The Diagnostic and Prognostic Value of BEN-Domain (BEND) Family Genes in Gastric Cancer

Jiaxin Fan
Shanxi Medical University
Min Yang
Shanxi Medical University
Chaojie Liang
First Hospital of Shanxi Medical University
Chaowei Liang
Shanxi Medical University
Jiansheng Guo (Sdyygjs2017@126.com)
First Hospital of Shanxi Medical University

Research Article

Keywords: gastric cancer, BEND family genes, diagnostic value, prognostic value, nomogram, The Cancer Genome Atlas

Posted Date: January 7th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1212308/v1

Abstract

BEND(BEN domain-containing protein)is a domain protein-coding gene, whose abnormal expression is related to the occurrence of malignant tumors. But studies on gastric cancer are rare. We attempted to investigate the role of BEND family genes in evaluating the prognosis of gastric cancer and guiding clinical treatment. We analyzed the BEND family genes expression, prognostic value, and drug sensitivity in pan-cancer, and the correlation between their expression and tumor microenvironment of gastric cancer, stemness index, immune subtypes, and clinicopathological characteristics were analyzed. We constructed a model using BEND3P1 and BEND6 to evaluate the prognosis of gastric cancer patients. Multivariate Cox proportional risk model analysis showed that risk score is an independent risk factor for gastric cancer patients. To assess the value of risk score for prognosis, patients were divided into high-risk and low-risk groups based on median risk scores, and survival analyses were performed. The results showed that the OS of patients with high-risk scores is significantly lower. We also constructed a nomogram to predict individual survival probability using the BEND risk score and clinical case characteristics. In conclusion, the BEND family genes can predict the prognosis and guide the treatment of gastric cancer patients.

1. Introduction

In 2020, it is estimated that there will be 19.3 million new cancer cases worldwide, with gastric cancer ranking fifth (5.6%) in incidence and fourth (7.7%) in mortality[1]. According to an article published in The Lancet in 2020, the incidence of gastric cancer in China in 2015 was 9.4%, ranking third[2]. Although the incidence of gastric cancer has decreased due to the improvement of surgical operations, the development of radiotherapy, chemotherapy, immunotherapy, and other therapies, the mortality rate is still high[3]. At present, the prognosis of gastric cancer is mainly based on the TNM stage, Borrmann classification, histological type, invasion depth, lymph node metastasis, tumor markers, etc. Although there have been a large number of studies on the pathogenesis of gastric cancer, it is still not completely clear. Due to the development of genomic technology, the role of genes in the signal transduction pathway has been explored, and their mutations may be related to the occurrence of tumors and may lead to changes in drug sensitivity. Therefore, further study of individual predictive markers at the molecular level is of great significance for clinical diagnosis and treatment, and prognostic judgment.

BEND (BEN domain-containing protein) is a protein-encoding gene. Its encoding product is a domain protein that plays an important role in the maintenance of DNA chromatin structure.BEND family has 9 members:BEND2,BEND3,BEND4,BEND5,BEND6,BEND7,and BEND3P1,BEND3P2,BEND3P3.Previous studies have shown that BEND3 is located in human 6q21, which is often deleted in leukemia and lymphoma[4] and is also associated with several other malignant tumors, including breast cancer, melanoma, ovarian cancer, and prostate cancer[5–7]. Sathyan KM et al.found that overexpression of BEND3 can inhibit cell transcription through its interaction with histone deacetylase and Sall4, and the overexpression of BEND3 can prevent the cell cycle process and make the cell stagnate in the S phase[8].

The protein synthesized by BEND5 can induce embryonic stem cells to differentiate into primordial germ progenitor-like cells[9]. According to 2 case reports, fusion mutations of BEND2 and some genes may induce the occurrence of spinal astroblastoma and high-grade neuroepithelial tumor[10,11].BEND6 is an independent factor of neural BEN that binds to mammalian CSL protein CBF1 and antagonizes Notch-dependent target activation, and its reduction increases the self-renewal of NSC in the wild-type neocortex[12]. At present, there are few studies on BEND family genes in gastric cancer. We first explored the important role of BEND family genes in gastric cancer.

Our research scheme is to explore the role of BEND family genes in gastric cancer by using bioinformatics methods. Using the data from the TCGA database, we analyzed the BEND gene expression in pan-cancer and the correlation between BEND gene expression and clinicopathological features and prognosis.GO analysis and KEGG analysis revealed the regulatory mechanism of BEND family genes in gastric cancer. A prognostic model was established based on BEND-related genes to calculate the risk score of each patient and to explore the relationship between risk score and overall survival rate. The potential association of BEND family genes with the tumor microenvironment(TME), Stemness index, drug sensitivity, and tumor-infiltrating immune cells was also investigated. In addition, we also studied the relationship between BEND family genes expression and immune subtypes and the clinical relevance of STAD patients. Finally, we constructed the nomograph to predict the one-year and three-year survival rates of gastric cancer patients, providing a basis for clinical research. This study revealed the potential role of BEND family genes in gastric cancer, which helps us to further understand the specific pathogenesis of gastric cancer.

2. Materials And Methods

2.1. Source and processing of data

Gene expression RNAseq(HTSEQ-FPKM),clinicopathological data,immune subtypes, survival data,and Stemness index(based on RNA and DNA methylation)of 33 tumors were obtained from UCSCXena online database(https://xena.ucsc.edu/,from TCGA database)[13,14]."*","**","***",indicate P value<0.05,<0.01,<0.001 ,respectively.The box diagram and heatmap were further designed using R-package "ggpubr" and "pheatmap" respectively. R-package "corrplot" was used to analyze the correlation between BEND family genes. Perl programming language was used to integrate data, match gene expression information with patient clinical information, and delete unknown or incomplete clinical information. Since the data is obtained through a public database, there are no ethical issues.

2.2. Analyze the expressions of BEND family genes

The expression levels of BEND family genes in various cancers including gastric cancer were analyzed using the data of The Cancer Genome Atlas (TCGA).In TCGA data, R-package "survival", "Beeswarm" and"Limma" were used to analyze the differential expression of BEND family genes between gastric cancer and normal tissues, and the P-value was determined by the Wilcox test. The analysis results are

visualized using scatter plots and paired plots. In addition, by GEPIA database mapping box a scatter diagram and the Human Protein Atlas(HPA) (https://www.proteinatlas.org/)in the immunohistochemical results, we further verify the BEND family genes differentially expressed in gastric cancer. R-package "Survival", "Survminer" and "timeROC" were used to analyze TCGA data, draw ROC curves, and evaluate the accuracy of BEND family genes in predicting the survival of gastric cancer patients. R software (V.4.0.4) was used to analyze the above contents.

2.3. Clinical correlation analysis

Gastric cancer tissue samples with clinicopathological characteristics such as age, stage, depth of invasion, lymph node metastasis, and distant metastasis were selected for correlation analysis with BEND family genes expression. Age and distant metastasis were analyzed using the Wilcox test. The remaining clinical features were analyzed using the Kruskal test.

2.4. Survival analysis

Survival data of each sample were obtained from the TCGA database, and the relationship between BEND family genes expression and clinical outcome was further analyzed, and the overall survival (OS) was fully evaluated. Kaplan-Meier method and log-rank test were used for survival analysis($P<0.05$). The median expression level of BEND family genes was selected as the cutoff value of the human cancer dichotomy, and each patient was divided into a high-risk group and a low-risk group. According to the high and low-risk values, R-package "SurvMiner" and "survival" were used to describe the survival curve.COX analysis was also performed to determine the relationship between BEND family genes expression and the prognosis of pan-cancers. Finally, use R-package"survival" and "forestplot" to draw the forest map.

In addition, we verified the association between the survival rate of pan-cancer patients and clinical outcomes through the online database Kaplan-Meier Plotter(https://kmplot.com/analysis/), including OS,disease-free survival (DFS).In COX analysis, $P <0.05$ is set as the threshold value. The Kaplan-Meier Plotter database was also used to assess the impact of BEND family genes on the survival of patients affected by 21 different types of cancer.

We used the Cox proportional hazards regression model to analyze the TCGA data to evaluate whether BEND family genes can be used as a prognostic factor for cancer. We use R-package "Survival" and "Survminer" to analyze the data and draw forest diagrams.

2.5. Tumor immune microenvironment and stem cell analysis

The stromal and immune cell scores were used to predict tumor purity and the presence of infiltrating stromal and immune cells in pan-cancers, using the ESTIMATE computations in R-Package "ESTIMATE" and "LIMMA" (using BEND family genes expression data).ESTIMATE score=stromal score+immune score[15].A higher ESTIMATE score corresponds to a lower tumor purity. The Stemness index describes the similarity between tumor cells and tumor stem cells. With the "cor. Test" command and R-package"limma",

Spearman's method was used to analyze the correlation between the expression of BEND family genes and RNA Stemness index(RNAss) and DNA Stemness index(DNAss). Both indicators are visualized through R-package"Corrplot".The correlation of the expression BEND family genes with tumor immune microenvironment(TME) and Stemness index was analyzed by R-package "reshape2", "GGpubr", "GGplot2" and "Limma".

2.6. Gene Enrichment Analysis (GSEA)

GSEA(version 4.1.0) was used to analyze the signal pathways related to BEND family genes in gastric cancer. BEND family genes expression level was used to determine gene expression enrichment among different phenotypes.Select the annotation of gene sets(c2.cp.kegg.v7.2.symbols.gmt)as a reference gene set.Gene sequence analysis 1000 times. Normalized enrichment score(NES), P-value, and FDR value were used to rank the enrichment paths of each group. Eventually, multiple cancer-related pathways were identified.

2.7. Construct a prognostic model of BEND family genes

Firstly, we used LASSO regression analysis to screen genes of BEND family genes. By running 1000 times of cross-validation likelihood method, the optimal value for penalization coefficient lambda(λ) is screened out. This method can avoid overfitting the signature. In the case where λ is the smallest, select the most suitable gene to construct the signature. We then performed multivariate Cox analysis on these genes to establish a prognostic model. We used the formula of a risk score to calculate the risk score of each patient(risk score=coef gene1×gene1 expression+coef gene2×gene2 expression+.+coef gene Ñ × gene Ñ expression). The risk score was obtained by weighting gene expression levels and regression coefficients(coef). The hazard ratio (HR) of the multivariate Cox regression analysis was log-transformed to calculate the coef value, and the expression level of each gene related to the prognostic characteristics was defined as the expression level of the Ñ gene. Gastric cancer patients were divided into high-risk groups and low-risk groups according to the median risk score.

We used the K-M method to analyze the difference in overall survival rate between high-risk and low-risk groups and draw a survival curve. We then plotted ROC curves using the "survivalROC" package to assess the ability of risk score and other clinical features to predict gastric cancer and to assess sensitivity and specificity(AUC values) by the area below the curve. To determine whether risk score can be used as an independent predictor of prognosis of gastric cancer, we included age, gender, grade, stage, TNM stage, and risk score into univariate and multivariate Cox regression analysis.

2.8. Build nomogram

Nomograms can intuitively calculate the survival rate of gastric cancer patients[16], which is of great value in clinical application. We screened the prognostic factors of gastric cancer patients, constructed a nomogram using R-package "surviva" and"RMS" to predict 1-year and 3-year survival rates of gastric cancer patients, and drew calibration curves to evaluate the accuracy of the nomogram. Finally, the ROC curve of the nomogram was drawn and the area under the curve(AUC) was calculated.

2.9. Correlation analysis between BEND family genes expression level and immune infiltration

To further explore the mechanism of BEND family genes involved in the pathological progression of gastric cancer, we investigated the relationship between BEND family genes expression and immune infiltration. "E1071", "preprocessCore", and "Limma"packages were used to determine the content of immune cells in each sample, and then the "Corrplot" was used to plot the histogram showing the results of immune infiltration. Patients are divided into high-expression groups and low-expression groups according to the expression level of BEND family genes, and then used "Limma" package and "Vioplot" package to detect differences in immune cells (pFilter=0.05). "Limma", "ggplot2", "ggpubr", and "ggExtra" packages were used to detect the correlation between BEND family genes expression and the content of the immune cell (pFilter=0.05). Finally, we intersected the results of the two tests to obtain immune cells associated with BEND family genes expression.

2.10. Drug sensitivity analysis and immune subtype analysis

Drug sensitivity processing data is downloaded from CellMiner™ database[17,18] (version:2020.3,number according to the library:2.4.2, HTTPS:// discover.nci.nih.gov/cellminer/home.do).Data processing and results were visualized using R-packages"impute", "Limma", "GGplot2", and "GGpubr".The correlation analysis of immune subtypes of BEND family genes mainly uses R packages "limma", "ggplot2", and "reshape2".

2.11. Statistical analysis

R software(version 4.0.4) was used for statistical analysis. Wilcox test and Kruskal test were used to determine the relationship between clinicopathological parameters and BEND family genes expression level. Kaplan-Meier analysis was used to study the relationship between BEND family genes expression level and survival rate.Univariate and multivariate survival analyses were performed by Cox proportional risk regression model,and P<0.05 indicated statistical significance."*","**"and"***" respectively represent P value<0.05,<0.01 and <0.001.

3. Result

3.1. BEND family genes expression level in pan-cancers

We first analyzed RNA sequencing data present in TCGA using R software to evaluate the transcriptional levels of BEND in different human tumors(Fig.1A). According to the results, BEND3, BEND5 and, BEND7 were highly expressed in pan-cancers, BEND3P1, BEND3P3, and BEND6 were moderately expressed, while BEND2, BEND3P2, and BEND4 were low expressed in tumor tissues. Further analysis showed that BEND5 expression is the highest in CHOL. The expression level of BEND7 was highest in CHOL. The expression of BEND3 was highest in CHOL. The expression of BEND4 was highest in prostatic PRAD. The expression of

BEND6 was highest in LUSC. The expression of BEND3P3 was highest in CHOL. The expression of BEND3P1 was highest in CHOL(Fig.1B).BEND5 and BEND6 were the two genes with the most significant positive correlation(Correlation coefficient=0.32, Fig.1C), while BEND3 and BEND5 were the two genes with the most significant negative correlation(Correlation coefficient=-0.16, Fig.1C).

We also used R software to analyze the RNA sequencing data in the TCGA database to evaluate the differential expression of BEND family genes in pan-cancers. The 11,075 mRNA expression profiles of 33 cancers from TCGA showed that BEND2 (Fig.2A) is highly expressed in many types of tumors, including breast invasive carcinoma(BRCA), Kidney Chromophobe(KICH), kidney renal clear cell carcinoma(KIRC), kidney renal papillary cell carcinoma(KIRP), Liver hepatocellular carcinoma(LIHC), lung adenocarcinoma(LUAD), lung adenocarcinoma (LUSC), Prostate adenocarcinoma(PRAD) and Thyroid carcinoma(THCA).BEND3(Fig.2B) in bladder urothelial carcinoma(BLCA),BRCA,Cholangiocarcinoma(CHOL), colon adenocarcinoma(COAD),glioblastoma multiforme(GBM),head and neck squamous cell carcinoma(HNSC),LIHC,LUAD,LUSC,PRAD,Rectum adenocarcinoma(READ),Stomach adenocarcinoma(STAD),THCA and uterine corpus endometrial carcinoma (UCEC) had higher expressions;meanwhile,it had lower expression in KICH, KIRC, and KIRP. BEND3P1(Fig.2C) was significantly expressed in BLCA, CHOL, COAD, HNSC, LIHC, READ, STAD and was low in KICH, KIRC, KIRP, THCA, and UCEC. BEND3P2(Fig.2D) was highly expressed in HNSC, STAD, and THCA; while it was low in CHOL, KICH, KIRP, LIHC, LUAD, and LUSC. BEND3P3(Fig.2E) was significantly expressed in CHOL, but in BLCA, BRCA, COAD, GBM, KICH, KIRC, KIRP, LUAD, LUSC, PRAD, READ, THCA, and UCEC had lower expression levels. BEND4(Fig.2F) was significantly expressed in PRAD, but in BLCA, BRCA, COAD, ESCA, GBM, KIRP, LIHC, LUAD, LUSC, READ, THCA, and UCEC was low in expression. BEND5 (Fig.2G) was significantly expressed in CHOL and in BRCA, COAD, ESCA, HNSC, KICH, KIRP, LUAD, LUSC, PRAD, READ, STAD, THCA, and UCEC was low in expression. BEDN6(Fig.2H) was significantly expressed in BRCA, CHOL, ESCA, HNSC, KIRP, LIHC, LUAD, and LUSC, and was low in COAD, GBM, PRAD, READ, UCEC. BEDN7(Fig.2I) was significantly expressed in CHOL and LIHC, which was low in BRCA, HNSC, KICH, KIRC, KIRP, LUAD, LUSC, PRAD, THCA, and UCEC.

Then we used the HPA database to verify the differential expression of the BEND family genes in gastric cancer and normal tissues (Fig.3). According to the results, BEND3 was highly expressed in gastric cancer tissues, while BEND5 was down-regulated in gastric cancer tissues. Due to the lack of BEND3P1, BEND3P2, and BEND3P3 data in the HPA database, further immunohistochemical verification of their expression is needed.

3.2. Prognostic value of BEND family genes in pan-cancers

Next, we analyzed the prognostic value of BEND family genes for pan-cancers through different databases. The Kaplan-Meier survival curve suggested(Fig.4) that the expression of BEND family genes is related to the prognosis of several TCGA cancers.Compared with the low-gene expression group,the survival rate of patients in the high expression group was higher.The expression of BEND2 was protective

in PRAD (OS:N=499,P=0.042).The expression of BEND3 in ACC(OS:N=79, P=0.007),KIRC(OS:N=531,P=0.033),LIHC(OS:N=368,P=0.021),MESO(OS:N=84,P=0.015),SARC(OS:N=260,P□0.001),SKCM(OS:N=457,P=0.040) was an unfavorable effect,however,which in LAML(OS:N=132,P=0.001),READ(OS:N=158,P=0.010) and UVM(OS:N=80,P=0.027) was protective.The expression of BEND3P1 played an unfavorable role in COAD(OS:N=448,P=0.047),LAML(OS:N=132,P=0.046) and UCEC (OS:N=544,P=0.035),while played a protective role in BLCA(OS:N=406,P=0.047).The expression of BEND3P2 played an protective effect in DLBC(OS:N=47□P=0.012)□HNSC(OS:N=501P=0.019)□LAML(OS:N=132,P=0.017)□ PAAD(OS:N=177,P=0.045).BEND3P3 played an unfavorable role in MESO(OS:N=84,P=0.048),and which in KIRC(OS:N=531□P=0.003)□LGG(OS:N=524,P□0.001) and LUAD(OS:N=513□P=0.048) was protective.The expression of BEND4 had an unfavorable effect in DLBC(OS:N=47,P=0.008) and UVM (OS:N=80,P=0.037),and it played a protective role in SKCM(OS:N=457, P<0.001). The expression of BEND5 played an unfavorable role in ACC(OS:N=79,P=0.008)□READ(OS:N=158,P=0.012),and in BLCA(OS:N=406,P□0.001)□BRCA(OS:N=1082,P=0.005)□CESC(OS:N=293,P=0.048)□ KIRP(OS:N=286,P=0.006)□LUAD(OS:N=513,P=0.002)□MESO(OS:N=84,P=0.015)□ PAAD(OS:N=177,P=0.014)□PRAD(OS:N=499,P=0.007)□TGCT(OS:N=139,P=0.048)□UVM(OS:N=80,P=0.034) it played an protective role.The expression of BEND6 in BLCA(OS:N=406,P=0.013)□KIRC(OS:N=531,P□ 0.001)□MESO(OS:N=84,P□0.001)□UCEC(OS:N=544,P=0.006)□UVM(OS:N=80,P=0.009) was unfavorable.The expression of BEND7 in BLCA(OS:N=406,P=0.038)□KIRC(OS:N=531,P=0.036)□ LGG(OS:N=524,P=0.002) waws protective,however which was adverse in HNSC(OS:N=501,P=0.016)□ UCEC(OS:N=544,P=0.018).

Next, we further conducted a Cox analysis on the prognostic value of BEND family genes expression for pan-cancers. According to the results□Fig.5□, the expression of BEND2 was an unfavorable factor for the prognosis of ACC, LICH, OV(HR>1, P<0.05). The expression of BNED3 was a protective factor for the prognosis of LAML, READ, UVM(HR<1, P<0.05), and its expression was a disadvantage factor of ACC, KIRP, LIHC, MESO, SARC, SKCM(HR>1, P<0.05). The expression of BEND3P1 was a protective factor for the prognosis of GBM(HR<1, P<0.05) and adversely affected the prognosis of COAD, ESCA, KIRP, UCEC(HR>1, P<0.05). The expression of BEND3P2 had no obvious correlation with the prognosis of pan-cancers. The expression of BEND3P3 was a protective factor for the prognosis of KIRC and LGG(HR<1, P<0.05), but it was an unfavorable factor for the prognosis of MESO and PCPG(HR>1, P<0.05). The expression of BEND4 was a protective factor for the prognosis of LGG(HR<1, P<0.05), and its expression was an unfavorable factor for the prognosis of DLBC, KIRP, PCPG, SARC, THCA(HR>1, P<0.05). The expression of BEND5 was a protective factor for the prognosis of BLCA, HNSC, KIRC, KIRP, LUAD, MESO, PAAD, PRAD, SKCM, UVM(HR<1, P<0.05), and its expression was an unfavorable factor for the prognosis of STAD(HR>1, P<0.05). The expression of BEND6 was an unfavorable factor for the prognosis of BLCA, KIRC, KIRP, LUAD, MESO, READ, SARC, STAD, UCEC(HR>1, P<0.05). The expression of BEND7 was a protective factor for the prognosis of GBM, KIRC, LGG, READ(HR<1, P<0.05), and its expression was an unfavorable factor for the prognosis of UCEC.

3.3. Relationship between the BEND family genes and tumor microenvironment correlation□tumor stem cells

The BEND family genes expression was positively or negatively correlated with Stromal Score(Fig.6A), Immune Score(Fig.6B), and Estimate Score(Fig.6C)in pan-cancers patients. Similarly, BEND family genes expression was positively or negatively correlated with Tumor Purity, RNAss, and DNAss in pan-cancers.BNED3 showed a significantly negative correlation with Stromal Score, Immune Score, and Estimate Score in various cancers, and showed a significantly positive correlation with Tumor Purity, RNAss, and DNAss, suggesting that BEND3 is involved in the occurrence of various tumors.BEND4, BEND5, and BEND6 were positively correlated with Stromal Score, Immune Score, and Estimate Score in various cancers, and significantly negatively correlated with Tumor Purity, RNAss, and DNAss.

3.4. Drug sensitivity analysis of BEND family genes

To analyze the potential correlation between BEND family genes expression and drug sensitivity in different cancer cell lines from the CellMiner™ database, we performed a correlation analysis. The results showed (Fig.7) that the expression of BEND4 is positively correlated with the sensitivity of Fludarabine, XK-469, Asparaginase, Benddamustine, Nelarabine, Chlorambucil, Chelerythrine, and Cytarabine. The expression of BEND5 was positively correlated with the sensitivity of Nelarabine, XK469, Chelerythrine, and PX-316. The expression of BEND7 was negatively correlated with the sensitivity of Tamoxifen, Pipamperone, Ixabepilone, VinorelbinePipamperone, Ixabepilone, and Vinorelbine.

3.5. Association between BEND family genes expression and Immune subtype in Stomach adenocarcinoma□STAD□

To analyze the potential correlation between BEND family genes expression and immune subtypes in STAD, we performed correlation analysis.Immune subtypes include C1(Wound Healing),C2(IFN-G dominant),C3(inflammatory),C4(lymphocyte depleted),C5(immunologically quiet) and C6(TGF-b dominant).According to the results, the expressions of BEND3, BEND3P3, BEND4, BEND5, and BEND6 were all related to the immune subtypes of STAD(Fig.8).BEND3 was highly expressed in C1, C2, and C4, and moderately expressed in C3 and C6.BEND4 was moderately expressed in C1, C2, and C3, and was lowly expressed in C4 and C6.BEND3P3 was highly expressed in C3 and moderately expressed in other subtypes. The expression of BEND5 was highest in C3, moderate in C1, C2, and C6, and low in C4. The expression of BEND6 was highest in C3 and C6, and moderate in other subtypes.

3.6. Association between BEND family genes expression and Clinical relevance□Age□Gender□Grade□Stage and TNM Stage□in Stomach adenocarcinoma□STAD□

We analyzed the correlation between BEND family genes expression in STAD and clinical relevance, including Age, Gender, Grade, Stage, and TNM stage. According to the results(Fig.9), the expressions of BEND3, BEND5, and BEND6 were correlated with Age. The expression of BEND3 was higher in gastric cancer patients that are > 65 years old, while BEND5 and BEND6 expression were higher in ≤65-years olds. The expressions of BEND4, BEND5, and BEND6 were significantly correlated with Grade. With the gradual increase of Grade, the expression of BEND4 and BEND5 gradually increased, and the expression was the highest in G3. The expression of BEND6 was high in G3 and low in G1 and G2. The expression of BEND4, BEND5, and BEND6 was significantly correlated with Stage.BEND4 was highly expressed in Stage Ⅲ and Stage Ⅳ, BEND5 was highly expressed in Stage Ⅱ, Stage Ⅲ, and Stage Ⅳ, and BEND6 was highly expressed in all stages.BEND3P2, BEND4, BEND5, and BEND6 were correlated with the T stage.BEND3P2 and BEND4 showed low expression in all T stages, while BEND5 and BEND6 showed high expression in T2, T3, and T4 stages. There was no correlation between the BEND family genes expression and gender, N and M stage.

3.7. Correlation between BEND family genes expression and TME, Stemness index in STAD

We further explored the relationship between BEND family genes expression and TME and Stemness scores of gastric cancer and conducted correlation analysis. The results showed that the expression of BEND3 is positively correlated with RNAss(Fig.10), while the expressions of BEND3P3, BEND4, BEND5, and BEND6 are negatively correlated with RNAss. The expression of BEND3 was positively correlated with DNAss, while the expressions of BEND3P3, BEND4, BEND5, and BEND6 were negatively correlated with DNAss. The expressions of BEND3 and BEND7 were negatively correlated with Stromal Score, while the expressions of BEND3P3, BEND4, BEND5, and BEND6 were positively correlated with Stromal Score. The expressions of BEND3 and BEND7 were negatively correlated with Immune Score, while the expressions of BEND4, BEND5, and BEND6 were positively correlated with Immune Score. The expressions of BEND3 and BEND7 were negatively correlated with Estimate Score, while BEND3P3, BEND4, BEND5, and BEND6 were positively correlated with Estimate Score.

3.8. An BEND risk model predicted OS and DFS in patients with STAD

We compared the relationship between the BEND family genes expression level and the risk of tumor occurrence in STAD patients, and the expressions of BEND3P1 and BEND6 were significantly correlated with the risk of tumor occurrence (Fig.11A), so a prognostic model composed of BEND3P1 and BEND6 was constructed. The risk score of each patient was calculated according to the risk score calculation formula: risk score =(-0.272087×BEND3P1 expression)+ (0.6450065×BEND6 expression). Patients were assigned to the high-risk groups (n=175) and the low-risk group(n=175) based on the median risk score. We constructed a heat map to show the expression of two genes in the high-risk and low-risk groups

(Fig.10A), where the expression of BEND6 in the high-risk groups was higher than that in the low-risk groups, while the expression of BEND3P1 was lower than that in the low-risk groups.Fig.11B showed the distribution of risk scores of STAD patients. After patients were divided into two groups, the risk scores increased from left to right.Fig.11C showed the distribution of survival status and survival time of patients with different risk scores. K-M curve was used to compare the difference in overall survival rate between high-risk groups and low-risk groups(Fig.11D, P=1.249E-02). The results showed that the OS of gastric cancer patients with high-risk scores is significantly lower than that of patients with low-risk scores. According to the univariate analysis between clinical characteristics, risk score, and OS of gastric cancer patients(Fig.11E, F), we drew a time-dependent ROC curve(Fig.11G) to evaluate the predictive value of various pathological features on the prognosis of gastric cancer.The risk score(AUC=0.619),age(AUC=0.578),gender(AUC=0.540),Grade(AUC=0.562),Stage (AUC=0.602),T stage(AUC=0.563),M stage (AUC=0.528) and N stage (AUC=0.577) both had high sensitivity and specificity. At the same time, we constructed a nomogram to predict OS and DFS using the BEND risk score and clinical case characteristics(Fig.11H). We calculated the score of gastric cancer patients according to the nomogram, then added them to obtain the total score to predict 1-year and 3-year survival rates of gastric cancer patients, which was beneficial to guide clinical treatment. The closer the calibration curve is to the diagonal, the more accurate the prediction result is. The calibration curve of the nomogram showed good accuracy in predicting 1-year and 3-year survival rates(Fig.11I, J). The ROC curves of 1 year(AUC=0.619) and 3 years(AUC=0.555) also showed that the predictive ability of the nomogram is very accurate(Fig.11K).

3.9.The relationship between STAD overall survival rate and clinicopathological characteristics by Univariate and multivariate COX analyses.

Univariate and multivariate Cox proportional hazards regression analysis were used to evaluate whether risk model could be used as an independent predictor of adverse survival in STAD patients.According to the results□Table.1□,Age(HR=1.024;95%CI,1.005-1.043;p=0.013),Gender(HR=1.562;95%CI,1.020-2392;p=0.040),Stage(HR=1.532;95%CI,1.213-1.935;p□0.001),Tumor size(HR=1.303;95%CI,1.025-1.656;p=0.030),Lymph node(HR=1.254;95%CI,1.056-1.490□p=0.010),Distant metastasis(HR=2.133;95%CI,1.109-4.103;p=0.023) and BEND risk score(HR=2.342;95%CI,1.195-4.589;p=0.013) were risk factors for prognosis. Multivariate analysis showed Age(HR=1.041;95%CI,1.019-1.062;p□0.001),Gender(HR=1.647;95%CI,1.066-2.544;p=0.025) and BEND risk score(HR=3.276;95%CI,1.607-6.680;p=0.001) are independent risk factors for prognosis.

3.10. Relationship between BEND scores and gene expression profiles

We used cBioportal to analyze the mutation and copy number variation of the BEND family genes.The mutation frequencies of BEND2,BEND3,BEND4,BEND5,BEND6 and BEND7 were 3%,7%,2.9%,1.6%,2.9%,1.8%,respectively(Fig. 12A).There was a missense mutation during BEND2, BEND3, BEND4, BEND5, BEND6, and BEND7. Then we performed survival analysis on the two genes that constructed the risk model and found that the high expression of BEND6 suggests a poor prognosis(Fig.12B, P=0.027), and the AUC curve showed that it has a high diagnostic value(Fig.12C, AUC at 1 years=0.579; AUC at 2 years=0.607; AUC at 1 years=0.653), while the predictive effect of BEND3P1 was poor(Fig.12E, F).In addition, we found multiple cancer-related KEGG pathways through Gene Set Enrichment Analysis(GSEA). The BEND6-related signaling pathways were as follows(Fig.12D): BASAL-CELL-CARCINOMA, DNA-REPLICATION, ECM-RECEPTOR-INTERACTION, MELANOMA, and TGF-BETA-SIGNALING-PATHWAY. The BEND3P1-related signaling pathways were as follows（Fig.12G）: BASAL-CELL-CARCINOMA, COLORECTAL-CANCER, DNA-REPLICATION, MELANOMA, NON-SMALL-CELL-LUNG-CANCER, NOTCH-SIGNALING-PATHWAY, and T-CELL-RECEPTOR-SIGNALING-PATHWAY.

3.11. Relationship between BEND family gene expression and Immune infiltration

Tumor-infiltrating lymphocytes have been proven to be independent prognostic factors in cancer patients, so we investigated the relationship between BEND family genes expression and immune infiltration. The bar chart showed the number of immune cells in each sample(Fig.13A). For STAD patients, the expression levels of BEND6 and BEND3P1 were mostly positively correlated with the infiltration of different immune cells. For tissues with high expression of BEND6(Fig.13B), the infiltration level of B cell naïve(P=0.026), T cells follicular helper(P=0.021), Monocytes(P<0.001), Macrophages M2(P=0.029), Mast cells resting(P<0.001) was higher than that of tissues with low expression of BEND6.Plasma cells(P=0.034), T cells CD4 memory activated(P=0.008), and Macrophages MO(P=0.042) showed higher infiltration in tissues with low BEND expression. Correlation test analysis(Fig.13C) showed that the BEND6 expression is significantly positively correlated with the infiltration levels of B cell(R=0.19,P=0.00028),Dendritic cells resting(R=0.12,P=0.017), Macrophages M2(R=0.11,P=0.037),Mast cells resting(R=0.33,P=8.9e-11) and Monocytes (R=0.21,P=4.1e-05),which was negatively correlated with the infiltration of Macrophages MO(R=-0.19,P=0.00027),T cells follicular helper(R=-0.11,P=0.037), Neutrophils(R=-0.13,P=0.012),Plasma cells(R=-0.11,P=0.032) and T cells CD4 memory activated(R=-0.18,P=7e-04).After taking the intersection results of the two tests, it could be concluded that the expression level of BEND6 is significantly related to B cells naïve, Plasma cells, T cells CD4 memory activated, T cells follicular helper, Macrophages M0, Macrophages M2, Dendritic cells resting and Neutrophils(Fig.13D).

For tissues with high expression of BEND3P1(Fig.14A), the infiltration level of T cells follicular helper(P 0.001) and Macrophages M1(P=0.035) was higher.However,the infiltration level of NK cells resting(P=0.019),Dendritic cells activated(P=0.030),Neutrophils(P=0.003) was higher in tissues with low BEND3P1 expression.Correlation test analysis(Fig.14C) showed BEND6 expression level is significantly positively correlated with infiltration level of Macrophages M1 (R=0.1,P=0.05),NK cells

activated(R=0.11,P=0.033) and T cells follicular helper (R=0.19,P= 0.00028).However the BEND6 expression level was negatively correlated with infiltration level of Mast cells resting(R=-0.11, P=0.035),Neutrophils(R=-0.17,P=0.00088) and NK cells resting(R=-0.18,P=0.00075). After taking the intersection of the two test results, it could be concluded that the expression level of BEND3P1 is significantly correlated with T cells follicular helper, NK cells resting, NK cells activated, Macrophages M1, Mast cells activated, and Neutrophils(Fig.14B). All in all, BEND family genes play an important role in the immune infiltration of STAD.

Discussion

Gastric cancer is one of the most common malignant tumors in the world. Although the rate of early diagnosis of gastric cancer has increased with the improvement of diagnostic methods, the prognosis of gastric cancer is still not optimistic. Surgery, radiotherapy, and chemotherapy are currently more effective methods for the treatment of gastric cancer, but for advanced patients, these methods have limited effects. Therefore, it is imperative to find the pathogenesis of gastric cancer and explore the early diagnosis methods of gastric cancer patients from the genetic level to evaluate the prognosis and guide clinical treatment.

In this study, we used an independent database from TCGA to study the BNED family genes (BEND2, BEND3, BEND3P1, BEND3P2, BEND3P3, BEND4, BEND5, BEND6, BEND7) expression in 10327 primary cases with 33 different cancer types including tumors, adjacent tissues, and normal tissues. Previous studies have shown that BEND family genes are related to the occurrence of some human tumors and may be fused with other genes to induce different cancers. This study found that BEND family genes are up-regulated in a variety of cancers, among which BEND3, BEND3P1, and BEND3P2 were significantly expressed in STAD(Fig.2). Therefore, we explored the value of BEND family genes as tumor markers in the diagnosis and prognosis of STAD and constructed models to help evaluate the research and development of BEND family genes targeted therapy.

According to our research, the expression of BEND family genes was related to the prognosis of gastric cancer, among which BEND5 and BEND6 were unfavorable factors for the prognosis of gastric cancer patients(Fig.5). According to the results of Fig.4, the expression of BEND family genes had nothing to do with the prognosis of gastric cancer. This contradictory result may be related to the different data collection methods and potential mechanisms related to biological characteristics.

Another important finding was that the expression of BEND family genes was correlated with TME and Stemness index in pan-cancers.TME is a component of tumors, which promotes tumor development, helps tumor growth, and suppresses immunity through complex intercellular signal transduction[19,20]. The Estimate algorithm is based on a single-sample gene set enrichment analysis to generate three scores: Stromal Score(to capture the presence of matrix in tumor tissue), Immune Score(to represent the infiltration of immune cells in tumor tissue), and Estimated Score(to infer tumor purity)[15]. Exploring potential therapeutic targets can help reshape TME, and then inhibit the growth of tumor cells or enhance

the efficacy of other drugs[21]. After analyzing the pan-cancers data from the TCGA database, we knew that the immune components in TME had an impact on the prognosis of cancer patients. In particular, the ratio of matrix and immune components in TME was significantly related to STAD(Fig.10). These results emphasize the importance of exploring the interaction between tumor cells and immune cells. Reshaping the tumor microenvironment and regulating the abnormal components in TME can provide a new perspective for future cancer treatment.

The stemness index can reflect the significant difference between cancer and normal tissues and reveals the degree of intratumoral heterogeneity at the level of stemness. In addition, a higher stemness index is associated with a worse prognosis and greater oncogenic dedifferentiation[22]. RNAsi/RNAss reflects gene expression, DNAsi/DNAss reflects epigenetic characteristics. The stemness index is related to the prognosis of cancer patients and can predict clinical outcomes[23], it is positively related to the level of activated immune cells, such as cancer-associated fibroblasts, M2 macrophages, etc[24]. The application of the stemness index shows the sensitivity of tumors to immunotherapy[25] and can predict the prognosis of specific immunotherapy[26]. Previous studies have shown that mRNAsi can predict the overall survival(OS), clinical characteristics, gene mutation status, immune cell infiltration, and tumor microenvironment composition of gastric cancer patients[27]. Our research found that in pan-cancers, BEND family gene expression was also significantly positively or negatively correlated with RNAss and DNAss, especially in STAD.

We also analyzed the potential correlation between BEND family genes expression in STAD and immune subtypes, and the results showed(Fig.8) that BEND3 is highly expressed in C1, C2, and C4, and was moderately expressed in C3 and C6. BEND4 was moderately expressed in C1, C2, C3 and it was low in C4, C6.BEND3P3 was highly expressed in C3 and was moderately expressed in the other subtypes.BEND5 had the highest expression in C3 and it showed moderate expression in C1, C2, C6, while showing low expression in C4.BEND6 had the highest expression in C3, C6, and was moderately expressed in other subtypes. Previous studies have shown that in patients with kidney renal clear cell carcinoma, subtype C3 is associated with the best 10-year OS rate. The decrease in the proportion of subtype C3 is related to the increase in pathological grades, furthermore, the increase in monocytes, CD8 T cells, macrophages, and activated NK cells of such patients is closely related to killing tumors[28]. The research of Soldevilla B et al. found that subtype C2 has higher activation of pathways related to the immune system, apoptosis, DNA repair, mTOR signaling, and oxidative phosphorylation[29].C1 and C2 subtypes of gastric cancer had a good prognosis, which was related to high tumor lymphocyte infiltration(TIL), on the contrary, the survival rate of patients with C3 subtypes was poor[30]. These results indicated that the expression of BEND family genes is closely related to the subtypes of gastric cancer patients, and the prognosis of patients can be predicted by their differences in expression.

Through research, we found that BEND family genes are highly expressed in gastric cancer. Age, Gender, Stage, Tumor size, Lymph node, Distant metastasis, and risk score were risk factors for prognosis. Multivariate analysis showed that Age, Gender, and risk score are independent risk factors for prognosis.

Gene enrichment analysis found multiple enrichment pathways related to cancer, such as BASAL-CELL-CARCINOMA, DNA-REPLICATION, MELANOMA, etc. At the same time, we constructed a risk model through BEND6 and BEND3P1. The study found the risk score is an independent risk factor that affects the prognosis and can be used to predict the OS and DFS of STAD. Therefore, we speculate that the two genes that constitute the prognostic model are involved in the progression of STAD. In addition, to facilitate the better evaluation of the prognosis of patients with gastric cancer in the clinic, We have constructed a nomogram with a good predictive effect.

Previous literature has shown that Tumor Infiltrating Lymphocytes(TILs) is considered to be closely related to the proliferation and elimination of cancer cells[31], and it may participate in specific immune responses. At the same time, it could be released by PD-1 blockers to maintain tumor specificity T cell response[32]. Our research found that the expression of BEND6 is significantly related to the infiltration of B cells naïve, Plasma cells, T cells CD4 memory activated, T cells follicular helper, Macrophages M0, Macrophages M2, Dendritic cells resting, and Neutrophils. The expression of BEND3P1 is significantly related to the infiltration of T cells follicular helper, NK Cell resting, NK cells activated, Macrophages M1, Mast cells activated, and Neutrophils, suggesting that the role of BEND family genes in gastric cancer may be related to immune infiltration. T cells follicular helper is a new member of the Th subgroup family and a key participant in the production of T cell-dependent antibody response and B cell memory[33], but the specific immune mechanism is not particularly clear. Some studies have shown that its high expression is associated with more extensive breast cancer infiltration. The presence of Tfh cells that produce CXCL13 is related to the immune structure of the tissue near the tumor bed. These structures may be involved in sustainable and effective long-term anti-tumor immunity[34]. The current research results of Neutrophils have been relatively rich, which can reflect the individual's inflammatory state. In recent years, it has also been considered to play an important role in the occurrence of tumors, including growth, proliferation, or metastasis[35-37].PGE2 or protease released by Neutrophils can promote tumorigenesis and metastasis[38,39]. However, the HGF/MET-dependent nitric oxide released by it promotes the killing of cancer cells, thereby inhibiting tumor growth and metastasis[40]. The contradictory effect may be related to the microenvironment in which Neutrophils are located. In addition, the combination of Neutrophils and other inflammatory markers can be used to judge clinical results. For example, the Neutrophils to Lymphocytes ratio (NLR) can be used to predict the prognosis of gastric cancer patients and is an independent prognostic factor[41,42]. However, the specific relationship between BEND family genes and TILs is still unclear, and further research is needed to prove it.

At present, many studies have shown that abnormal gene expression is related to chemotherapy resistance[43,44]. Our research has found that the high expression of BEND family genes may be related to the sensitivity of multiple drugs. For example, the expression of BEND4 is significantly positively correlated with the sensitivity of Fludarabine, XK-469, Asparaginase, Benddamustine, Nelarabine, Chlorambucil, Chelerythrine, and Cytarabine. The expression of BEND5 is significantly positively correlated with the sensitivity of Nelarabine, XK-469, Chelerythrine, and PX-316. On the contrary, the expression of BEND7 is negatively correlated with the sensitivity of Tamoxifen, Pipamperone, Ixabepilone,

and Vinorelbine. The study of Beelen K et al. found that patients whose tumors express p-p70S6K cannot benefit from the adjuvant treatment of Tamoxifen[45]. The study of Chong K et al. showed that the expression of IGF-1 and ERα in breast cancers that may be resistant to Tamoxifen both increased[46]. Vinorelbine requires BRCA1 to initiate apoptosis in preclinical models and the deletion of BRCA1/MAD2L1 is a putative predictive marker of Vinorelbine resistance in mesothelioma[47]. It can be seen that the up-regulation or down-regulation of gene expression has a guiding role in the treatment of cancer patients. BEND family genes can provide new ideas for the research process of STAD, and detecting the expression of BEND4, BEND5, and BEND7 has guiding significance for the selection of clinical drugs.

Although we analyzed the correlation between BEND family genes and survival prognosis, tumor microenvironment, stemness index, immune infiltration, and drug sensitivity, there are still some shortcomings. First of all, this research is to conduct bioinformatics analysis on the expression of BEND family genes but does not conduct human, zoology, or cytology research which may lack certain clinical evidence. Secondly, although BNED family genes are related to immune cell infiltration and drug sensitivity, the specific mechanism is still unclear. Thirdly, whether the prognosis of the patient can be judged by the TME situation is also unknown. In general, BEND family genes can be used as reliable biomarkers for the diagnosis of gastric cancer. They play an important role in immune cell infiltration, which can guide the use of clinical drugs and evaluate the prognosis of gastric cancer patients. But the mechanism in gastric cancer is still needed to discuss.

Conclusion

Our research confirmed the expression profile of BEND family genes in pan-cancers and found that BEND family genes are related to the prognosis of gastric cancer. At the same time, it may be related to the tumor microenvironment, stemness index, immune subtype, and clinical correlation. In addition, the expression of BEND4, BEND5, and BEND7 in gastric cancer is related to drug sensitivity. We constructed a risk assessment model based on the two genes BEND3P1 and BEND6 to assess the prognosis of patients with gastric cancer. These results can provide new directions for potential therapeutic targets of gastric cancer through BEND family genes.

Declarations

Acknowledgments

We thank the databases of TCGA, Kaplan-Meier Plotter, PrognoScan, and CellMiner™ for the availability of the data.

Author contributions

Jiaxin Fan and Chaojie Liang: conception and design. Jiaxin Fan and Min Yang: acquisition, analysis, and interpretation of data. Jiaxin Fan and Chaowei Liang: figures drawing. Jiaxin Fan and Jiansheng Guo:

writing and revision of manuscript. Jiansheng Guo: study supervision. All authors read and approved the final manuscript.

Competing interests

The authors declare that there are no conflicts of interest.

Data availability

The authors certify that all the original data in this research could be obtained from the public database. The Gene expression RNAseq(HTSEQ-FPKM),clinicopathological data,immune subtypes, survival data,and Stemness index(based on RNA and DNA methylation)of 33 tumors were obtained from UCSCXena online database(https://xena.ucsc.edu/,from TCGA database). The immunohistochemical results were obtained from the Human Protein Atlas(HPA) (https://www.proteinatlas.org/).BEND family gene expression and survival in pan-cancer were verified in the online data-base: Kaplan-Meier Plotter (https://kmplot.com/analysis/). Drug sensitivity processed data was downloaded from the CellMiner™ database (Version: 2020.3, database: 2.4.2,https://discover.nci.nih.gov/cellminer/home.do). Other data used to support the findings of this study are included within the supplementary information files. All the data are available from the first author or corresponding author upon request.

References

1. Sung, H. et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 71, 209-249, doi:10.3322/caac.21660 (2021).
2. Wei, W. et al. Cancer registration in China and its role in cancer prevention and control. Lancet Oncol 21, e342-e349, doi:10.1016/s1470-2045(20)30073-5 (2020).
3. Hallowell, B. D. et al. Gastric cancer mortality rates among US and foreign-born persons: United States 2005-2014. Gastric cancer : official journal of the International Gastric Cancer Association and the Japanese Gastric Cancer Association 22, 1081-1085, doi:10.1007/s10120-019-00944-w (2019).
4. Zhang, Y. et al. A 3-cM commonly deleted region in 6q21 in leukemias and lymphomas delineated by fluorescence in situ hybridization. Genes, chromosomes & cancer 27, 52-58, doi:10.1002/(sici)1098-2264(200001)27:1<52::aid-gcc7>3.0.co;2-x (2000).
5. Orphanos, V. et al. Allelic imbalance of chromosome 6q in ovarian tumours. British journal of cancer 71, 666-669, doi:10.1038/bjc.1995.132 (1995).
6. Orphanos, V., McGown, G., Hey, Y., Boyle, J. M. & Santibanez-Koref, M. Proximal 6q, a region showing allele loss in primary breast cancer. British journal of cancer 71, 290-293, doi:10.1038/bjc.1995.58 (1995).
7. Hyytinen, E. R. et al. Defining the region(s) of deletion at 6q16-q22 in human prostate cancer. Genes, chromosomes & cancer 34, 306-312, doi:10.1002/gcc.10065 (2002).

8. Sathyan, K. M., Shen, Z., Tripathi, V., Prasanth, K. V. & Prasanth, S. G. A BEN-domain-containing protein associates with heterochromatin and represses transcription. Journal of cell science 124, 3149-3163, doi:10.1242/jcs.086603 (2011).
9. University, S. Y.-s. he use of Bend5 protein in enhancing the differentiation of embryonic stem cells into primordial germ progenitor-like cells. CN106011051A (2016-10-12).
10. Tsutsui, T. et al. Spinal cord astroblastoma with EWSR1-BEND2 fusion classified as HGNET-MN1 by methylation classification: a case report. Brain tumor pathology 38, 283-289, doi:10.1007/s10014-021-00412-3 (2021).
11. Burford, A. et al. The ten-year evolutionary trajectory of a highly recurrent paediatric high grade neuroepithelial tumour with MN1:BEND2 fusion. Scientific reports 8, 1032, doi:10.1038/s41598-018-19389-9 (2018).
12. Dai, Q. et al. BEND6 is a nuclear antagonist of Notch signaling during self-renewal of neural stem cells. Development (Cambridge, England) 140, 1892-1902, doi:10.1242/dev.087502 (2013).
13. Goldman, M. J. et al. Visualizing and interpreting cancer genomics data via the Xena platform. Nature biotechnology 38, 675-678, doi:10.1038/s41587-020-0546-8 (2020).
14. Casper, J. et al. The UCSC Genome Browser database: 2018 update. Nucleic acids research 46, D762-d769, doi:10.1093/nar/gkx1020 (2018).
15. Yoshihara, K. et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nature communications 4, 2612, doi:10.1038/ncomms3612 (2013).
16. Iasonos, A., Schrag, D., Raj, G. V. & Panageas, K. S. How to build and interpret a nomogram for cancer prognosis. J Clin Oncol 26, 1364-1370, doi:10.1200/jco.2007.12.9791 (2008).
17. Reinhold, W. C. et al. The NCI-60 Methylome and Its Integration into CellMiner. Cancer research 77, 601-612, doi:10.1158/0008-5472.Can-16-0655 (2017).
18. Luna, A. et al. CellMiner Cross-Database (CellMinerCDB) version 1.2: Exploration of patient-derived cancer cell line pharmacogenomics. Nucleic acids research 49, D1083-d1093, doi:10.1093/nar/gkaa968 (2021).
19. Zemek, R. M. et al. Sensitization to immune checkpoint blockade through activation of a STAT1/NK axis in the tumor microenvironment. Science translational medicine 11, doi:10.1126/scitranslmed.aav7816 (2019).
20. Wang, Y. et al. Cancer/testis Antigen MAGEA3 Interacts with STAT1 and Remodels the Tumor Microenvironment. International journal of medical sciences 15, 1702-1712, doi:10.7150/ijms.27643 (2018).
21. Mullins, S. R. et al. Intratumoral immunotherapy with TLR7/8 agonist MEDI9197 modulates the tumor microenvironment leading to enhanced activity when combined with other immunotherapies. J Immunother Cancer 7, 244, doi:10.1186/s40425-019-0724-8 (2019).
22. Zheng, H. et al. An absolute human stemness index associated with oncogenic dedifferentiation. Briefings in bioinformatics 22, 2151-2160, doi:10.1093/bib/bbz174 (2021).

23. Wei, R. et al. Integrative Analysis of Biomarkers Through Machine Learning Identifies Stemness Features in Colorectal Cancer. Frontiers in cell and developmental biology 9, 724860, doi:10.3389/fcell.2021.724860 (2021).
24. Xu, Q., Xu, H., Chen, S. & Huang, W. Immunological Value of Prognostic Signature Based on Cancer Stem Cell Characteristics in Hepatocellular Carcinoma. Frontiers in cell and developmental biology 9, 710207, doi:10.3389/fcell.2021.710207 (2021).
25. Xu, F. et al. A signature of immune-related gene pairs predicts oncologic outcomes and response to immunotherapy in lung adenocarcinoma. Genomics 112, 4675-4683, doi:10.1016/j.ygeno.2020.08.014 (2020).
26. Yi, L. et al. Integrative stemness characteristics associated with prognosis and the immune microenvironment in esophageal cancer. Pharmacological research 161, 105144, doi:10.1016/j.phrs.2020.105144 (2020).
27. Mao, D. et al. Identification of Stemness Characteristics Associated With the Immune Microenvironment and Prognosis in Gastric Cancer. Frontiers in oncology 11, 626961, doi:10.3389/fonc.2021.626961 (2021).
28. Wang, Q. et al. Tumor microenvironment immune subtypes for classification of novel clear cell renal cell carcinoma profiles with prognostic and therapeutic implications. Medicine 100, e24949, doi:10.1097/md.0000000000024949 (2021).
29. Soldevilla, B. et al. The correlation between immune subtypes and consensus molecular subtypes in colorectal cancer identifies novel tumour microenvironment profiles, with prognostic and therapeutic implications. Eur J Cancer 123, 118-129, doi:10.1016/j.ejca.2019.09.008 (2019).
30. Kim, J. Y., Kim, W. G., Kwon, C. H. & Park, D. Y. Differences in immune contextures among different molecular subtypes of gastric cancer and their prognostic impact. Gastric cancer : official journal of the International Gastric Cancer Association and the Japanese Gastric Cancer Association 22, 1164-1175, doi:10.1007/s10120-019-00974-4 (2019).
31. Galon, J., Angell, H. K., Bedognetti, D. & Marincola, F. M. The continuum of cancer immunosurveillance: prognostic, predictive, and mechanistic signatures. Immunity 39, 11-26, doi:10.1016/j.immuni.2013.07.008 (2013).
32. Demaria, O. & Vivier, E. Immuno-Oncology beyond TILs: Unleashing TILCs. Cancer cell 37, 428-430, doi:10.1016/j.ccell.2020.03.021 (2020).
33. Crotty, S. Follicular helper CD4 T cells (TFH). Annual review of immunology 29, 621-663, doi:10.1146/annurev-immunol-031210-101400 (2011).
34. Gu-Trantien, C. et al. CD4 follicular helper T cell infiltration predicts breast cancer survival. The Journal of clinical investigation 123, 2873-2892, doi:10.1172/jci67428 (2013).
35. Powell, D., Lou, M., Barros Becker, F. & Huttenlocher, A. Cxcr1 mediates recruitment of neutrophils and supports proliferation of tumor-initiating astrocytes in vivo. Scientific reports 8, 13285, doi:10.1038/s41598-018-31675-0 (2018).

36. Deryugina, E. I. et al. Tissue-infiltrating neutrophils constitute the major in vivo source of angiogenesis-inducing MMP-9 in the tumor microenvironment. Neoplasia (New York, N.Y.) 16, 771-788, doi:10.1016/j.neo.2014.08.013 (2014).
37. Coffelt, S. B., Wellenstein, M. D. & de Visser, K. E. Neutrophils in cancer: neutral no more. Nature reviews. Cancer 16, 431-446, doi:10.1038/nrc.2016.52 (2016).
38. Xiao, Y. et al. Cathepsin C promotes breast cancer lung metastasis by modulating neutrophil infiltration and neutrophil extracellular trap formation. Cancer cell 39, 423-437.e427, doi:10.1016/j.ccell.2020.12.012 (2021).
39. Antonio, N. et al. The wound inflammatory response exacerbates growth of pre-neoplastic cells and progression to cancer. The EMBO journal 34, 2219-2236, doi:10.15252/embj.201490147 (2015).
40. Finisguerra, V. et al. MET is required for the recruitment of anti-tumoural neutrophils. Nature 522, 349-353, doi:10.1038/nature14407 (2015).
41. Miyamoto, R. et al. The neutrophil-to-lymphocyte ratio (NLR) predicts short-term and long-term outcomes in gastric cancer patients. European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology 44, 607-612, doi:10.1016/j.ejso.2018.02.003 (2018).
42. Cupp, M. A. et al. Neutrophil to lymphocyte ratio and cancer prognosis: an umbrella review of systematic reviews and meta-analyses of observational studies. BMC Med 18, 360, doi:10.1186/s12916-020-01817-1 (2020).
43. Yano, S. et al. Hepatocyte growth factor induces gefitinib resistance of lung adenocarcinoma with epidermal growth factor receptor-activating mutations. Cancer research 68, 9479-9487, doi:10.1158/0008-5472.Can-08-1643 (2008).
44. Mullighan, C. G. et al. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. The New England journal of medicine 360, 470-480, doi:10.1056/NEJMoa0808253 (2009).
45. Beelen, K. et al. Phosphorylated p-70S6K predicts tamoxifen resistance in postmenopausal breast cancer patients randomized between adjuvant tamoxifen versus no systemic treatment. Breast cancer research : BCR 16, R6, doi:10.1186/bcr3598 (2014).
46. Chong, K., Subramanian, A., Sharma, A. & Mokbel, K. Measuring IGF-1, ER-α and EGFR expression can predict tamoxifen-resistance in ER-positive breast cancer. Anticancer research 31, 23-32 (2011).
47. Busacca, S. et al. BRCA1/MAD2L1 Deficiency Disrupts the Spindle Assembly Checkpoint to Confer Vinorelbine Resistance in Mesothelioma. Molecular cancer therapeutics 20, 379-388, doi:10.1158/1535-7163.Mct-20-0363 (2021).

Tables

Table.1. Univariate and multivariate Cox analyses of overall survival prediction, based on BEND scores calculated by cases from TCGA

Parameters	univariate analysis				multivariate analysis			
	HR	95%CI		P-value	HR	95%CI		P-value
		Lower	Upper			Lower	Upper	
Age (continuous)	1.024	1.005	1.043	0.013	1.041	1.019	1.062	0.000
Gender	1.562	1.020	2392	0.040	1.647	1.066	2.544	0.025
Grade	1.339	0.923	1.943	0.124	1.333	0.907	1.960	0.143
stage	1.532	1.213	1.935	0.000	1.343	0.856	2.106	0.200
Tumor size	1.303	1.025	1.656	0.030	1.109	0.793	1.550	0.547
Lymph node	1.254	1.056	1.490	0.010	1.087	0.845	1.399	0.515
Distant metastasis	2.133	1.109	4.103	0.023	2.217	0.958	5.131	0.063
BEND riskScore	2.342	1.195	4.589	0.013	3.276	1.607	6.680	0.001

Figures

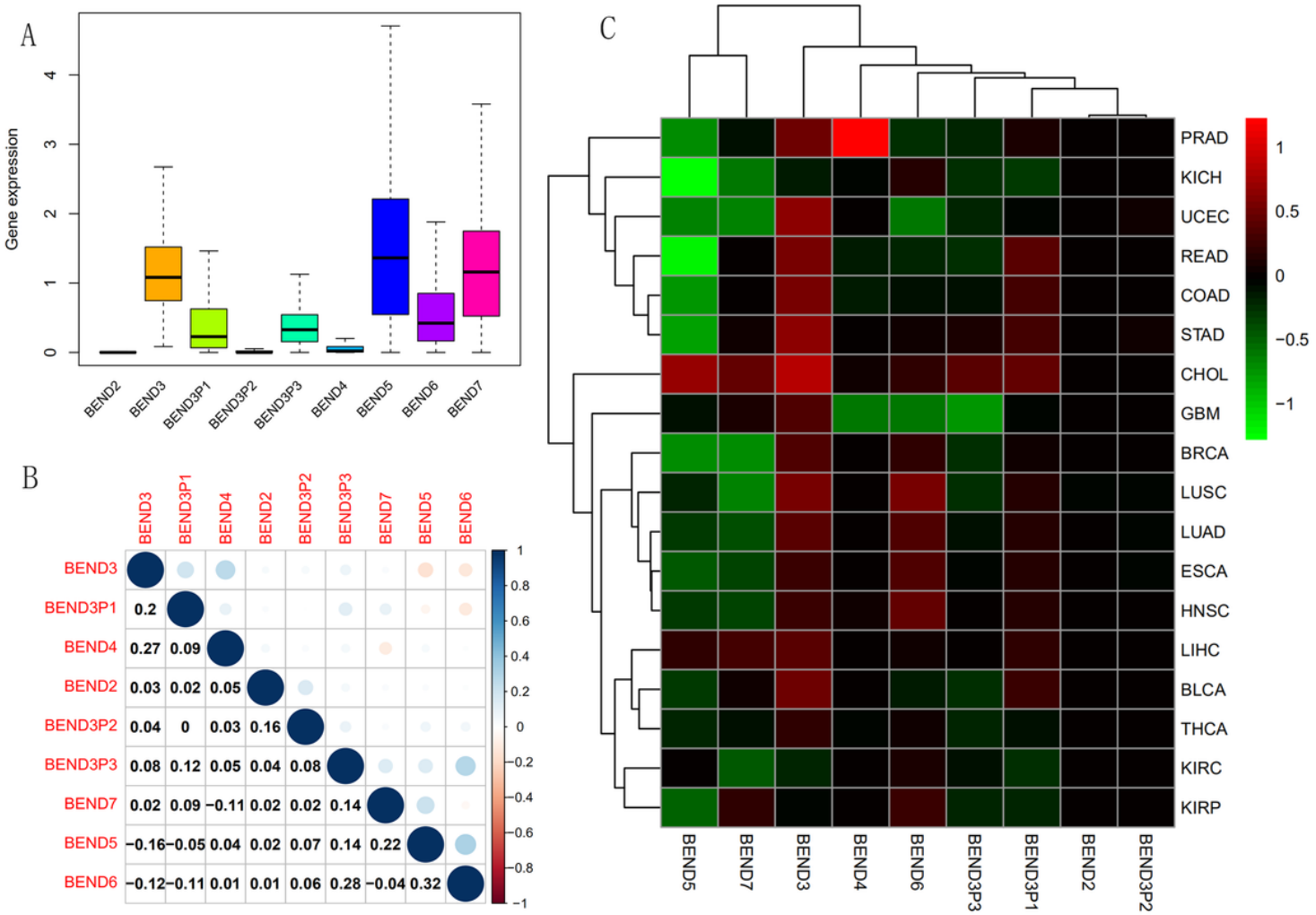

Figure 1

BEND family gene expression levels and correlation in different cancer types from TCGA. (A) Increased or decreased expression of BEND family gene in various cancer. (B) BEND family gene expression levels in different cancer types from TCGA data. The color in each small rectangle represents a high or low expression of the BEND family gene in each cancer. The red and green indicate the high or low expression, respectively. (C) The correlation between the BEND family gene. Blue dots indicate a positive correlation and red dots indicate a negative correlation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

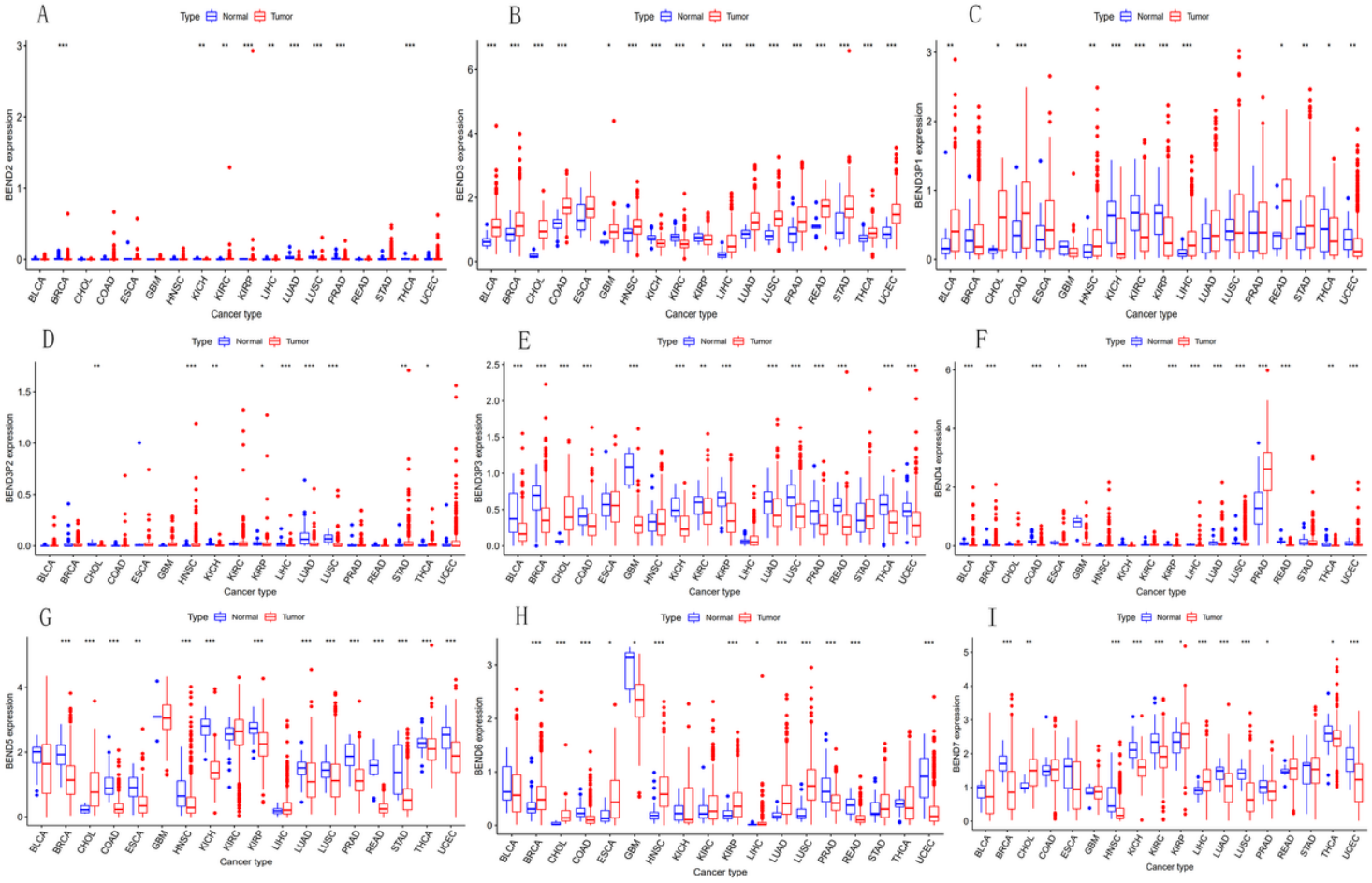

Figure 2

BEND family gene expression levels in different cancer types and normal tissue. (A) BEND2, (B) BEND3, (C)BEND3P1 ,(D) BEND3P2,(E)BEND3P3,(F)BEND4,(G)BEND5,(H)BEND6,(I)BEND7. The red rectangle box represents gene expression levels in tumor tissue and the blue rectangle box represents normal tissue. * P <0.05, ** P <0.01, and *** P <0.001. Red cancer names indicate high expression and blue cancer names indicate low expression of the corresponding BEND family gene. BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CHOL, cholangiocarcinoma; COAD, colon adenocarcinoma; ESCA, esophageal carcinoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; KICH, Kidney Chromophobe; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; PRAD, prostate adenocarcinoma; READ, rectum adenocarcinoma; STAD, stomach adenocarcinoma; THCA, thyroid carcinoma; UCEC, uterine corpus endometrial carcinoma. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

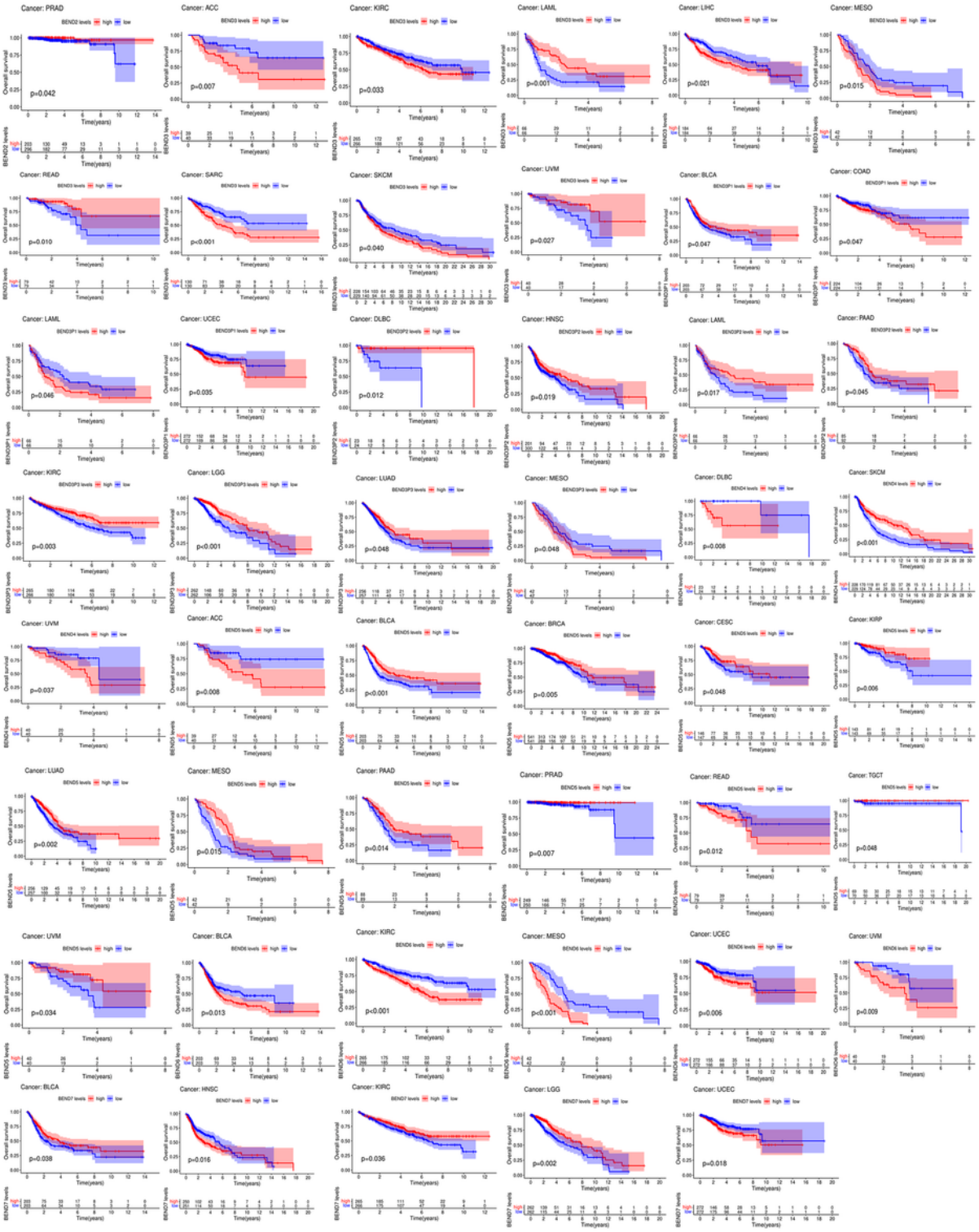

Figure 3

the differential expression of the BEND family genes in gastric cancer and normal tissues from the HPA database.

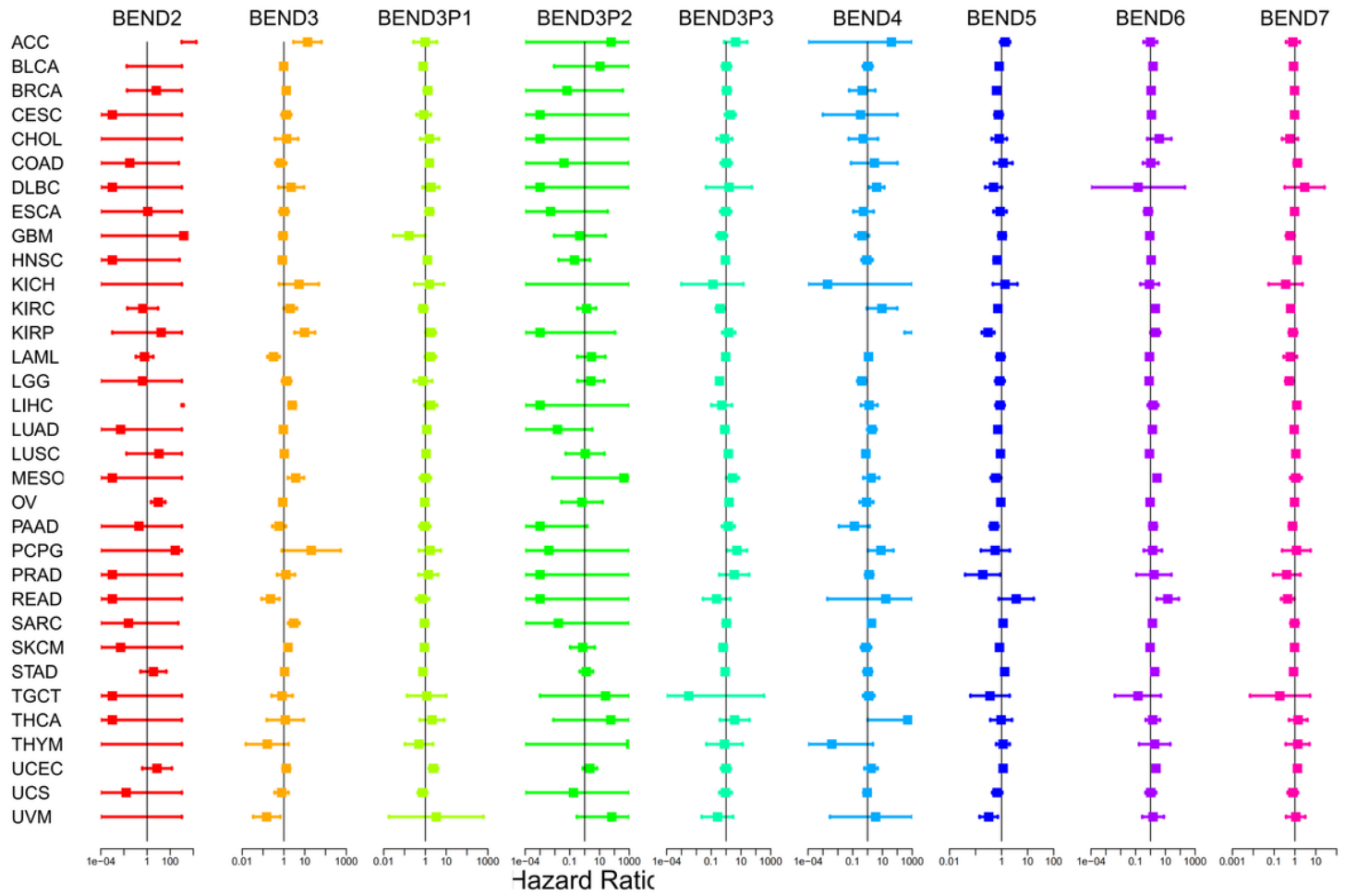

Figure 4

Kaplan-Meier survival curves comparison of high and low expression of BEND family gene in pan-cancers.OS survival curves of BEND2 in different cancers: PRAD (n=499);OS survival curves of BEND3 in different cancers: ACC(n=79); KIRC(n=531); LAML(n=132); LIHC(n=368); MESO(n=84); READ(n=158); SARC(n=260); SKCM(n=457); UVM(n=80);OS survival curves of BEND3P1 in different cancers: BLCA(n=406); COAD(n=448); LAML(n=132); UCEC(n=544); OS survival curves of BEND3P2 in different cancers: DLBC(n=47); HNSC(n=501); LAML(n=132); PAAD(n=177); OS survival curves of BEND3P3 in different cancers: KIRC(n=531); LGG(n=524); LUAD(n=513); MESO(n=84); OS survival curves of BEND4 in different cancers: DLBC(n=47); SKCM(n=457); UVM(n=80); OS survival curves of BEND5 in different cancers: ACC(n=79); BLCA(n=406); BRCA(n=1082); CESC(n=293); KIRP(n=286); LUAD(n=513); MESO(n=84); PAAD(n=197); PRAD(n=499); READ(n=158); TGCT(n=139); UVM(n=80); OS survival curves of BEND6 in different cancers: BLCA(n=406); KIRC(n=531); MESO(n=84); UCEC(n=544); UVM(n=80); OS survival curves of BEND7 in different cancers: BLCA(n=406); HNSC(n=501); KIRC(n=531); LGG(n=524); UCEC(n=544).ACC, Adrenocortical carcinoma; BLCA, Bladder Urothelial Carcinoma; BRCA,Breast invasive carcinoma; CESC,Cervical squamous cell carcinoma and endocervical adenocarcinoma; COAD,Colon adenocarcinoma; DLBC,Lymphoid Neoplasm Diffuse Large B-cell Lymphoma; HNSC,Head and Neck squamous cell carcinoma; KIRC,Kidney renal clear cell carcinoma; KIRP,Kidney renal papillary cell

carcinoma; LAML,Acute Myeloid Leukemia; LGG,Brain Lower Grade Glioma; LIHC,Liver hepatocellular carcinoma; LUAD,Lung adenocarcinoma; MESO,Mesothelioma; PAAD,Pancreatic adenocarcinoma; PRAD,Prostate adenocarcinoma; READ,Rectum adenocarcinoma; SARC,Sarcoma; SKCM,Skin Cutaneous Melanoma; TGCT,Testicular Germ Cell Tumors; UCEC,Uterine Corpus Endometrial Carcinoma; UVM,Uveal Melanoma.

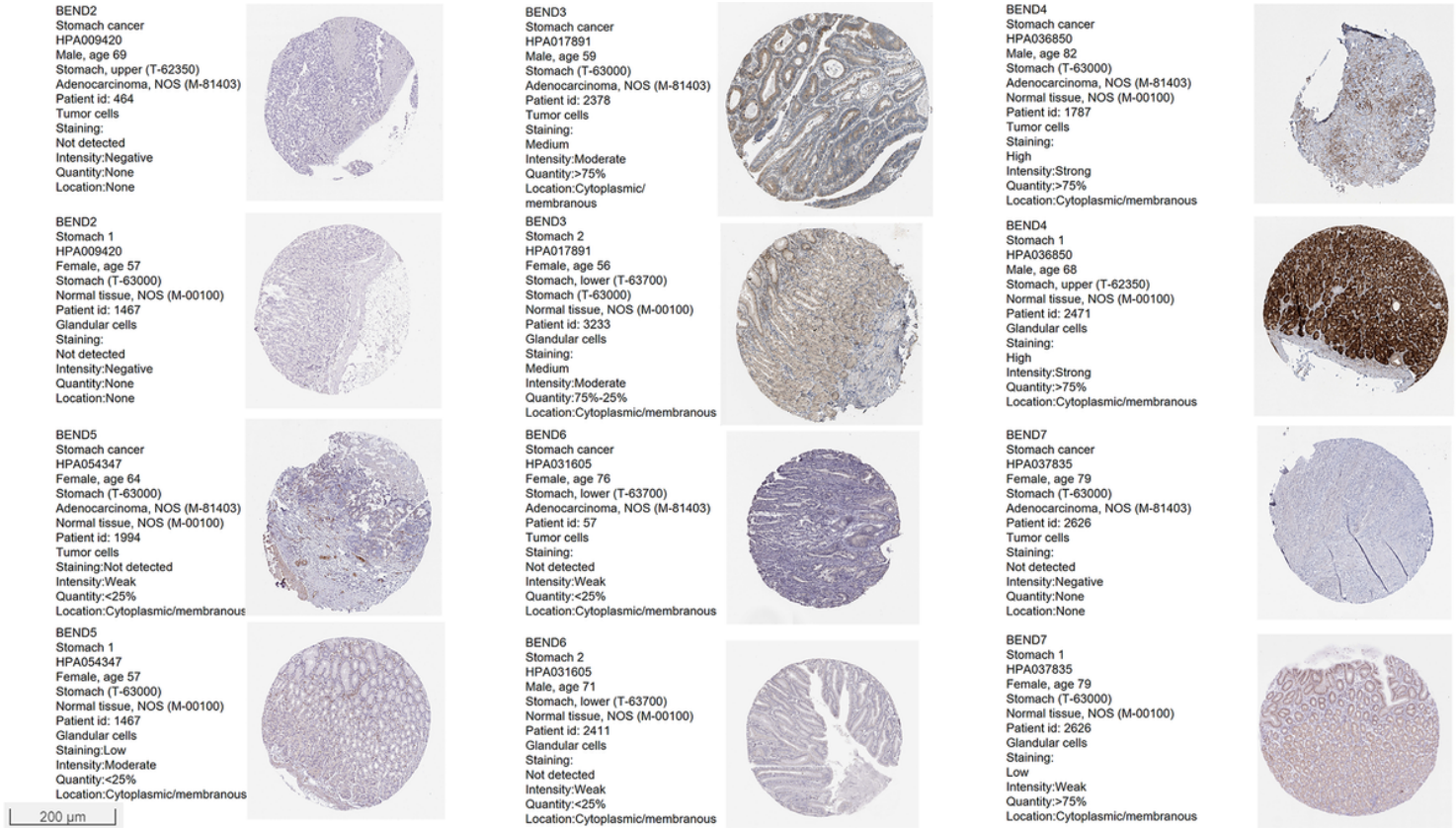

Figure 5

Correlation analysis of BEND family gene expression with survival by the COX method in different types of cancers. Different colored lines indicate the risk value of different genes in tumors, hazard ratio <1 represents a low risk, and hazard ratio >1 represents a high risk.

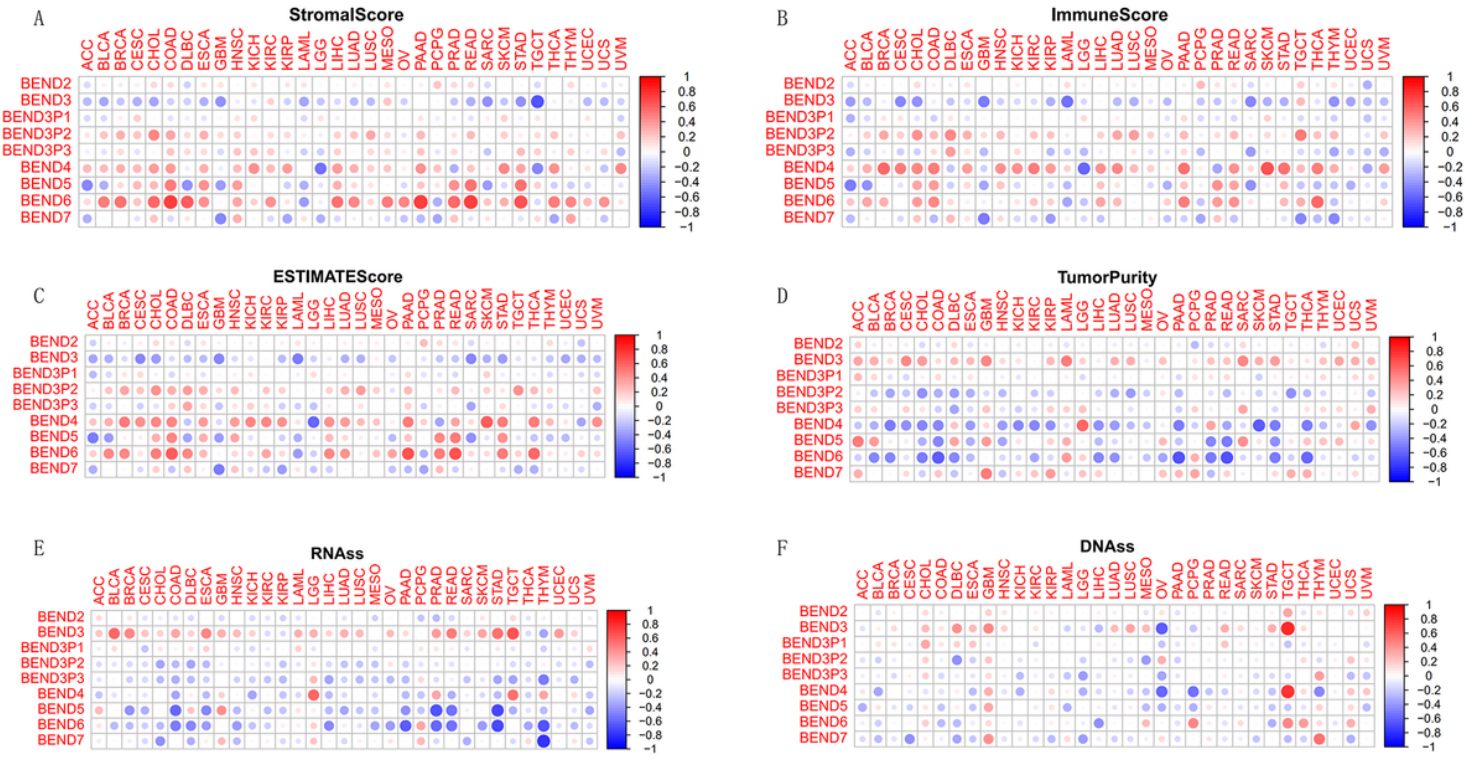

Figure 6

Correlation of BEND family gene expression with the tumor microenvironment, Stemness index in pan-cancers. (A, B, C) BEND family gene expression associated with Stromal Score, Immune Score and ESTIMATE Scores in different cancers. Red dots indicate a positive correlation between gene expression in the tumor and stromal score, and blue dots indicate a negative correlation. (D,E,F) BEND family gene expression associated with Tumor Purity,RNAss, and DNAss in different cancers. Red dots indicate a positive correlation between gene expression in the tumor and immune score, and blue dots indicate a negative correlation. RNAss, RNA stemness index DNAss, DNA stemness index. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

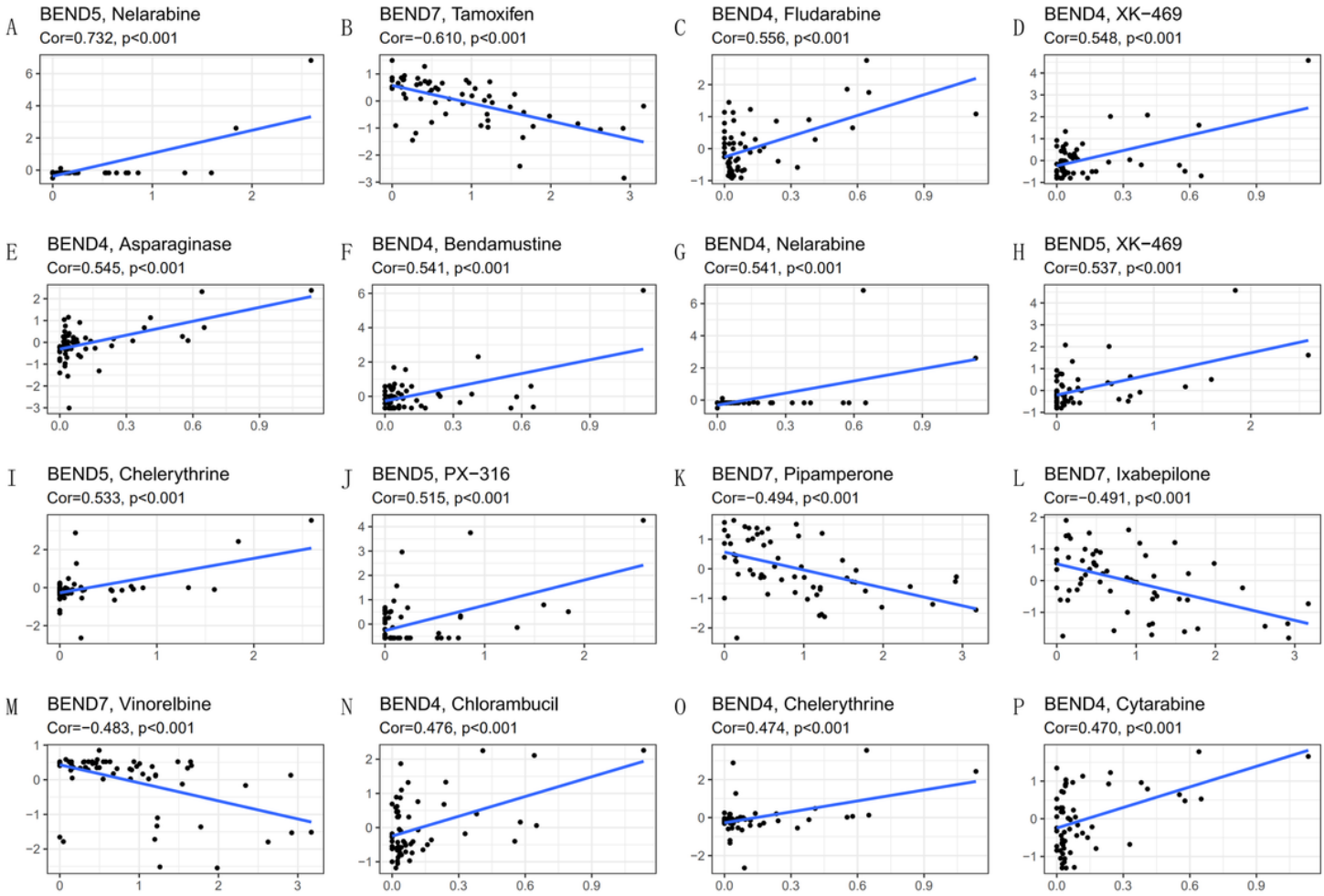

Figure 7

Drug sensitivity analysis of BEND family gene. BEND4 expression was positively associated with drug sensitivity of Fludarabine(C)、XK-469(D)、Asparaginase(E)、Benddamustine(F)、Nelarabine(G)、Chlorambucil(N)、Chelerythrine(O) and Cytarabine(P); BEND5 expression was positively associated with drug sensitivity of Nelarabine(A)、XK-469(H)、Chelerythrine(I) and PX-316(J); BEND7 expression was negatively associated with drug sensitivity of Tamoxifen(B)、Pipamperone(K)、Ixabepilone(L) and Vinorelbine(M). Cor, correlation coefficient.

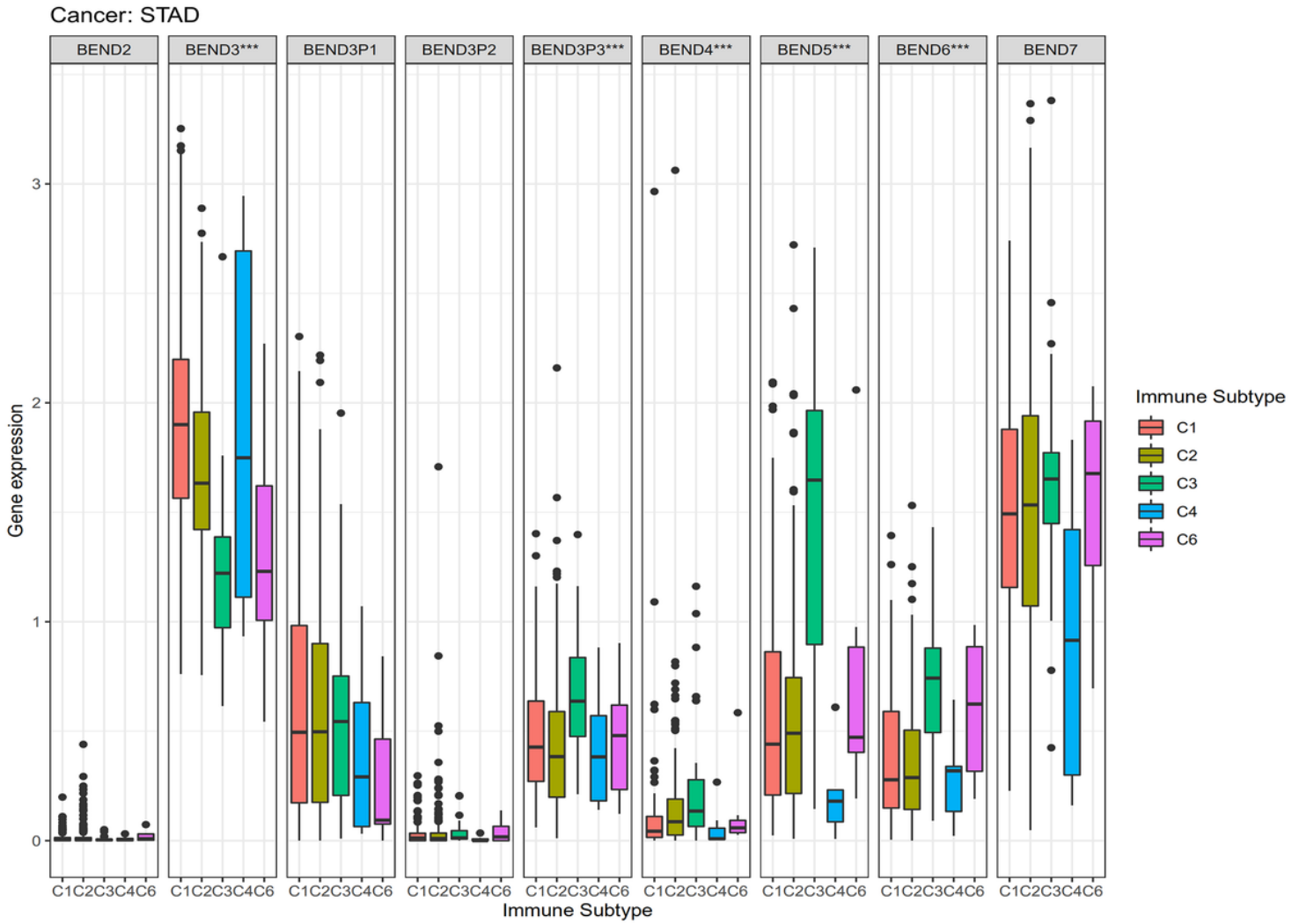

Figure 8

BEND family gene expression level of different immune subtypes in STAD. The X-axis represents the immune subtype, the Y-axis represents gene expression. C1, wound healing; C2, IFN-g dominant; C3, inflammatory; C4, lymphocyte depleted; C5, immunologically quiet; C6, TGF-b dominant. P<0.05; ** P<0.01; *** P<0.001.

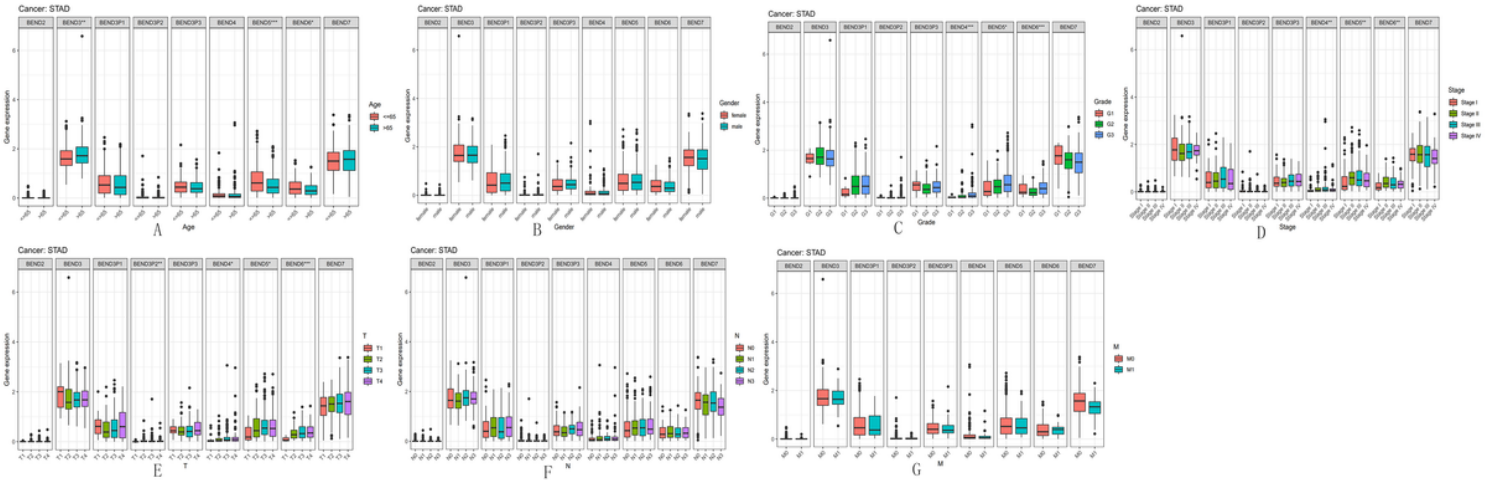

Figure 9

BEND family gene expression level of Clinical relevance in STAD. The X-axis represents different ages、Gender、Grade、Stage and TNM Stage, the Y-axis represents gene expression. P<0.05; ** P<0.01; *** P<0.001.

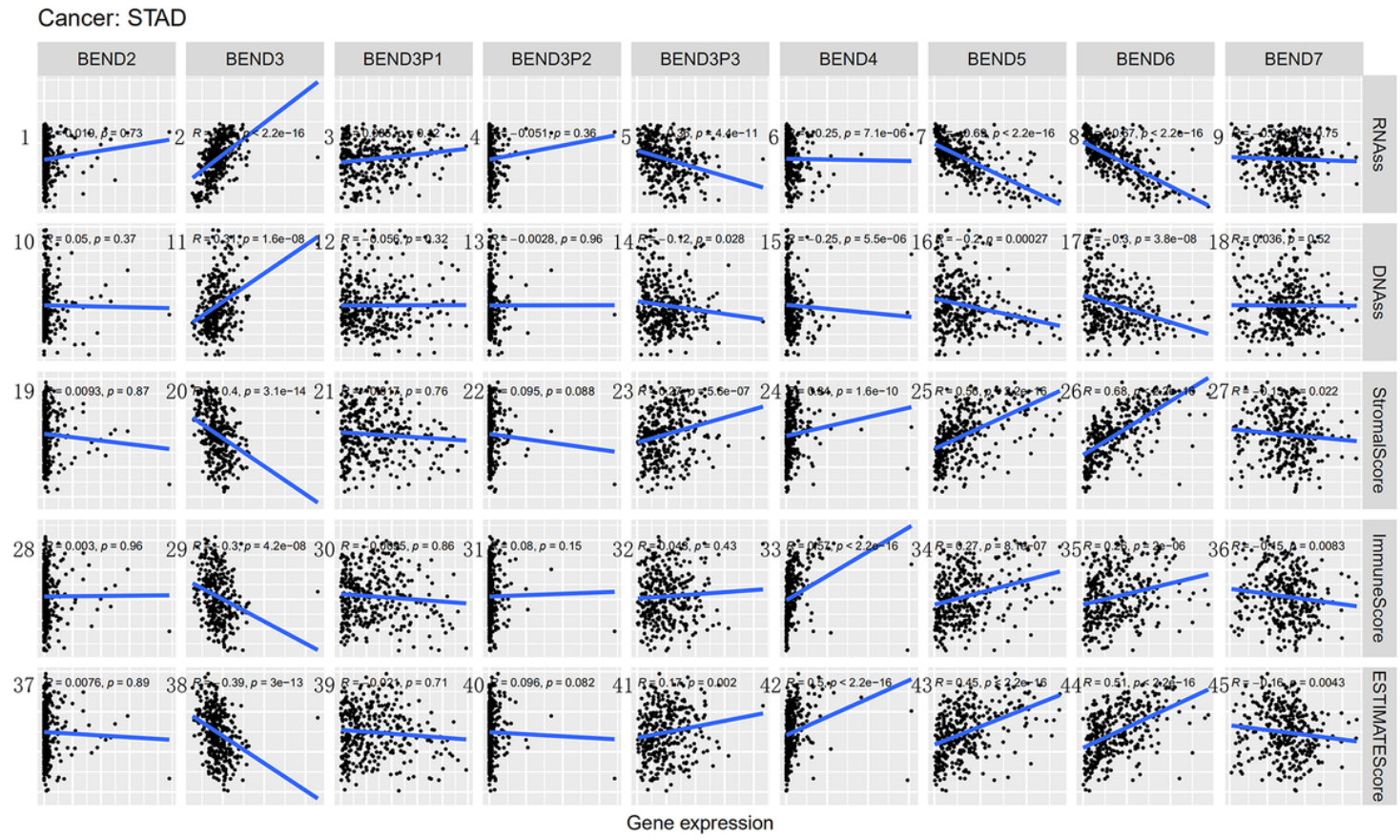

Figure 10

Correlation analysis of BEND family gene expression with Stemness index, tumor microenvironment in STAD. (1–9) Correlation analysis between BEND family gene expression and RNAss; (10-18) Correlation analysis between BEND family gene expression and DNAss; (19–27) Correlation analysis between BEND family gene expression and stromal score; (28-36) Correlation analysis between BEND family gene expression and immune score; (37-45) Correlation analysis between BEND family gene expression and ESTIMATE score. Light brown background indicates that the gene is significantly related to the corresponding index (P<0.05). Gray background indicates that the gene does not correlate with the corresponding index (P>0.05). R means correlation value, a positive number means positive correlation, a negative number means negative correlation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

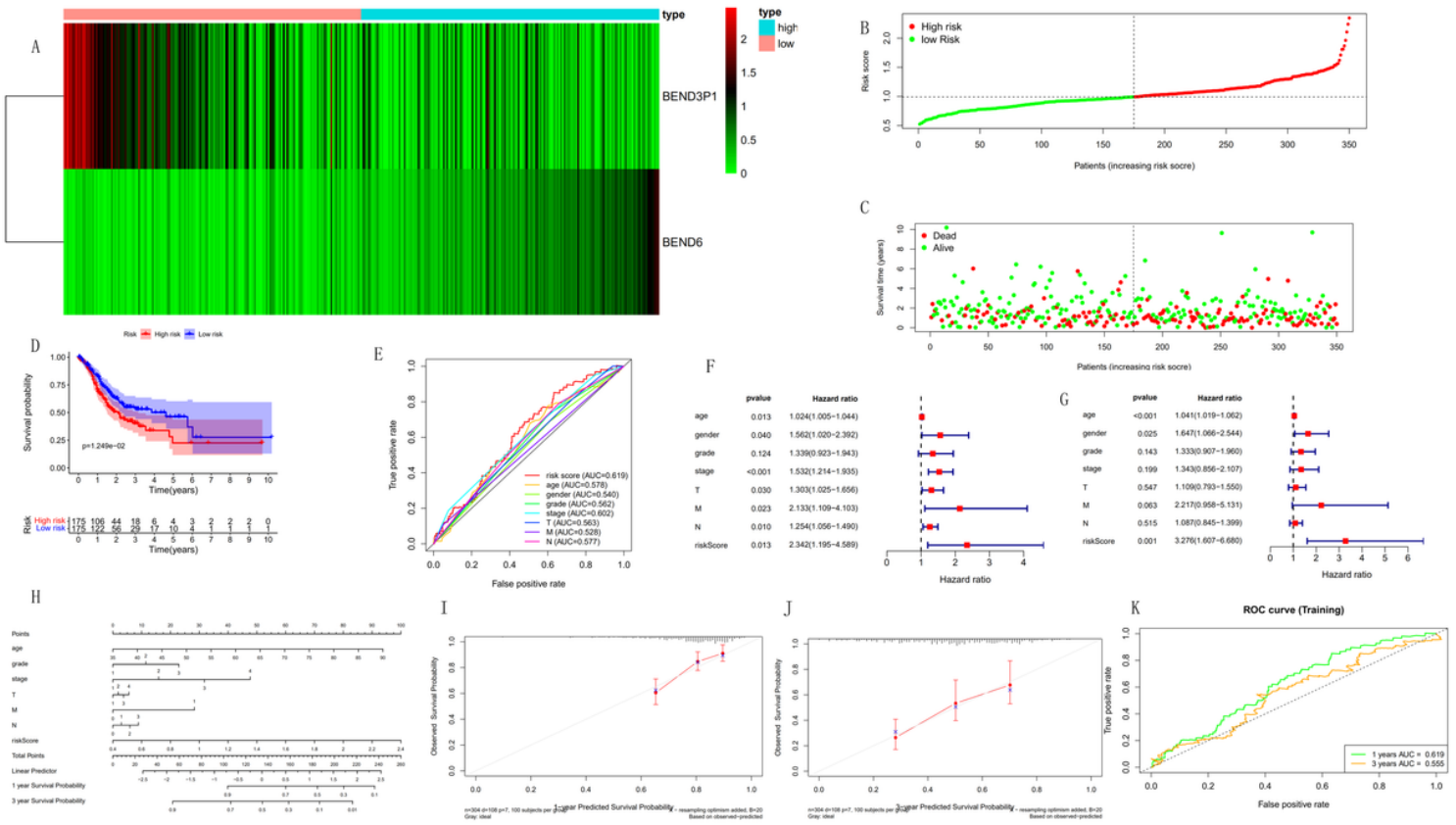

Figure 11

A BEND risk model predicted OS and DFS in patients with STAD.（A）The heatmap of the 2 related gene expression profiles in high- and low-risk STAD patients. (B)The distribution of risk scores of high- and low-risk STAD patients. (C)The scatter plot shows the correlation between survival time and risk score. （D）The K-M curve reflects that the OS of high-risk STAD patients is significantly lower than that of low-risk patients (P =1.249e-02). （E）The forest plot reflects the univariate Cox analysis of the relationship between the clinical features, risk score, and OS of STAD. (F)The forest plot reflects multivariate Cox analyzed the relationship between the clinical features, risk score, and OS of STAD patients. risk score (P =0.006) is an independent prognostic risk factor for STAD.（G）ROC curve reflects the high predictive value of the risk score (AUC = 0.619). （H）Calculate the scores of each item of STAD patients according to the nomogram, and the total scores obtained after addition can predict the 1- and 3-year survival probability. （I）（J）The 1- and 3-year calibration curves of the nomogram. （K）The ROC curves of 1-and 3-year nomogram (AUC = 0.619 of 1 year, AUC = 0.555 of 3 years).

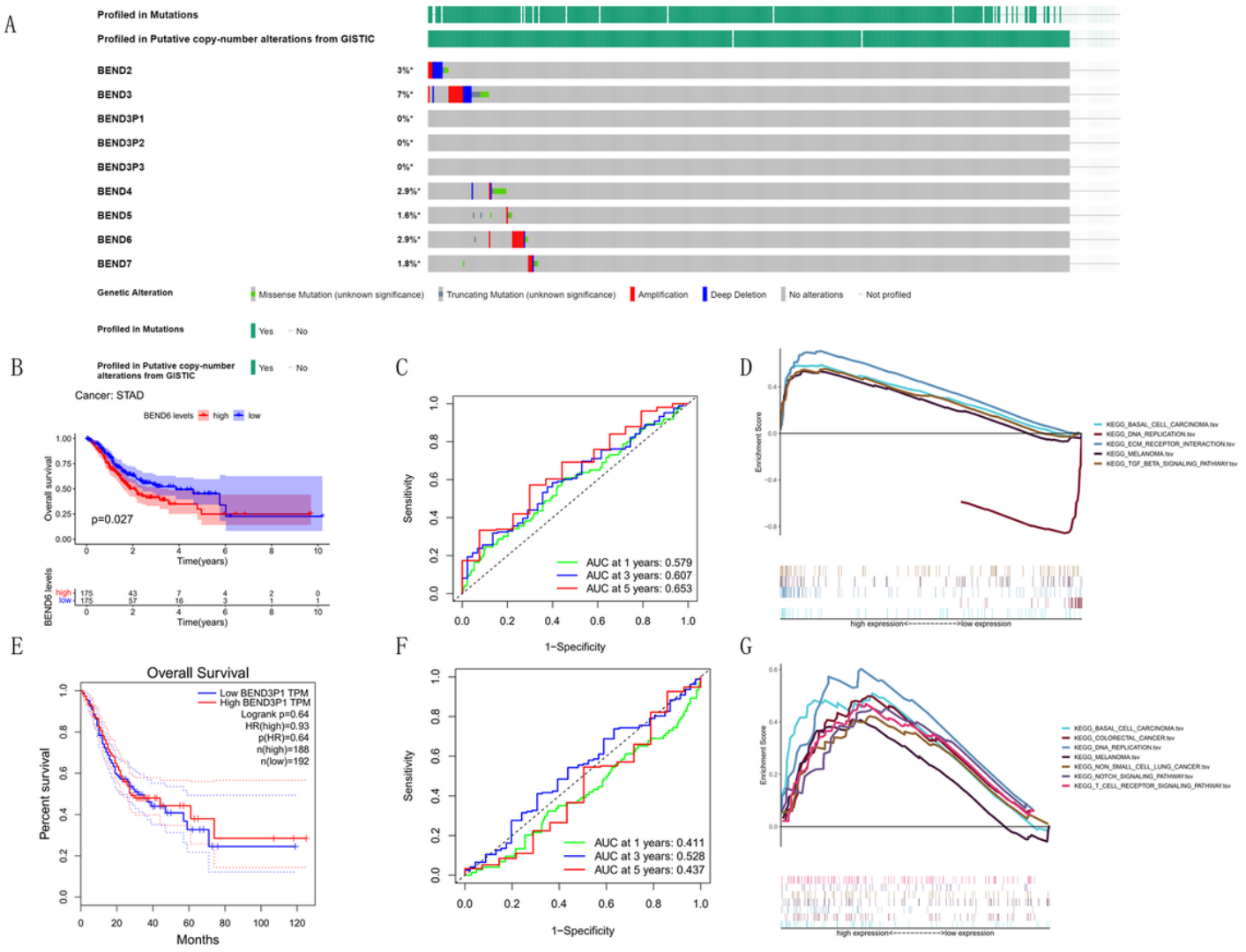

Figure 12

Relationship between BEND scores and gene expression profiles

Figure 13

Expression of BEND6 related to tumor immune infiltration. (A)The number of immune cells in each sample.（B）Results of differential analysis of immune cells. Green represents the low expression group and red represents the high expression group. （C）Correlation analysis of immune cells. （D）The intersection of the two test results.

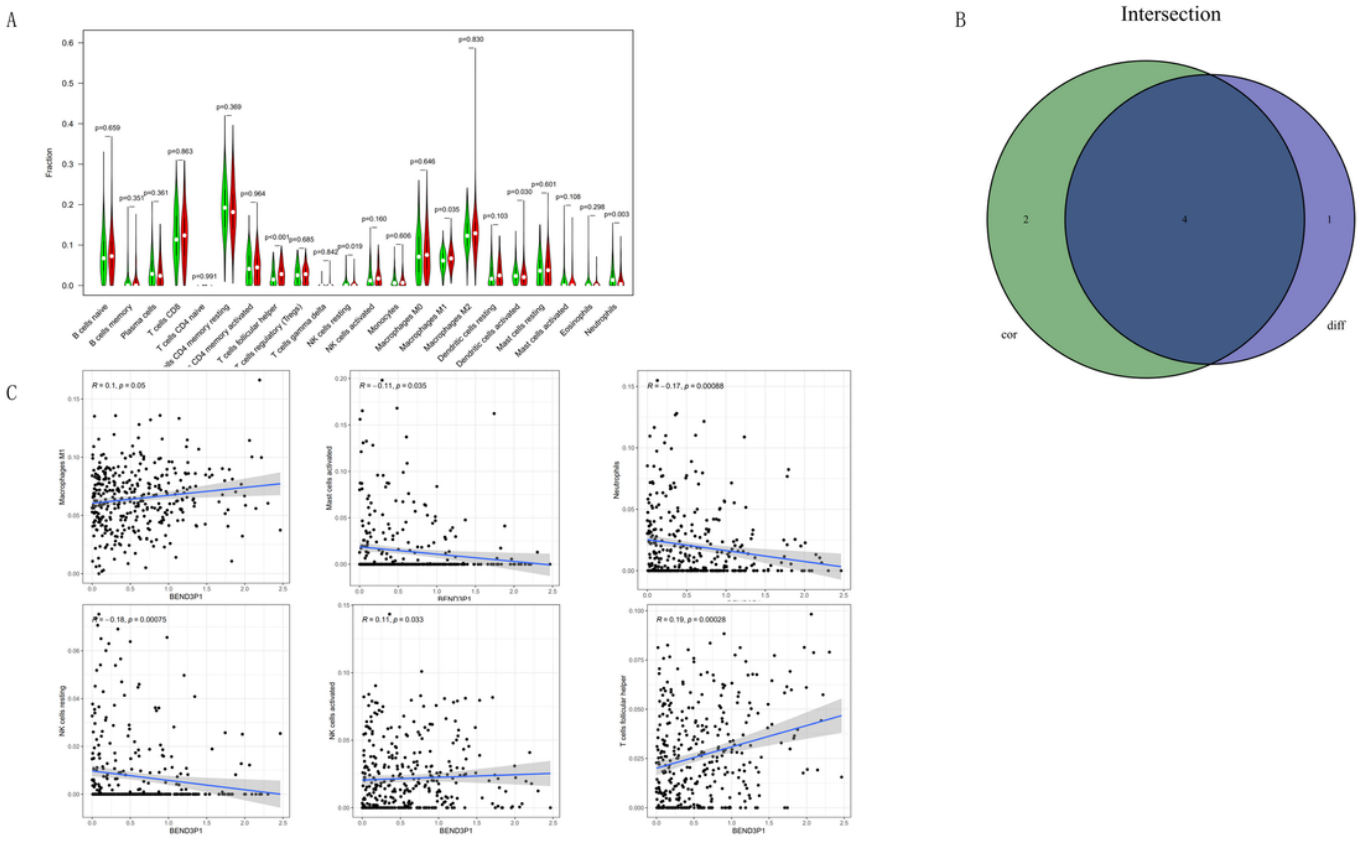

Figure 14

Expression of BEND3P1 related to tumor immune infiltration. (A)Results of differential analysis of immune cells. Green represents the low expression group and red represents the high expression group. (B)The intersection of the two test results. (C)Correlation analysis of immune cells.