duration of 10 minutes.

Cell culture and stabilized cell line construction

Human gastric cancer cell lines (AGS, HGC-27, MKN-45, MGC-803, and BGC-823) and the normal gastric mucosal epithelial cell line GES-1 were sourced from the American Type Culture Collection (ATCC). The cells were cultured in 1640 medium (Gibco), supplemented with fetal bovine serum and penicillin-streptomycin. Incubation occurred at 37°C with a CO₂ concentration maintained at 5%. For transfection, GC cells and co-transfection reagents (GeneChem, Shanghai, China) were added to 12-well plates at specific ratios. Notably, MKN-45 and BGC-823 cells were utilized for gene knockdown experiments. Stably expressing cells were subsequently selected through screening with 5 µg/ml puromycin.

Cell proliferation capacity assay

The cell proliferative capacity was determined using a cell counting kit-8 (CCK8) assay and a plate clone formation assay. Stably transformed cells were added to 96-well plates at 3000 cells/well. The absorbance at 450 nm was measured using an enzyme marker after adding the CCK8 reagent at 0 h, 24 h, 48 h, 72 h, and 96 h

respectively. The plate clone formation assay was performed by adding 1000 cells/well to 6-well plates and changing the medium every 3 d for a total of approximately 10 d. Cells were fixed with paraformaldehyde and stained with crystal violet after clone formation.

Wound healing assays

GC cells were cultured in 6-well plates, and when the cells were fully grown, the tip of a pipette was used to make a scratch approximately 2 mm wide. Low-serum medium containing 2% fetal bovine serum was added, and the culture was continued. Scratch healing was observed at 0 h and 24 h.

Transwell assay

Migration and invasion capabilities of gastric cancer (GC) cells were assessed using the Transwell assay. Cells were suspended in a serum-free medium, with 30,000 cells added to the upper layer of each chamber. The lower layer of the chamber received a complete medium containing 10% FBS. For the assessment of cell invasiveness using the Transwell assay, Matrigel (BD, USA) was added to the chambers beforehand, while the Transwell assay without Matrigel was employed to evaluate the migration ability of the cells. Following a 48-hour incubation period, the chambers were carefully removed, washed with PBS, fixed using 4% paraformaldehyde, and subsequently stained with a crystal violet staining solution.

Cell line-derived xenograft

The immunodeficient mice used in this study were purchased from Empharmatech (Jiangsu, China). Breeding conditions and all operational procedures were performed in accordance with the requirements of the Animal Ethics Committee of Anhui Medical University. We selected 4-week-old nude mice for experiments after acclimatization in a specific pathogen-free (SPF) environment for seven days. Stable BANF1 knockdown MKN-45 cells and MKN-45 control cells were injected subcutaneously into nude mice, and the number and volume of cells injected into each nude mouse were 5 million and 200 µl. The volume of subcutaneous tumors in the nude mice was measured every two days, and the formula for calculating the volume of the tumors was $V = ab^2/2$. The mice were sacrificed at the end of the experiment and disposed of.

TUNEL assay for apoptosis detection

The principle of apoptosis detection by TUNEL assay is that when DNA breaks during apoptosis, the exposed

3'-OH is labeled with TMR-5-dUTP catalyzed by TdT, which can be fluorescently microscopically observed. Paraffin sections of nude mouse subcutaneous tumors were stained using a one-step TUNEL apoptosis detection kit (Servicebio, Wuhan, China).

Statistical analysis

Disparities in molecular expression between tumor and normal tissues in publicly available data were assessed through the Wilcoxon rank-sum test. Kaplan–Meier survival curves were employed to illustrate the prognostic differences between high and low BANF1 expression groups, with log-rank tests utilized for inter-group comparisons. The correlation between BANF1 expression and immune cell infiltration was scrutinized using Spearman's method, and variations in immune infiltration results between high and low BANF1 expression groups were evaluated through the Wilcoxon rank-sum test. Differences in BANF1 expression across various cancer types were examined using the Wilcoxon rank-sum test. All bioinformatics analyses were executed using R software (v4.1.2). Experimental procedures were replicated a minimum of three times. Western blot (WB) results were quantified with ImageJ software, and inter-group differences were analyzed using the *t*-test in GraphPad Prism 8.02. A significance threshold of $p < 0.05$ was applied. Bar graphs depict the mean \pm standard deviation (SD) based on the outcomes of the three experiments.

Abbreviations

GC: Gastric cancer; GEO: Gene Expression Omnibus; TCGA: The Cancer Genome Atlas; BANF1: barrier to autointegration factor 1; DEGs: differentially expressed genes; PCA: principal component analysis; WGCNA: Weighted co-expression network analysis; LASSO: least absolute shrinkage and selection operator; RF: random forest; SVM-RFE: support vector machine-recursive feature elimination; ESTIMATE: Estimation of Stromal and Immune Cells in Malignant Tumors Using Expression Data; ssGSEA: single sample Gene Set Enrichment Analysis; GS: gene significance; MM: module membership; ROC: Receiver Operating Characteristic; AUC: area under the curve; OS: overall survival; FP: first progression survival; PPS: post progression survival; FC: fold change; scRNA-seq: single cell RNA sequencing; TISCH: Tumor Immune Single-cell Hub; UMAP: uniform manifold approximation and projection; GSEA: gene set enrichment analysis; TME: tumor microenvironment; t-SNE: t-Distributed Stochastic Neighbor Embedding; TPM: transcripts per kilobase million; PMSF: Phenylmethanesulfonyl fluoride; TBST: Tris buffer Saline containing 0.1% Tween 20; DAPI: 4',6-diamidino-2-

phenylindole; ATCC: American Type Culture Collection; RPMI: Roswell Park Memorial Institute; CCK8: cell counting kit-8; SPF: specific pathogen Free; TdT: Terminal Deoxynucleotidyl Transferase; TUNEL: TdT mediated dUTP Nick End Labeling; TIDE: Tumor Immune Dysfunction and Exclusion; CDX: cell line-derived xenograft; cGAS-STING: Cyclic GMP-AMP synthase (cGAS)-stimulator of the interferon genes (STING); WB: western blot.

AUTHOR CONTRIBUTIONS

KWH and WC designed this study. YMX and XW performed data analysis and experimental validation and drafted the manuscript. WWY and LZ contributed to data collection and animal husbandry. All authors contributed to revision of the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

Public data for this study were obtained from GEO and TCGA, clinical samples were obtained from the First Affiliated Hospital of Anhui Medical University, and we thank the patients and surgeons. The study was supported by the Immunology Laboratory of Anhui Medical University and we thank them for their help in our study.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest related to this study.

ETHICAL STATEMENT AND CONSENT

The GC tissue samples used in this study were obtained with informed consent from patients and approved by the Medical Ethics Committee from The First Affiliated Hospital of Anhui Medical University (approval No. Quick-PJ 2023-03-22). The Animal experiment was authorized by the Animal Ethics Review Committees of Anhui Medical University (approval No. LLSC20220216).

FUNDING

This study was funded by the Key Research and Development Project of Anhui Provincial Science and Technology Department (202004j07020036).

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185

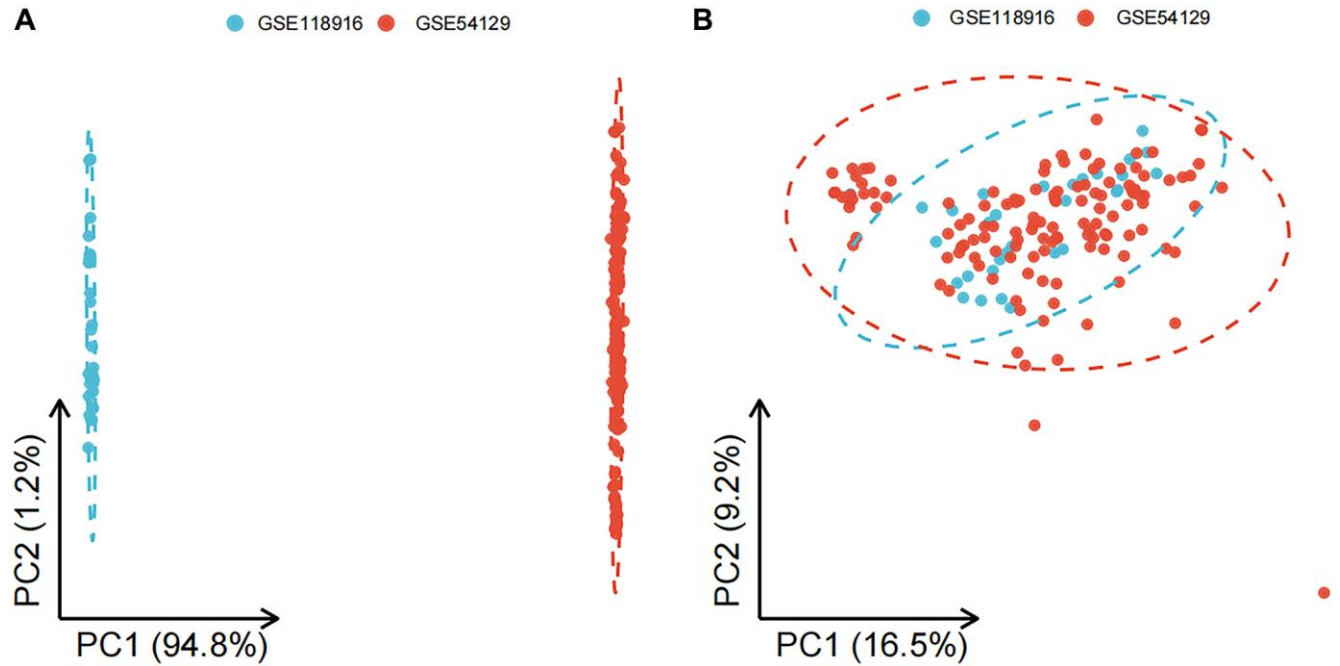
- Countries. *CA Cancer J Clin.* 2021; 71:209–49.
<https://doi.org/10.3322/caac.21660>
PMID:33538338
2. Costa L, Corre S, Michel V, Le Luel K, Fernandes J, Ziveri J, Jouvion G, Danckaert A, Mouchet N, Da Silva Barreira D, Torres J, Camorlinga M, D'Elios MM, et al. USF1 defect drives p53 degradation during *Helicobacter pylori* infection and accelerates gastric carcinogenesis. *Gut.* 2020; 69:1582–91.
<https://doi.org/10.1136/gutjnl-2019-318640>
PMID:31822580
 3. Ajani JA, D'Amico TA, Bentrem DJ, Chao J, Cooke D, Corvera C, Das P, Enzinger PC, Enzler T, Fanta P, Farjah F, Gerdes H, Gibson MK, et al. Gastric Cancer, Version 2.2022, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw.* 2022; 20:167–92.
<https://doi.org/10.6004/jnccn.2022.0008>
PMID:35130500
 4. Lin Y, Zheng Y, Wang HL, Wu J. Global Patterns and Trends in Gastric Cancer Incidence Rates (1988-2012) and Predictions to 2030. *Gastroenterology.* 2021; 161:116–27.e8.
<https://doi.org/10.1053/j.gastro.2021.03.023>
PMID:33744306
 5. Smyth EC, Nilsson M, Grabsch HI, van Grieken NC, Lordick F. Gastric cancer. *Lancet.* 2020; 396:635–48.
[https://doi.org/10.1016/S0140-6736\(20\)31288-5](https://doi.org/10.1016/S0140-6736(20)31288-5)
PMID:32861308
 6. Rose M, Burgess JT, O'Byrne K, Richard DJ, Bolderson E. The role of inner nuclear membrane proteins in tumorigenesis and as potential targets for cancer therapy. *Cancer Metastasis Rev.* 2022; 41:953–63.
<https://doi.org/10.1007/s10555-022-10065-z>
PMID:36205821
 7. Burgess JT, Cheong CM, Suraweera A, Sobanski T, Beard S, Dave K, Rose M, Boucher D, Croft LV, Adams MN, O'Byrne K, Richard DJ, Bolderson E. Barrier-to-autointegration-factor (Banf1) modulates DNA double-strand break repair pathway choice via regulation of DNA-dependent kinase (DNA-PK) activity. *Nucleic Acids Res.* 2021; 49:3294–307.
<https://doi.org/10.1093/nar/gkab110>
PMID:33660778
 8. Gorjánác M. Nuclear assembly as a target for anti-cancer therapies. *Nucleus.* 2014; 5:47–55.
<https://doi.org/10.4161/nucl.27928>
PMID:24637400
 9. Cabanillas R, Cadiñanos J, Villameytide JA, Pérez M, Longo J, Richard JM, Alvarez R, Durán NS, Illán R, González DJ, López-Otín C. Néstor-Guillermo progeria syndrome: a novel premature aging condition with early onset and chronic development caused by BANF1 mutations. *Am J Med Genet A.* 2011; 155A:2617–25.
<https://doi.org/10.1002/ajmg.a.34249>
PMID:21932319
 10. Zhang G. Expression and Prognostic Significance of BANF1 in Triple-Negative Breast Cancer. *Cancer Manag Res.* 2020; 12:145–50.
<https://doi.org/10.2147/CMAR.S229022>
PMID:32021431
 11. Li J, Wang T, Pei L, Jing J, Hu W, Sun T, Liu H. Expression of VRK1 and the downstream gene BANF1 in esophageal cancer. *Biomed Pharmacother.* 2017; 89:1086–91.
<https://doi.org/10.1016/j.biopha.2017.02.095>
PMID:28298069
 12. Mao L, Zhang Y, Mo W, Yu Y, Lu H. BANF1 is downregulated by IRF1-regulated microRNA-203 in cervical cancer. *PLoS One.* 2015; 10:e0117035.
<https://doi.org/10.1371/journal.pone.0117035>
PMID:25658920
 13. Zeng H, Ran X, An L, Zheng R, Zhang S, Ji JS, Zhang Y, Chen W, Wei W, He J, and HBCR Working Group. Disparities in stage at diagnosis for five common cancers in China: a multicentre, hospital-based, observational study. *Lancet Public Health.* 2021; 6:e877–87.
[https://doi.org/10.1016/S2468-2667\(21\)00157-2](https://doi.org/10.1016/S2468-2667(21)00157-2)
PMID:34838194
 14. Wang FH, Zhang XT, Li YF, Tang L, Qu XJ, Ying JE, Zhang J, Sun LY, Lin RB, Qiu H, Wang C, Qiu MZ, Cai MY, et al. The Chinese Society of Clinical Oncology (CSCO): Clinical guidelines for the diagnosis and treatment of gastric cancer, 2021. *Cancer Commun (Lond).* 2021; 41:747–95.
<https://doi.org/10.1002/cac2.12193>
PMID:34197702
 15. Alsina M, Arrazubi V, Diez M, Tabernero J. Current developments in gastric cancer: from molecular profiling to treatment strategy. *Nat Rev Gastroenterol Hepatol.* 2023; 20:155–70.
<https://doi.org/10.1038/s41575-022-00703-w>
PMID:36344677
 16. Joshi SS, Badgwell BD. Current treatment and recent progress in gastric cancer. *CA Cancer J Clin.* 2021; 71:264–79.
<https://doi.org/10.3322/caac.21657>
PMID:33592120
 17. Zhang D, Wang M, Zhang Y, Xia C, Peng L, Li K, Yin H, Li S, Yang X, Su X, Huang H. Novel insight on marker genes and pathogenic peripheral neutrophil subtypes in acute pancreatitis. *Front Immunol.* 2022; 13:964622.
<https://doi.org/10.3389/fimmu.2022.964622>
PMID:36072587

18. Feng X, Zhang Y, Du M, Li S, Ding J, Wang J, Wang Y, Liu P. Identification of diagnostic biomarkers and therapeutic targets in peripheral immune landscape from coronary artery disease. *J Transl Med.* 2022; 20:399.
<https://doi.org/10.1186/s12967-022-03614-1>
PMID:[36064568](https://pubmed.ncbi.nlm.nih.gov/36064568/)
19. Ahmadih-Yazdi A, Mahdavinezhad A, Tapak L, Nouri F, Taherkhani A, Afshar S. Using machine learning approach for screening metastatic biomarkers in colorectal cancer and predictive modeling with experimental validation. *Sci Rep.* 2023; 13:19426.
<https://doi.org/10.1038/s41598-023-46633-8>
PMID:[37940644](https://pubmed.ncbi.nlm.nih.gov/37940644/)
20. Zhang J, Wang H, Yuan C, Wu J, Xu J, Chen S, Zhang C, He Y. ITGAL as a Prognostic Biomarker Correlated With Immune Infiltrates in Gastric Cancer. *Front Cell Dev Biol.* 2022; 10:808212.
<https://doi.org/10.3389/fcell.2022.808212>
PMID:[35399517](https://pubmed.ncbi.nlm.nih.gov/35399517/)
21. Suresh S, Vinet M, Dakroub R, Lesage L, Ye M, Fayyad-Kazan H, Nicolas A, Meseure D, Dubois T. Expression, Localization and Prognosis Association of MEP50 in Breast Cancer. *Cancers (Basel).* 2022; 14:4766.
<https://doi.org/10.3390/cancers14194766>
PMID:[36230689](https://pubmed.ncbi.nlm.nih.gov/36230689/)
22. Wu Z, Wang Y, Li J, Wang H, Tuo X, Zheng J. MCM10 is a Prognostic Biomarker and Correlated With Immune Checkpoints in Ovarian Cancer. *Front Genet.* 2022; 13:864578.
<https://doi.org/10.3389/fgene.2022.864578>
PMID:[35664337](https://pubmed.ncbi.nlm.nih.gov/35664337/)
23. Zhao J, Cheng M, Gai J, Zhang R, Du T, Li Q. SPOCK2 Serves as a Potential Prognostic Marker and Correlates With Immune Infiltration in Lung Adenocarcinoma. *Front Genet.* 2020; 11:588499.
<https://doi.org/10.3389/fgene.2020.588499>
PMID:[33244319](https://pubmed.ncbi.nlm.nih.gov/33244319/)
24. Ren Z, Geng J, Xiong C, Li X, Li Y, Li J, Liu H. Downregulation of VRK1 reduces the expression of BANF1 and suppresses the proliferative and migratory activity of esophageal cancer cells. *Oncol Lett.* 2020; 20:1163–70.
<https://doi.org/10.3892/ol.2020.11654>
PMID:[32724356](https://pubmed.ncbi.nlm.nih.gov/32724356/)
25. Kisling SG, Atri P, Shah A, Cox JL, Sharma S, Smith LM, Ghersi D, Batra SK. A Novel HOXA10-Associated 5-Gene-Based Prognostic Signature for Stratification of Short-term Survivors of Pancreatic Ductal Adenocarcinoma. *Clin Cancer Res.* 2023; 29:3759–70.
<https://doi.org/10.1158/1078-0432.CCR-23-0825>
PMID:[37432996](https://pubmed.ncbi.nlm.nih.gov/37432996/)
26. Zhang Y, Zhang Z. The history and advances in cancer immunotherapy: understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications. *Cell Mol Immunol.* 2020; 17:807–21.
<https://doi.org/10.1038/s41423-020-0488-6>
PMID:[32612154](https://pubmed.ncbi.nlm.nih.gov/32612154/)
27. Kwon J, Bakhoun SF. The Cytosolic DNA-Sensing cGAS-STING Pathway in Cancer. *Cancer Discov.* 2020; 10:26–39.
<https://doi.org/10.1158/2159-8290.CD-19-0761>
PMID:[31852718](https://pubmed.ncbi.nlm.nih.gov/31852718/)
28. Samson N, Ablasser A. The cGAS-STING pathway and cancer. *Nat Cancer.* 2022; 3:1452–63.
<https://doi.org/10.1038/s43018-022-00468-w>
PMID:[36510011](https://pubmed.ncbi.nlm.nih.gov/36510011/)
29. Wang Y, Luo J, Alu A, Han X, Wei Y, Wei X. cGAS-STING pathway in cancer biotherapy. *Mol Cancer.* 2020; 19:136.
<https://doi.org/10.1186/s12943-020-01247-w>
PMID:[32887628](https://pubmed.ncbi.nlm.nih.gov/32887628/)
30. Jiang M, Chen P, Wang L, Li W, Chen B, Liu Y, Wang H, Zhao S, Ye L, He Y, Zhou C. cGAS-STING, an important pathway in cancer immunotherapy. *J Hematol Oncol.* 2020; 13:81.
<https://doi.org/10.1186/s13045-020-00916-z>
PMID:[32571374](https://pubmed.ncbi.nlm.nih.gov/32571374/)
31. Li A, Yi M, Qin S, Song Y, Chu Q, Wu K. Activating cGAS-STING pathway for the optimal effect of cancer immunotherapy. *J Hematol Oncol.* 2019; 12:35.
<https://doi.org/10.1186/s13045-019-0721-x>
PMID:[30935414](https://pubmed.ncbi.nlm.nih.gov/30935414/)
32. Guey B, Wischnewski M, Decout A, Makasheva K, Kaynak M, Sakar MS, Fierz B, Ablasser A. BAF restricts cGAS on nuclear DNA to prevent innate immune activation. *Science.* 2020; 369:823–8.
<https://doi.org/10.1126/science.aaw6421>
PMID:[32792394](https://pubmed.ncbi.nlm.nih.gov/32792394/)
33. Broussard G, Ni G, Zhang Z, Li Q, Cano P, Dittmer DP, Damania B. Barrier-to-autointegration factor 1 promotes gammaherpesvirus reactivation from latency. *Nat Commun.* 2023; 14:434.
<https://doi.org/10.1038/s41467-023-35898-2>
PMID:[36746947](https://pubmed.ncbi.nlm.nih.gov/36746947/)
34. Wang M, Huang Y, Chen M, Wang W, Wu F, Zhong T, Chen X, Wang F, Li Y, Yu J, Wu M, Chen D. Inhibition of tumor intrinsic BANF1 activates antitumor immune responses via cGAS-STING and enhances the efficacy of PD-1 blockade. *J Immunother Cancer.* 2023; 11:e007035.
<https://doi.org/10.1136/jitc-2023-007035>
PMID:[37620043](https://pubmed.ncbi.nlm.nih.gov/37620043/)

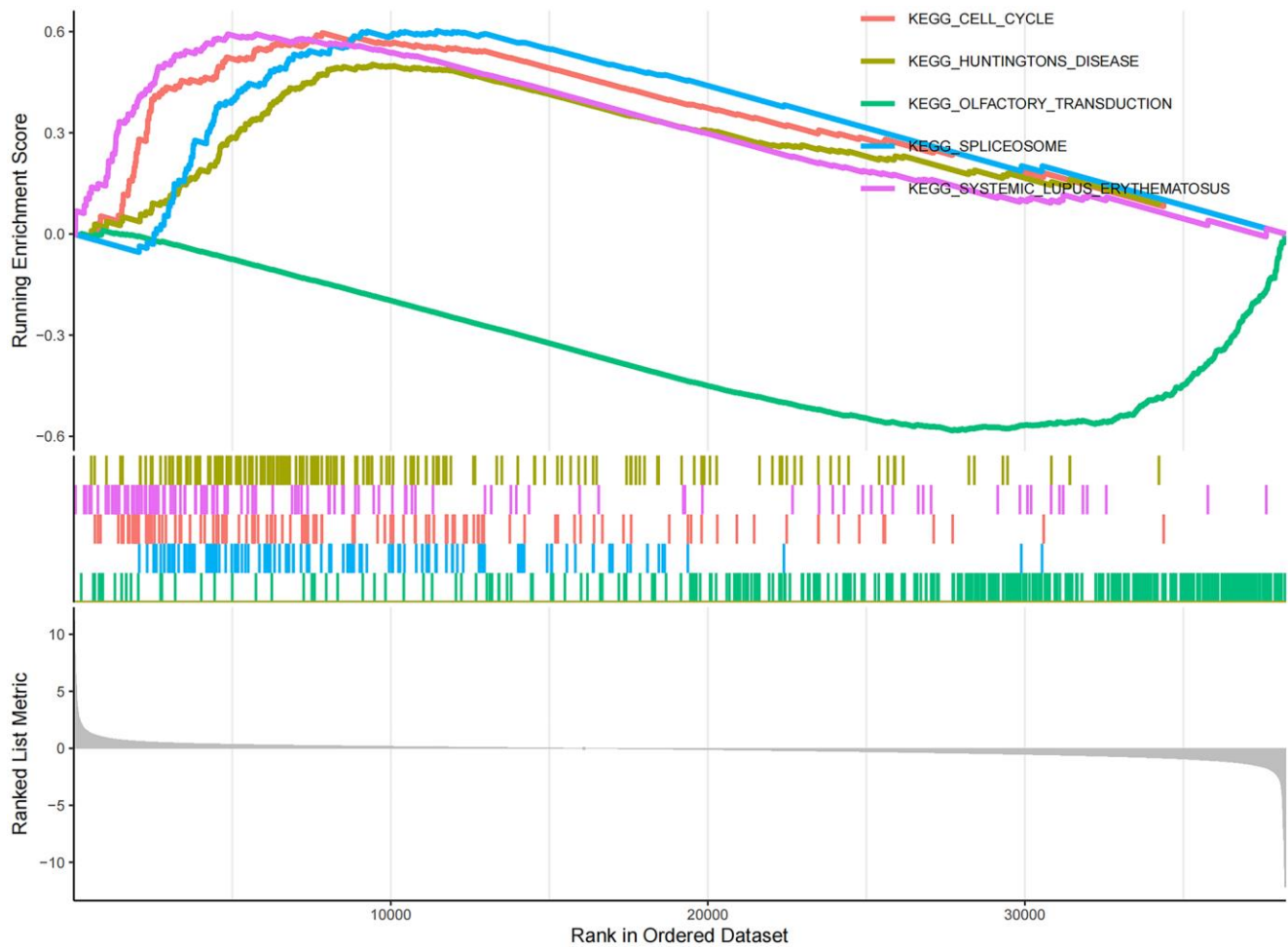
35. Li K, Zhang A, Li X, Zhang H, Zhao L. Advances in clinical immunotherapy for gastric cancer. *Biochim Biophys Acta Rev Cancer*. 2021; 1876:188615. <https://doi.org/10.1016/j.bbcan.2021.188615> PMID:34403771
36. Janjigian YY, Shitara K, Moehler M, Garrido M, Salman P, Shen L, Wyrwicz L, Yamaguchi K, Skocyzlas T, Campos Bragagnoli A, Liu T, Schenker M, Yanez P, et al. First-line nivolumab plus chemotherapy versus chemotherapy alone for advanced gastric, gastro-oesophageal junction, and oesophageal adenocarcinoma (CheckMate 649): a randomised, open-label, phase 3 trial. *Lancet*. 2021; 398:27–40. [https://doi.org/10.1016/S0140-6736\(21\)00797-2](https://doi.org/10.1016/S0140-6736(21)00797-2) PMID:34102137
37. Janjigian YY, Kawazoe A, Yañez P, Li N, Lonardi S, Kolesnik O, Barajas O, Bai Y, Shen L, Tang Y, Wyrwicz LS, Xu J, Shitara K, et al. The KEYNOTE-811 trial of dual PD-1 and HER2 blockade in HER2-positive gastric cancer. *Nature*. 2021; 600:727–30. <https://doi.org/10.1038/s41586-021-04161-3> PMID:34912120
38. Feng S, Xu Y, Dai Z, Yin H, Zhang K, Shen Y. Integrative Analysis From Multicenter Studies Identifies a WGCNA-Derived Cancer-Associated Fibroblast Signature for Ovarian Cancer. *Front Immunol*. 2022; 13:951582. <https://doi.org/10.3389/fimmu.2022.951582> PMID:35874760
39. Li W, Zhang S, Zhao Y, Wang D, Shi Q, Ding Z, Wang Y, Gao B, Yan M. Revealing the Key MSCs Niches and Pathogenic Genes in Influencing CEP Homeostasis: A Conjoint Analysis of Single-Cell and WGCNA. *Front Immunol*. 2022; 13:933721. <https://doi.org/10.3389/fimmu.2022.933721> PMID:35833124
40. Wang K, Zheng Q, Liu X, Geng B, Dong N, Shi J. Identifying hub genes of calcific aortic valve disease and revealing the immune infiltration landscape based on multiple WGCNA and single-cell sequence analysis. *Front Immunol*. 2022; 13:1035285. <https://doi.org/10.3389/fimmu.2022.1035285> PMID:36405745
41. Liu C, Zhou Y, Zhou Y, Tang X, Tang L, Wang J. Identification of crucial genes for predicting the risk of atherosclerosis with system lupus erythematosus based on comprehensive bioinformatics analysis and machine learning. *Comput Biol Med*. 2023; 152:106388. <https://doi.org/10.1016/j.combiomed.2022.106388> PMID:36470144
42. Wei P, Dong M, Bi Y, Chen S, Huang W, Li T, Liu B, Fu X, Yang Y. Identification and validation of a signature based on macrophage cell marker genes to predict recurrent miscarriage by integrated analysis of single-cell and bulk RNA-sequencing. *Front Immunol*. 2022; 13:1053819. <https://doi.org/10.3389/fimmu.2022.1053819> PMID:36439123
43. Geng R, Huang X, Li L, Guo X, Wang Q, Zheng Y, Guo X. Gene expression analysis in endometriosis: Immunopathology insights, transcription factors and therapeutic targets. *Front Immunol*. 2022; 13:1037504. <https://doi.org/10.3389/fimmu.2022.1037504> PMID:36532015

SUPPLEMENTARY MATERIALS

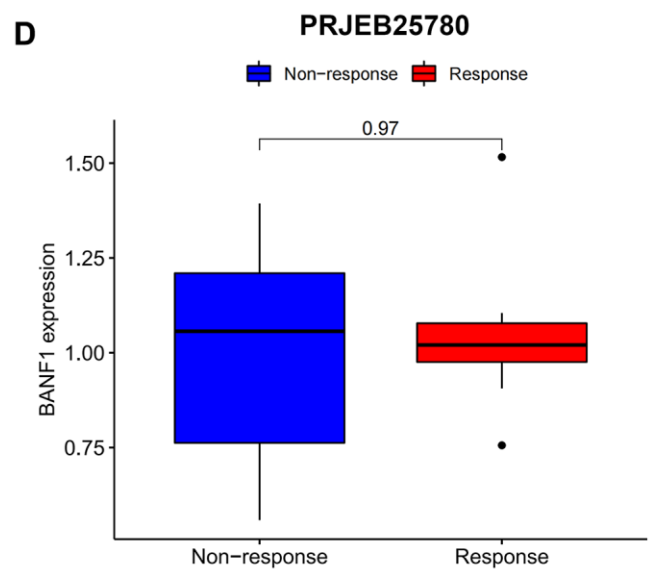
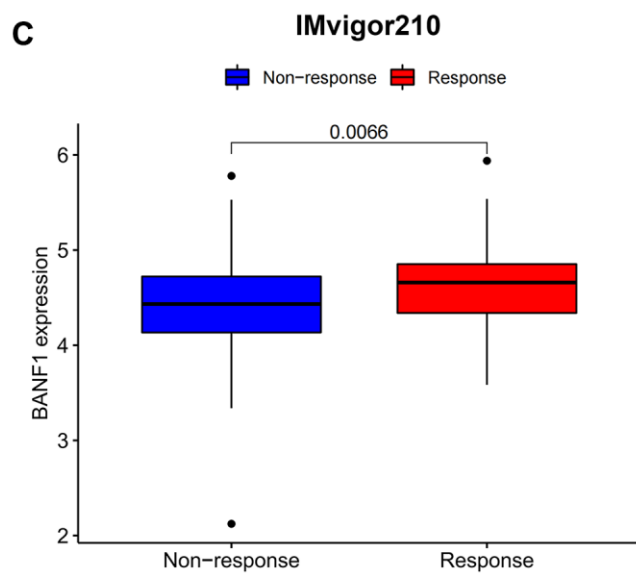
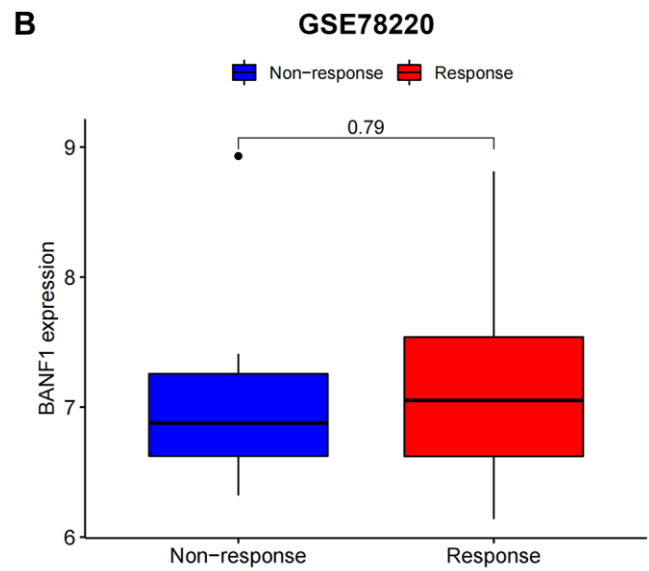
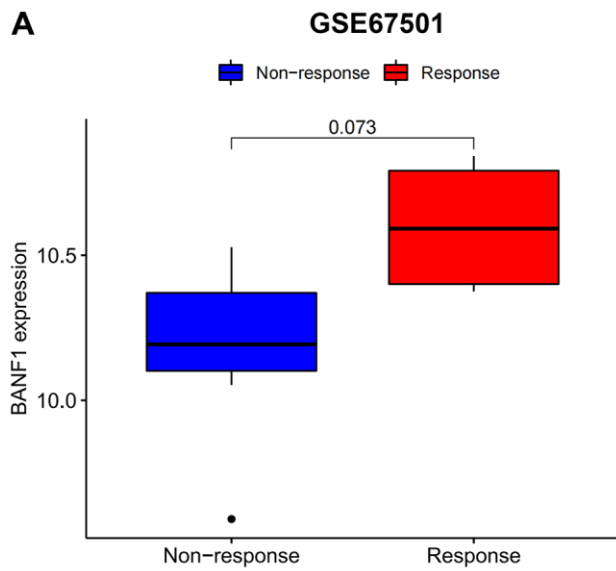
Supplementary Figures



Supplementary Figure 1. Principal component analysis to view the differences between samples. (A) Difference between raw data of GSE118916 and GSE54129. (B) Expression of data of GSE118916 and GSE54129 after removing batch effects.



Supplementary Figure 2. Gene set enrichment analysis. The samples were divided into two groups according to the high and low expression of BANF1, the differential genes between the two groups were calculated, and the signaling pathways in which BANF1 might be involved were speculated based on the analysis of the up- and down-regulation of the differential genes.



Supplementary Figure 3. Effect of BANF1 expression level on immunotherapy. Differences in BANF1 expression in renal cell carcinoma (A), melanoma (B), urothelial carcinoma (C), and gastric cancer (D) immunotherapy cohorts between response and non-response groups.