


Pan-cancer analysis of the prognostic significance and oncogenic role of GXYLT2

Yi-Bei Song, MM^a, Wen-Guang Bao, MM^b, Deng-He Liu, MM^a, Li-Qiang Wei, MM^a, Shu-Ting Yang, BD^a, Xue-Jing Miao, BD^a, Chun-Yu Lin, BD^a, Hong-Jun Li, BD^c, Dong Lan, MM^b, Hui-Min He, MD^{c,*} 

Abstract

Growing evidence supports an oncogenic role for glucoside xylosyltransferase 2 (GXYLT2) in a number of malignancies. To evaluate the prognostic value and oncogenic function of GXYLT2 in diverse cancer types, we analyzed sequencing data from public databases on 33 tumor tissues and their corresponding normal tissues. We found that GXYLT2 was overexpressed in a number of tumors, and that its expression was positively correlated with disease progression and mortality in several major cancer types including stomach adenocarcinoma (STAD). GXYLT2 was also linked to tumor size, grade, and the immune and molecular subtypes of STAD. GO and KEGG pathway analyses of GXYLT2 co-expressed genes in STAD suggested that GXYLT2 possibly plays a role in epithelial-mesenchymal transition, extracellular matrix production and degradation, angiogenesis, apoptosis, as well as in tumor inflammation, such as cytokine production and T cell activation. Finally, prognostic nomograms were created and validated for predicting 1, 3, and 5-year survival of patients with STAD. Our findings indicate that GXYLT2 may play a role in tumorigenesis and tumor immunity, and it may serve as a prognostic marker and potential immunotherapeutic target for STAD and some other types of cancer.

Abbreviations: BLCA = bladder urothelial carcinoma, BRCA = breast cancer, CHOL = cholangiocarcinoma, COAD = colon adenocarcinoma, DFI = disease-free interval, DFI = disease-free interval, DSS = disease-specific survival, DSS = disease-specific survival, ECM = extracellular matrix, ESCA = esophagus cancer, GBM = glioblastoma multiforme, GO = Gene Ontology, GTEx = The Genotype-Tissue Expression database, GXYLT2 = Glucoside Xylosyltransferase 2, HNSC = head and neck squamous cell carcinoma, HR = hazard ratio, IHC = immunohistochemistry, KEGG = Kyoto Encyclopedia of Genes and Genomes, KICH = kidney chromophobe, KIRC = kidney renal clear cell carcinoma, LGG = low-grade gliomas, LIHC = liver hepatocellular carcinoma, LUAD = lung adenocarcinoma, MSI = microsatellite instability, OS = overall survival, OS = overall survival, PAAD = pancreatic adenocarcinoma, PFI = progression-free interval, PFI = progression-free interval, PRAD = prostate adenocarcinoma, STAD = stomach adenocarcinoma, TCGA = The Cancer Genome Atlas, THCA = thyroid carcinoma, TIME = tumor immune microenvironment, TIMER = Tumor Immune Estimation Resource, TMB = tumor mutational burden, TME = tumor microenvironment, UCEC = uterine corpus endometrial carcinoma.

Keywords: GXYLT2, MSI, pan-cancer, prognosis, STAD, TMB, tumor immunity

1. Introduction

Glucoside xylosyltransferase 2 (GXYLT2), also known as GLT8D4, encodes a human protein of 443 amino acids with α -1, 3-D-xylosyltransferase activity.^[1] This enzyme shows common structural features of a type II Golgi glycosyltransferase,

including a short cytoplasmic domain at the N-terminus, a catalytic domain at the C-terminus, and a transmembrane domain. In terms of functions, GXYLT2 has been shown to transfer the first xylose to O-glucosylated residues on epidermal growth factor repeats of Notch, a key regulator of carcinogenesis.^[1–3]

Y-BS and W-GB contributed equally to this work.

The Guangxi Health Department Research Project (Z-A20230495); and the Clinical Research "Climbing" Program of the First Affiliated Hospital of Guangxi Medical University (YYZS2020023).

The authors have no conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

This project was approved by the Ethical Approval Committee of the First Affiliated Hospital of Guangxi Medical University (No. 2022-KT-transverse item-029), and the privacy of patients was well protected during the process of this project. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Supplemental Digital Content is available for this article.

^a Department of Clinical Laboratory, the First Affiliated Hospital of Guangxi Medical University, Key Laboratory of Clinical Laboratory Medicine of Guangxi

Department of Education, Nanning, China, ^b Department of Medical Oncology, the First Affiliated Hospital of Guangxi Medical University, Nanning, China, ^c Guangxi Medical University, Nanning, China.

*Correspondence: Hui-Min He, Guangxi Medical University, 22 Shuangyong Road, Nanning, Guangxi Zhuang Autonomous Region 530021, China (e-mail: hehuimin@sr.gxmu.edu.cn).

Copyright © 2023 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

How to cite this article: Song Y-B, Bao W-G, Liu D-H, Wei L-Q, Yang S-T, Miao X-J, Lin C-Y, Li H-J, Lan D, He H-M. Pan-cancer analysis of the prognostic significance and oncogenic role of GXYLT2. *Medicine* 2023;102:46(e35664).

Received: 15 August 2023 / Received in final form: 24 September 2023 / Accepted: 25 September 2023

<http://dx.doi.org/10.1097/MD.0000000000035664>

Further studies have implicated GXYLT2 in the development of a number of cancers. For example, GXYLT2 was found to be overexpressed in human stomach cancer, and its expression is correlated with disease progression and poor prognosis.^[4] In human bladder cancer, GXYLT2 was closely associated with tumor immune cell infiltration and immune gene expression.^[5] In *in vitro* studies, GXYLT2 overexpression in human breast and gastric cancer cells upregulated Notch/MARK signaling and accelerated cell proliferation and migration.^[6] Nonetheless, the prognostic significance and oncogenic role of GXYLT2 in most cancer types remain unclear.

Pan-cancer analysis aims to uncover common patterns of gene mutations or aberrant expression across diverse tumor types.^[7] According to recent pan-cancer analyses, a typical tumor type has, on average, 4 or 5 driver gene mutations.^[8,9] In recent years, pan-cancer analysis has gained traction as an effective approach to identifying new diagnostic/prognostic markers or therapeutic targets for cancer treatment.^[10,11] In this study, we performed a pan-cancer analysis of GXYLT2

expression across diverse tumor types using data from public cancer genome databases. The associations of GXYLT2 with disease progression, prognosis, tumor mutational burden (TMB), microsatellite instability (MSI), and tumor immune cell infiltration were investigated for each tumor type, with a special focus on stomach adenocarcinoma (STAD). Finally, we conducted Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses of GXYLT2 co-expressed genes in STAD to determine the functional classification of GXYLT2.

2. Methods

2.1. Data collection and differential expression analysis

GXYLT2 expression data for 33 types of human cancer and matching normal tissues were downloaded from OncoPrint^[12] and The Cancer Genome Atlas (TCGA).^[13] GXYLT2 expression data in normal human tissues were also extracted from

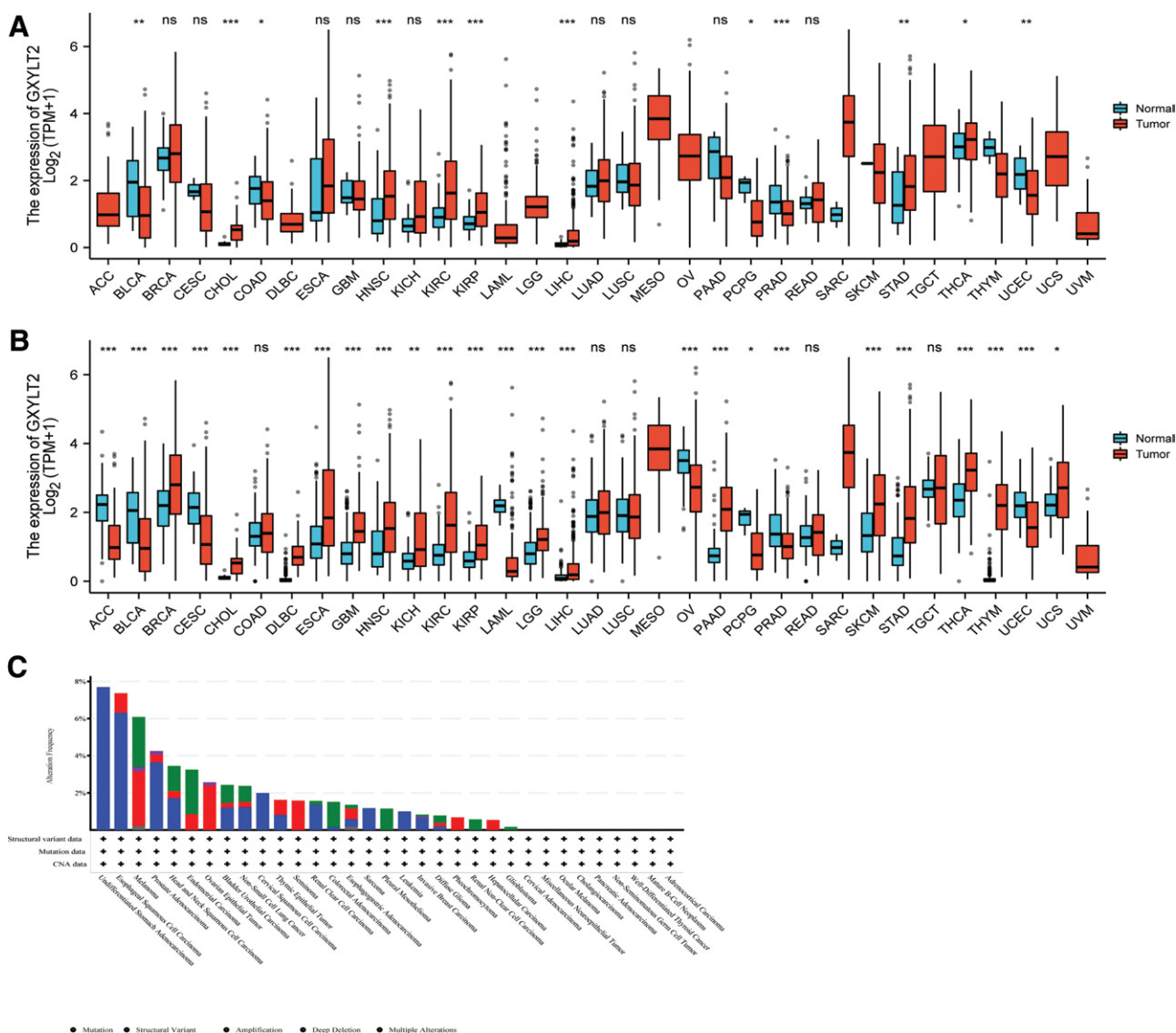


Figure 1. (A and B) GXYLT2 expression in 33 TCGA tumor types and corresponding normal control tissues. (A) GXYLT2 expression in tumor and normal tissues. All expression data were extracted from the TCGA database. **P* < .05; ***P* < .01; ****P* < .001, and ns indicates no significance. (B) GXYLT2 expression in tumors and normal tissues. Expression data for normal tissues were extracted from the TCGA and GTEx databases and merged. **P* < .05; ***P* < .01; ****P* < .001, and ns indicates no significance. (C) Genetic alterations of GXYLT2 in 33 TCGA tumor types analyzed through cBioPortal. GTEx = Genotype-Tissue Expression database, GXYLT2 = Glucoside Xylosyltransferase 2, TCGA = The Cancer Genome Atlas.

The Genotype-Tissue Expression database (GTEx; <https://www.genome.gov/Funded-Programs-Projects/Genotype-Tissue-Expression-Project>).^[14] GXYL2 protein expression data in tumor and normal tissues were extracted from The Human Protein Atlas (<https://www.proteinatlas.org>). All expression data were log2 transformed. Two-group *t* tests were used to evaluate the differences in expression between normal and malignant tissues, and *P* < .05 was considered statistically significant.

2.2. Analysis of the associations between GXYL2 and disease progression and survival

A Kaplan–Meier analysis was used to evaluate the relationship between GXYL2 expression and progression/prognosis markers including disease-free interval (DFI), progression-free interval (PFI), disease-specific survival (DSS), and overall survival (OS). The median value of GXYL2 expression was used as the cutoff between low and high

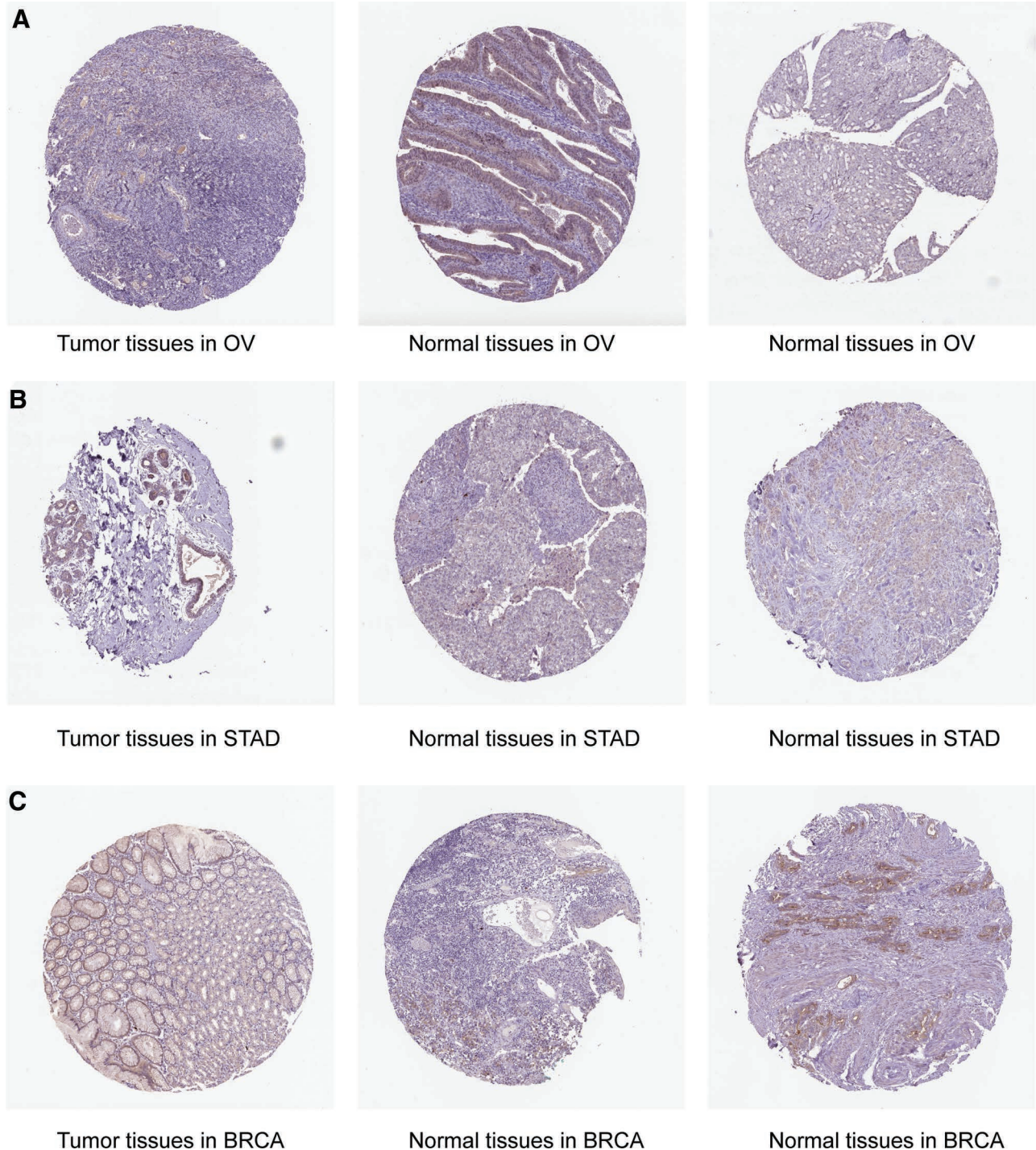


Figure 2. Representative IHC images from the HPA database showing GXYL2 protein expression in OV (A), STAD (B), and BRCA (C) tumor and normal tissues. Images on the left represent tumor tissues. Images in the middle and on the right represent normal tissues. BRCA = breast cancer, GXYL2 = Glucoside Xylosyltransferase 2, HPA = Human Protein Atlas, IHC = immunohistochemistry, OV = ovarian cancer, STAD = stomach adenocarcinoma.

GXYLT2 expression. In a univariate survival analysis, the hazard ratio (HR) and 95% confidence interval were calculated using the Kaplan–Meier plotter (<http://kmplot.com/> analysis).

2.3. Evaluation of genetic alterations in GXYLT2

Genetic alterations in GXYLT2 across all 33 TCGA tumor types were examined using the cBioPortal for Cancer Genomics (<https://www.cbioportal.org>). TMB was calculated using Perl

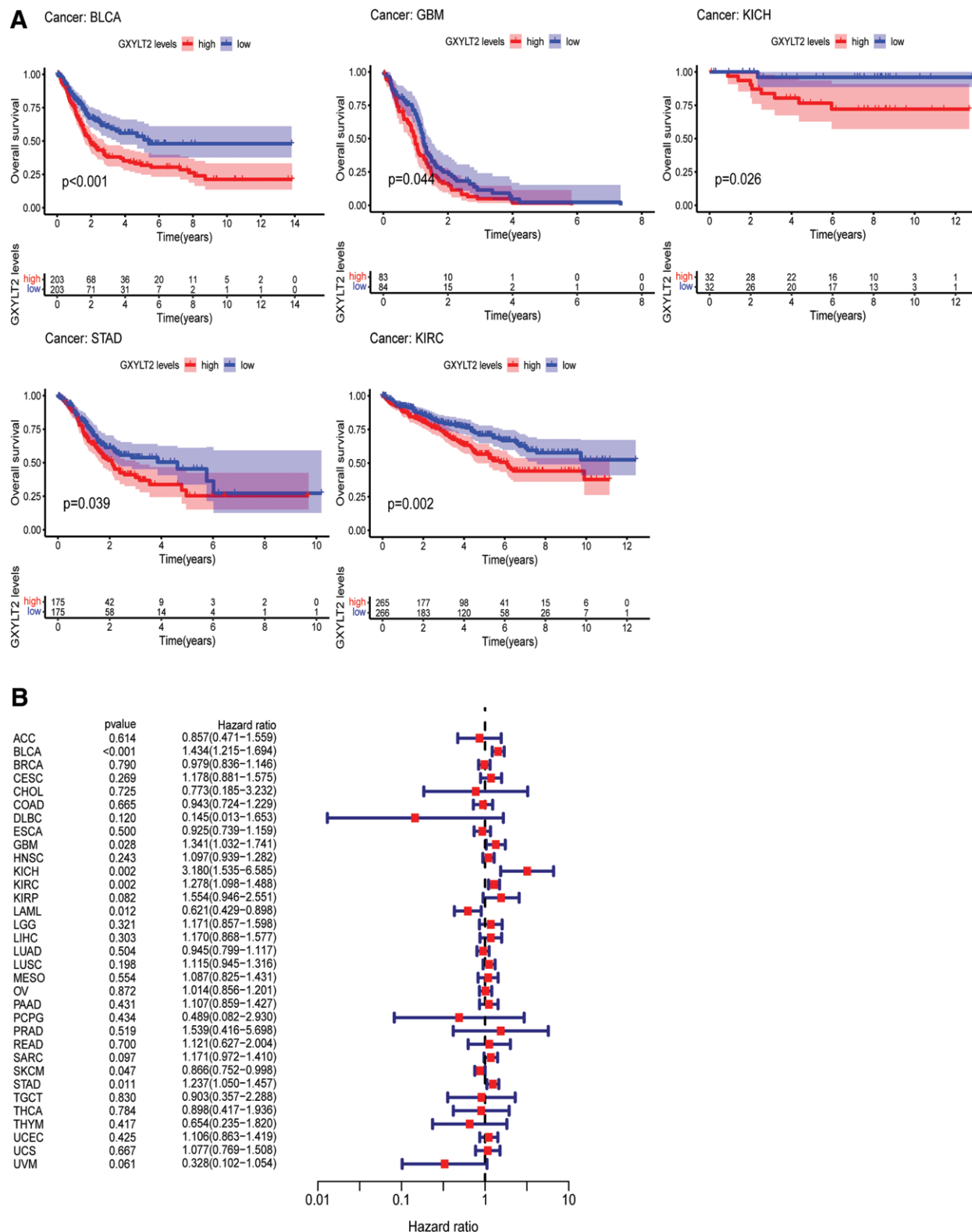


Figure 3. Correlation between GXYLT2 expression and OS. (A) Kaplan–Meier curves showing the correlation between GXYLT2 expression and OS in BLCA, GBM, KICH, STAD, and KIRC. (B) Forest plots showing the relationship between GXYLT2 expression and OS in the 33 TCGA tumor types. BLCA = bladder urothelial carcinoma, GBM = glioblastoma multiforme, GXYLT2 = Glucoside Xylosyltransferase 2, KICH = kidney chromophobe, KIRC = kidney renal clear cell carcinoma, OS = overall survival, STAD = stomach adenocarcinoma, TCGA = The Cancer Genome Atlas.

scripts based on the total number of somatic mutations per million bases. MSI scores were calculated based on DNA-seq data from TCGA. The correlation between GXYLT2 expression and TMB or MSI was assessed using Spearman's test by utilizing the cor.test tool package of R software (<https://www.r-project.org>). Radar plots showing the correlations were created using the radar chart function of the fmsb package in R.

2.4. Relationship between GXYLT2 expression and tumor immune microenvironment

Tumor immune infiltration plays an important role in cancer development. TIMER (Tumor Immune Estimation Resource) is

a public resource that provides cancer researches with computational tools to estimate tumor immune cell infiltration from gene expression profiles.^[15] In this study, we used TIMER to estimate the abundance of different types of tumor-infiltrating immune cells in the 33 TCGA tumors. Immune and stromal scores were computed using the ESTIMATE method, via the estimate and limma tool packages in R. Tumor purity and the percentages of stromal and specific immune cell types in tumor tissues were evaluated using CIBERSORT,^[16] xCELL,^[17] QUANTISEQ,^[18] EPIC,^[19] and MCPOUNTER.^[20] The relationship between GXYLT2 expression and immune cell infiltration was assessed using the ggplot2, ggpubr, and ggExtra packages in R.

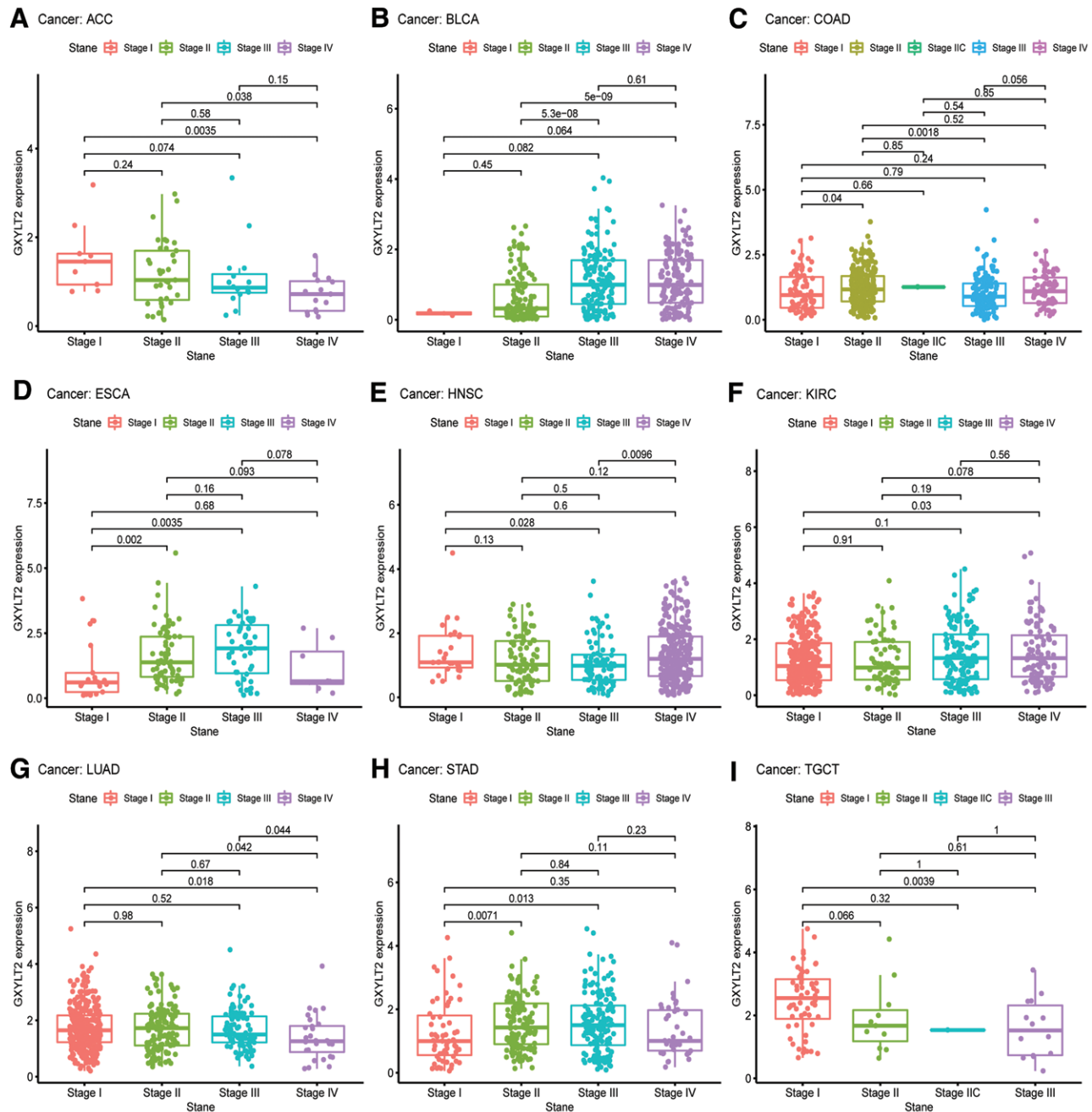


Figure 4. Correlation between GXYLT2 expression and tumor stage in ACC (A), BLCA (B), COAD (C), ESCA (D), HNSC (E), KIRC (F), LUAD (G), STAD (H), and TGCT (I). BLCA = bladder urothelial carcinoma, COAD = colon adenocarcinoma, GXYLT2 = Glucoside Xylosyltransferase 2, HNSC = head and neck squamous cell carcinoma, KIRC = kidney renal clear cell carcinoma, LUAD = lung adenocarcinoma, STAD = stomach adenocarcinoma, TGCT = testicular germ cell tumors.

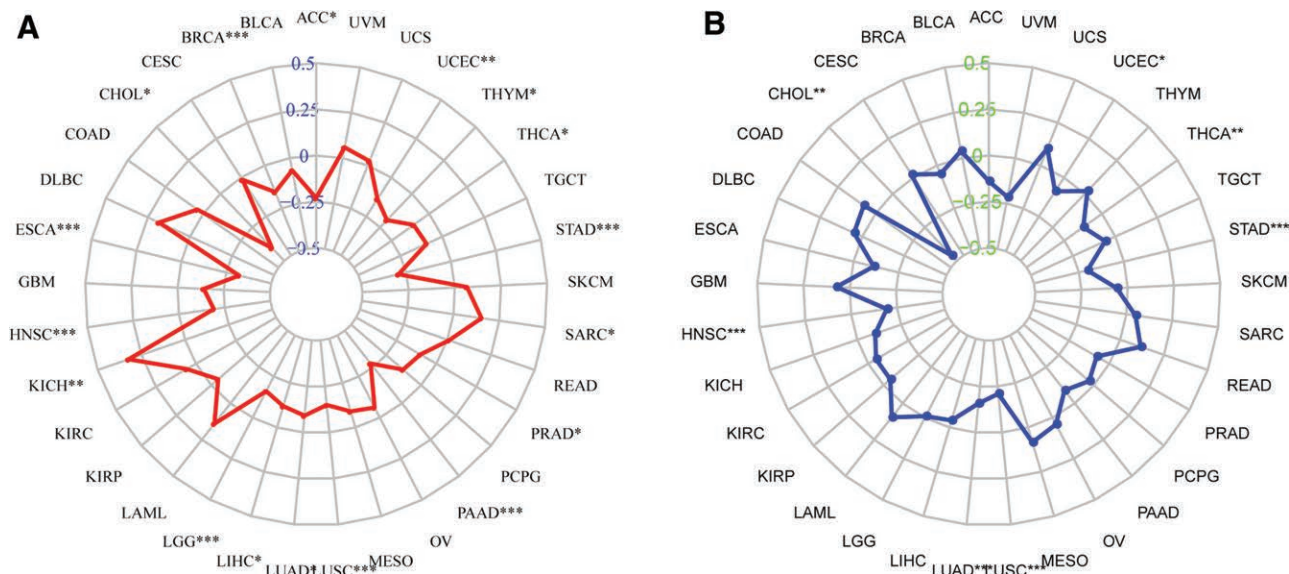


Figure 5. Radar plots showing association of GXYL2 expression with TMB (A) and MSI (B) in 33 TCGA tumor types. Lines and numbers in each plot represent correlation coefficients and ranges, respectively. * $P < .05$, ** $P < .01$, and *** $P < .001$. GXYL2 = Glucoside Xylosyltransferase 2, MSI = microsatellite instability, TCGA = The Cancer Genome Atlas, TMB = tumor mutational burden.

2.5. GO and KEGG pathway enrichment analysis

GXYLT2 co-expression analysis was conducted based on sequencing data from TCGA using the limma, reshape2, and RColorBrewer tool packages in R. Functional pathways linked to GXYL2 were predicted by GO and KEGG pathway enrichment analysis using the limma, org.Hs.e.g..db, clusterProfiler, and enrichplot packages in R.

2.6. Collection of human STAD specimens

Tumor tissues were collected from 100 patients with a pathologically confirmed STAD diagnosis and who underwent surgery between 2020 and 2022 at The First Affiliated Hospital of Guangxi Medical University (Nanning, Guangxi, China). The study protocol was approved by the Ethics Committee of The First Affiliated Hospital of Guangxi Medical University (No. 2022-KT-Transverse Item-029). All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. All patients provided written informed consent.

2.7. Immunohistochemistry

Immunohistochemistry (IHC) was conducted according to previously described procedures.^[21] In brief, paraffin-embedded STAD tissue sections were deparaffinized and rehydrated. Antigen retrieval was performed by soaking samples in citrate buffer (pH 6.0) at 95°C for 15 minutes followed by incubating in 0.3% hydrogen peroxide at room temperature for 15 minutes to inactivate endogenous peroxidases. After blocking in 5% goat serum (Thermo Fisher Scientific, Waltham, USA), GXYL2 was detected via IHC with an anti-GXYLT2 antibody (1:50; Abcam, Cambridge, UK). The proportion of positive gastric cells was estimated based on cytoplasm/nucleus staining intensity. Staining intensity was graded as follows: 0, negative (-); 1, weak (+); 2, medium (++); and 3, strong (+++). Based on the percentage of positive cells, the positive ratio scores are as follows: 1, 25% or less; 2, 26–50%; 3, 51–75%; and 4, 76% or more. We calculated the final score by multiplying staining intensity scores and positive ratio scores.

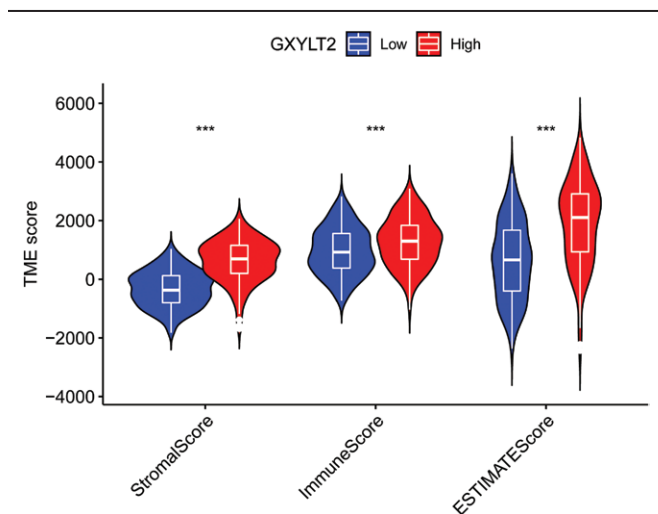


Figure 6. GXYL2 associations with stromal, immune, and ESTIMATE scores across 33 TCGA cancers. *** $P < .001$. GXYL2 = Glucoside Xylosyltransferase 2, TCGA = The Cancer Genome Atlas.

2.8. Statistical analysis

Methods used for differential expression and survival analyses were described in their corresponding methods sections. The correlations between 2 variables were evaluated using Spearman’s test or Pearson’s test, and $P < .05$ was considered to be statistically significant. All statistical studies were carried out using R software.

3. Results

3.1. Pan-cancer GXYL2 expression and genetic alterations

GXYLT2 expression data in 33 TCGA tumor types and their corresponding normal tissues were extracted from the TCGA or GTEx database as previously described and analyzed. Based on data from the TCGA database, GXYL2 expression was found to be upregulated in cholangiocarcinoma (CHOL), head and neck squamous cell carcinoma (HNSC), kidney renal clear

cell carcinoma (KIRC), kidney renal papillary cell carcinoma, liver hepatocellular carcinoma (LIHC), STAD, and thyroid carcinoma (THCA), but downregulated in bladder urothelial carcinoma (BLCA), uterine corpus endometrial carcinoma (UCEC), prostate adenocarcinoma (PRAD), colon adenocarcinoma (COAD), and pheochromocytoma and paraganglioma, compared with normal tissue controls (Fig. 1A). As expression data for normal tissues in some tumor types were not available from the TCGA, we extracted data for normal tissues from the GTEx database and used the R program to merge the data from both datasets for a more thorough comparison of GXYLT2 expression in tumor and corresponding normal tissues (Fig. 2B). The 2 figures show similar patterns of relative expression of GXYLT2 in tumor versus normal tissues, except for a few tumor types including breast cancer (BRCA), diffuse large B-cell lymphoma, esophagus cancer (ESCA), glioblastoma multiforme (GBM), kidney chromophobe (KICH), acute myeloid leukemia, low-grade gliomas (LGG), and pancreatic adenocarcinoma (PAAD) (Fig. 1A and B).

The genetic alterations of GXYLT2 across the 33 TCGA tumor types were detected through the cBioPortal. It was found that patients with undifferentiated STAD had the highest GXYLT2 alteration frequency (>7%), with “deep deletion”

being the predominant type (Fig. 1C). ESCA showed the second highest alteration frequency in the GXYLT2 gene (>6%) (Fig. 1C).

IHC images for GXYLT2 protein expression in tumor and normal tissues were extracted from the Human Protein Atlas database and analyzed. GXYLT2 protein was overexpressed in ovarian cancer (Fig. 2A), STAD (Fig. 2B), and BRCA (Fig. 2C), suggesting that GXYLT2 might play an oncogenic role in the development of these cancers.

3.2. Pan-cancer prognostic significance of GXYLT2

To evaluate the pan-cancer prognostic value of GXYLT2, we investigated the correlations between GXYLT2 expression and progression/prognosis markers including DFI, PFI, DSS, and OS across the 33 TCGA tumor types. Kaplan–Meier analyses revealed that high GXYLT2 was associated with low OS in BLCA ($P < .001$), GBM ($P = .044$), KICH ($P = .026$), KIRC ($P = .002$), and STAD ($P = .039$) (Fig. 3A), with high GXYLT2 showing the greatest predictive power for all-cause death in KICH (HR = 3.180) (Fig. 3B). In contrast, GXYLT2 was positively associated with OS in acute myeloid leukemia ($P = .012$) and skin cutaneous melanoma (SKCM) ($P = .047$)

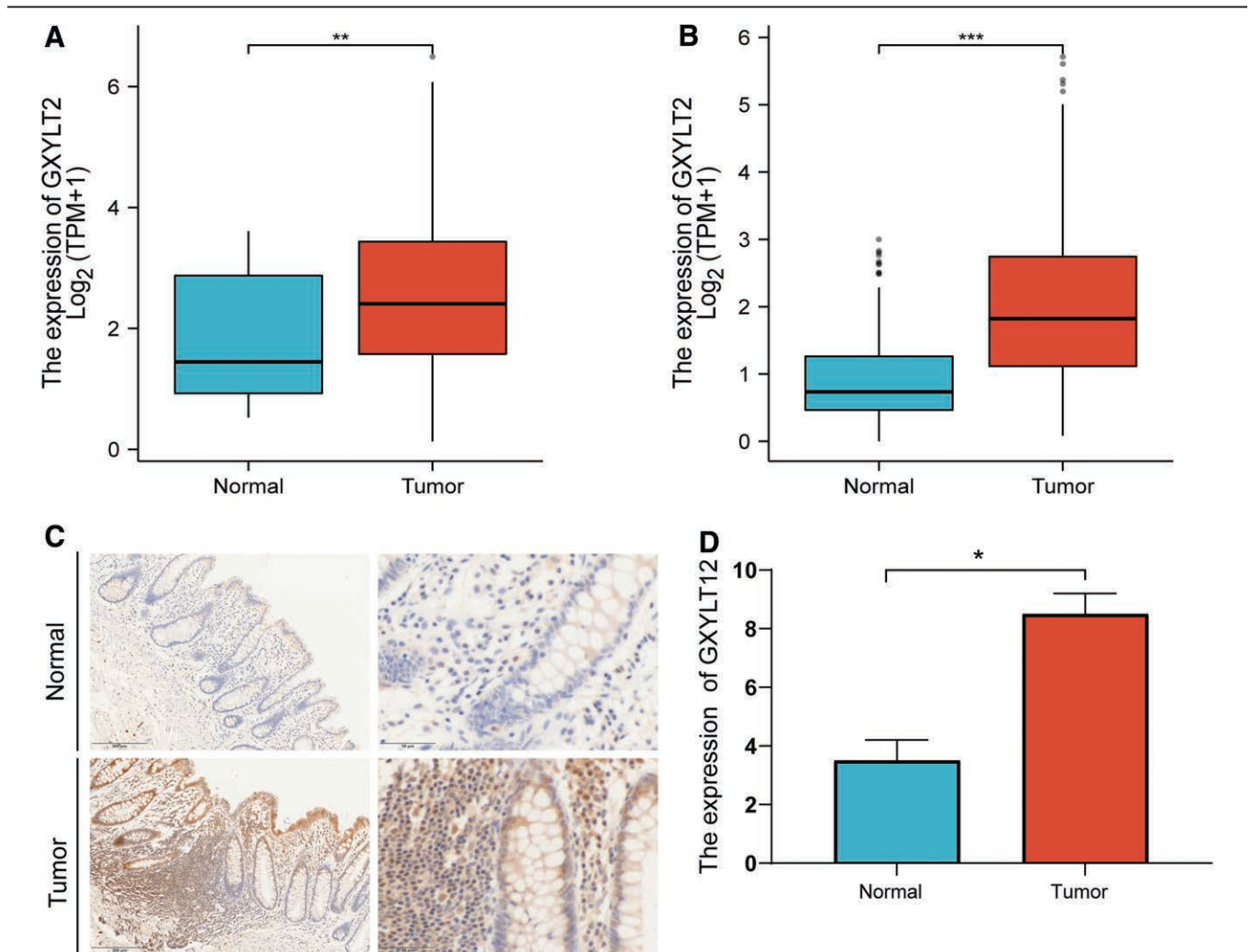


Figure 7. GXYLT2 expression in human STAD. (A) GXYLT2 mRNA expression in human STAD and normal tissues based on TCGA data. (B) GXYLT2 mRNA expression in human STAD and normal tissues based on combined GTEx and TCGA data. (C) Representative IHC images showing GXYLT2 protein expression in normal (upper panel) and STAD (lower panel) human tissue. IHC experiments were conducted in house. * $P < .05$, ** $P < .01$, and *** $P < .001$. GTEx = Genotype-Tissue Expression database, GXYLT2 = Glucoside Xylosyltransferase 2, IHC = immunohistochemistry, STAD = stomach adenocarcinoma, TCGA = The Cancer Genome Atlas.

(Fig. 3B). Similarly, high GXYLT2 was linked to low DSS in patients with BLCA ($P = .003$), GBM ($P = .026$), KIRC ($P < .001$), and STAD ($P = .013$) (Figure S1A, Supplemental Digital

Content, <http://links.lww.com/MD/K381>), with high GXYLT2 showing the greatest predictive power for DSS in KICH (HR = 3.450) (Figure S1B, Supplemental Digital Content, [**A**

Age: \$P = 0.13\$

Gender: \$P = 0.65\$

Grade: \$P = 0.84\$ \(G1 vs G2\), \$P = 0.39\$ \(G2 vs G3\), \$P = 0.00039\$ \(G1 vs G3\)

M: \$P = 0.95\$

T: \$P = 0.17\$ \(T1 vs T2\), \$P = 0.052\$ \(T2 vs T3\), \$P = 0.57\$ \(T3 vs T4\), \$P = 1.1e-05\$ \(T1 vs T3\), \$P = 4.8e-05\$ \(T1 vs T4\), \$P = 4.8e-05\$ \(T2 vs T4\)

B

Heatmaps for: GXYLT2, Age, Gender, Grade*, Stage*, T**, M, N

C

OS: HR = 1.67 \(1.18 - 2.36\), logrank \$P = 0.0032\$

PFS: HR = 2.83 \(1.48 - 5.4\), logrank \$P = 0.001\$

Number at risk \(OS\):

Time \(months\)	0	20	40	60	80	100	120
low	276	105	28	13	3	3	1
high	95	29	7	2	1	1	0

Number at risk \(PFS\):

Time \(months\)	0	20	40	60	80	100	120
low	157	69	17	9	2	2	1
high	58	22	7	1	1	1	0
</div>
<div data-bbox=)

Figure 8. GXYLT2 association with demographic and clinicopathological characteristics of STAD patients based on TCGA data. (A) GXYLT2 association with age, gender, grade, primary tumor size, and metastasis. (B) Heatmaps of age, gender, grade, stage, tumor size, lymph node metastasis (N), and distant metastasis (M) of patients with low and high GXYLT2 expression. (C) Kaplan-Meier survival curves of patients with low and high GXYLT2 expression. Left, OS; Right, PFS. * $P < .05$ and ** $P < .01$. GXYLT2 = Glucoside Xylosyltransferase 2, OS = overall survival, PFS = progression-free survival, TCGA = The Cancer Genome Atlas.

links.lww.com/MD/K381). In DLBC, however, the relationship between GXYLT2 and DSS was reversed ($P = .031$) (Figure S1, Supplemental Digital Content, <http://links.lww.com/MD/K381>). Moreover, high GXYLT2 was linked to low DFI in STAD ($P = .047$, Figure S2A, Supplemental Digital Content, <http://links.lww.com/MD/K382>) and low PFI in BLCA ($P = .018$), GBM ($P = .006$), KICH ($P = .005$), KIRC ($P < .001$), and STAD ($P = .018$) (Figure S3, Supplemental Digital Content, <http://links.lww.com/MD/K383>). Forest plots showing the association between GXYLT2 expression with DFI and PFI in the 33 TCGA tumor types are presented in Figures S2B and S3G, Supplemental Digital Content, <http://links.lww.com/MD/K382> and <http://links.lww.com/MD/K383>, respectively. These results suggest that GXYLT2 might serve as a prognostic marker for certain types of cancer.

3.3. Pan-cancer GXYLT2 correlation with disease stage

Of the 33 TCGA tumor types, GXYLT2 expression was significantly associated with tumor stage in adenoid cystic carcinoma (ACC) (Fig. 4A), BLCA (Fig. 4B), COAD (Fig. 4C), ESCA (Fig. 4D), HNSC (Fig. 4E), KIRC (Fig. 4F), lung

adenocarcinoma (LUAD) (Fig. 4G), STAD (Fig. 4H), and testicular germ cell tumors (TGCT) (Fig. 4I). In particular, GXYLT2 expression was higher in stage III than stage I tumors of ESCA ($P = .0035$), HNSC ($P = .028$), and STAD ($P = .013$). However, GXYLT2 expression was significantly lower in stage IV than stage I tumors of ACC ($P = .0035$), LUAD ($P = .018$), and TGCT ($P = .0039$). GXYLT2 expression was higher in stage II than stage I tumors of COAD ($P = .04$), ESCA ($P = .002$), and STAD ($P = .0071$), but no significant differences were detected between stage IV and stage I tumors of these malignancies.

3.4. Pan-cancer GXYLT2 association with TMB and MSI

Of the 33 TCGA tumor types, GXYLT2 expression was significantly linked to TMB in 17 malignancies, with a negative association detected in ACC, UCEC, THYM, THCA, STAD, PRAD, PAAD, LUSC, LUAD, LIHC, LGG, HNSC, ESCA, CHOL, and BRCA, and a positive association in SARC and KICH (Fig. 5A). GXYLT2 expression was found to be negatively associated with MSI in 7 cancer types including UCEC, THCA, STAD, LUSC, LUAD, HNSC, and CHOL (Fig. 5B).

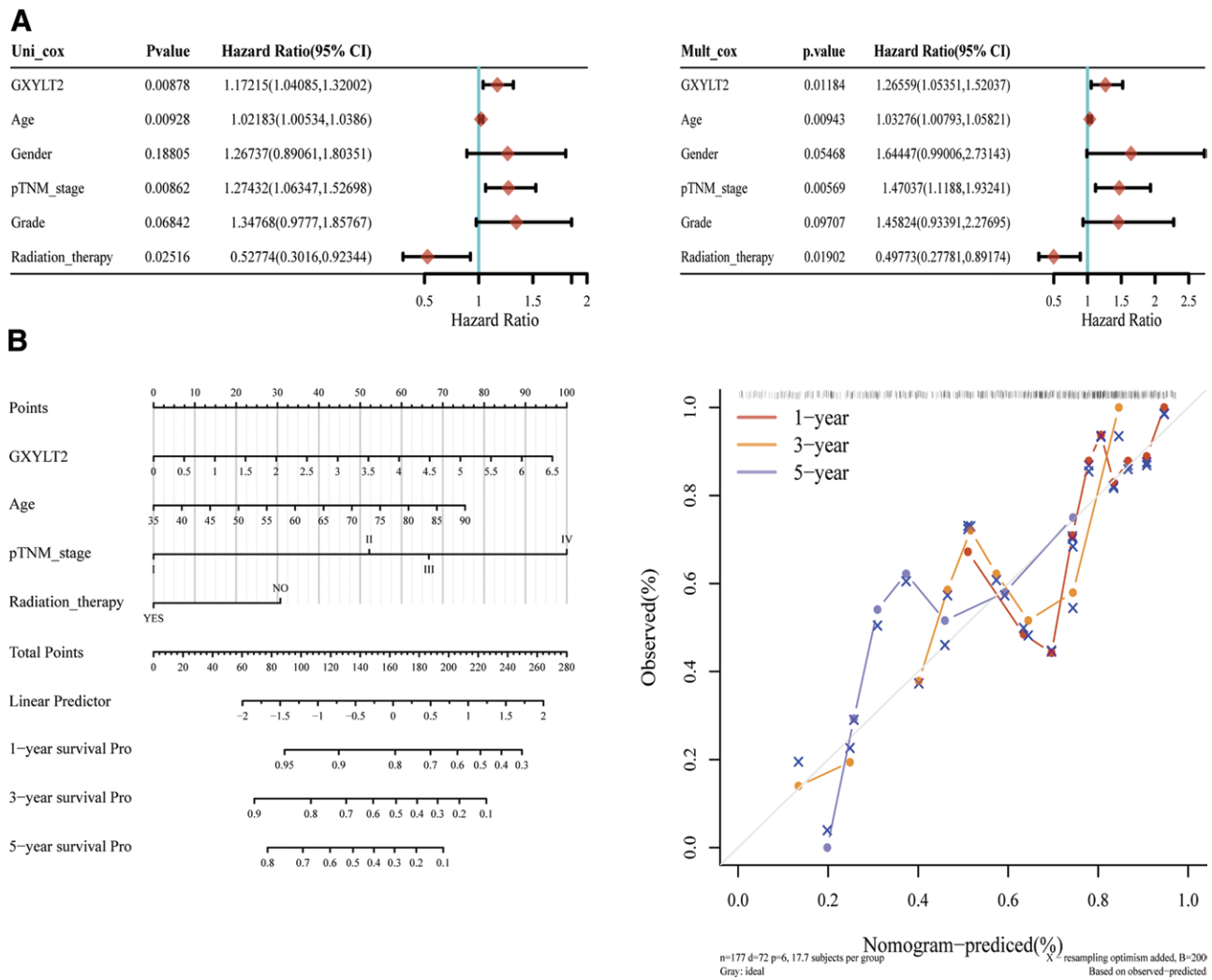


Figure 9. Prognostic nomograms for predicting the OS of STAD patients. (A) Univariate (left) and multivariate (right) Cox regression analyses of OS in a training cohort of patients. (B) Prognostic nomograms created from training data for predicting the 1-, 3-, and 5-year OS (left) and plots showing their performance with regard to their predictive accuracy in a validation cohort of patients (right). For predicting OS using a nomogram, the value for each predictor is determined by drawing a line upward to the reference point line, adding the points, and drawing a line downward from the total points line to find the predicted probability of node positivity. OS = overall survival, STAD = stomach adenocarcinoma.

3.5. Pan-cancer GXYLT2 association with tumor immune microenvironment (TIME)

In the TIME, the composition of infiltrating immune cells and their interactions in the tumor microenvironment (TME) niche play a critical role in cancer development and responses to therapies.^[22] Tumor stromal cells (e.g. cancer-associated fibroblasts) are key components of the TME and TIME that contribute to cancer formation and metastasis.^[23,24] In this study, the association between GXYLT2 with immune and stromal scores in the 33 TCGA tumor types were estimated using the ESTIMATE method. GXYLT2 expression was found to be positively linked to stromal scores in BLCA, BRCA, CHOL, COAD, DLBC, ESCA, GBM, HNSC, KIRC, kidney renal papillary cell carcinoma, LGG, LIHC, LUAD, LUSC, ovarian cancer, PAAD, PRAD, READ, SARC STAD, and UCS (Figure S4, Supplemental Digital Content, <http://links.lww.com/MD/K384>), as well as immune scores in BRCA, BLCA, COAD, CHOL, GBM, HNSC, KIRC, LIHC, LUAD, LUSC, PAAD, PRAD, and STAD (Figure S5, Supplemental Digital Content). However, GXYLT2 expression was negatively linked to immune scores in UCEC (Figure S5, Supplemental Digital Content, <http://links.lww.com/MD/K385>). Taking all 33 TCGA cancers into account, GXYLT2 was positively associated with composite stromal, immune, and ESTIMATE scores (Fig. 6). These data support a role for GXYLT2 in regulating the TME and TIME in cancer.

3.6. Pan-cancer GXYLT2 association with tumor immune cell infiltration

The TIME consists of tumor cells, different tumor-infiltrating immune cells, and cytokines, etc. To look further into the relationship between GXYLT2 and the TIME, we evaluated the relationship between GXYLT2 expression and the infiltration of 22 immune cell types in the 33 TCGA tumors using the TIMER

database. GXYLT2 expression was found to be positively correlated with the infiltration of CD8⁺ T cells (cytotoxic T lymphocytes), CD4⁺ T cells (helper T cells), neutrophils, myeloid dendritic cells, macrophages, and B cells in 18, 15, 23, 21, 24, and 11 tumor types, respectively (Figure S6A, Supplemental Digital Content, <http://links.lww.com/MD/K386>). Occasionally, a negative correlation between GXYLT2 and immune cell infiltration was detected, albeit at a much lower frequency (Figure S6A, Supplemental Digital Content, <http://links.lww.com/MD/K386>). Notably, the above results were confirmed by additional analysis using the xCELL (Figure S6B, Supplemental Digital Content, <http://links.lww.com/MD/K386>), QUANTISEQ (Figure S6C, Supplemental Digital Content), MCPOUNTER (Figure S6D, Supplemental Digital Content, <http://links.lww.com/MD/K386>), and EPIC (Figure S6E, Supplemental Digital Content, <http://links.lww.com/MD/K386>) algorithms.

3.7. Pan-cancer GXYLT2 correlation with gene expression

Across the 33 TCGA tumors, GXYLT2 expression was frequently found to be significantly correlated with the expression of RNA N6-methyladenosine demethylases, especially the fat mass and obesity-associated protein, where a positive association was detected in 25 tumors (Figure S7A, Supplemental Digital Content, <http://links.lww.com/MD/K387>). In addition, GXYLT2 expression was closely correlated with the expression of all 4 members of the DNA methyltransferase family (Figure S7B, Supplemental Digital Content, <http://links.lww.com/MD/K387>).

3.8. GXYLT2 was upregulated in STAD

Gene expression data extracted from the TCGA database revealed significant GXYLT2 upregulation in human STAD tumors compared to normal tissues (Fig. 7A), and the same

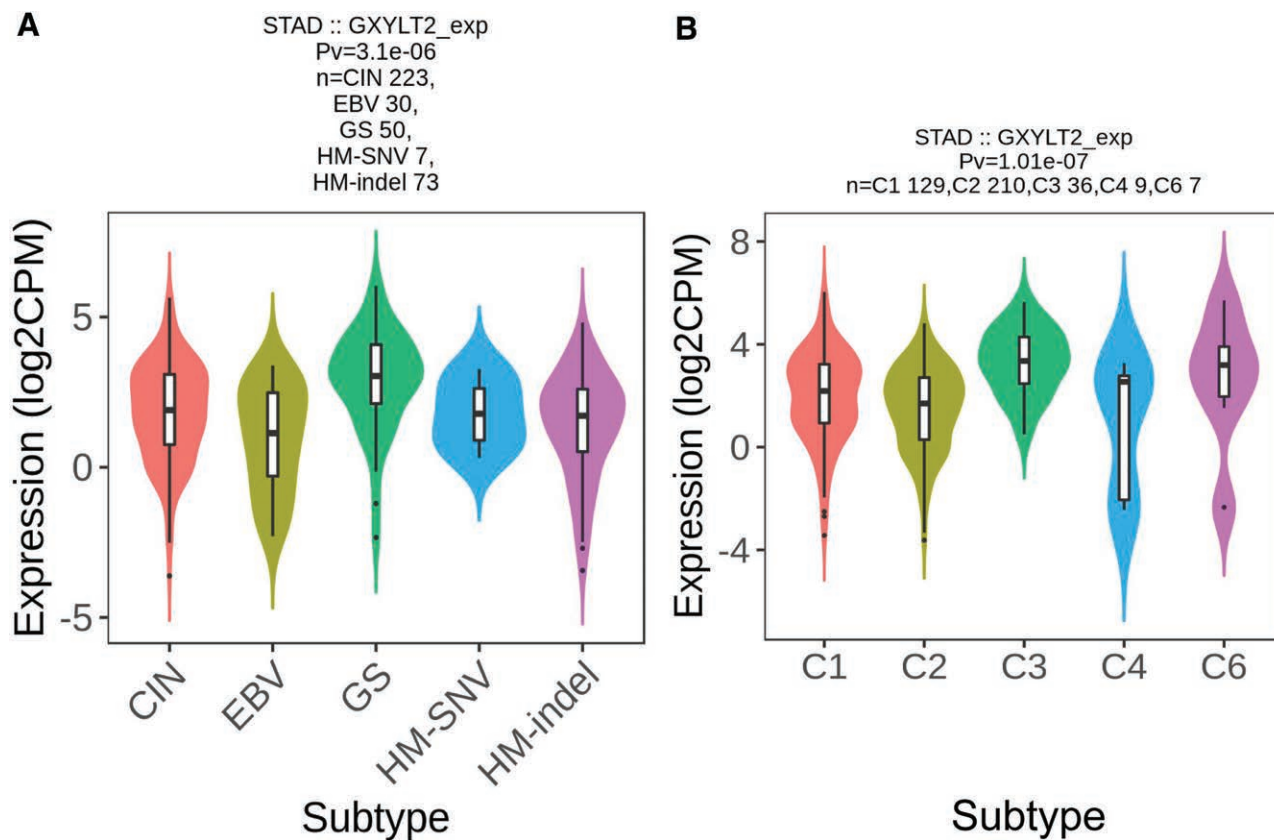


Figure 10. GXYLT2 expression was associated with STAD molecular (A) and immune (B) subtypes. GXYLT2 = Glucoside Xylosyltransferase 2, STAD = stomach adenocarcinoma.

result was obtained by combining data from the GTEx and TCGA databases (Fig. 7B). To confirm these findings, we conducted IHC experiments in our own laboratory to evaluate GXYLT2 protein expression in human STAD and normal tissues. The results confirmed GXYLT2 upregulation in human STAD (Fig. 7C).

3.9. GXYLT2 correlation with clinicopathological characteristics of STAD

Next, we evaluated the association of GXYLT2 expression with demographic and clinicopathological characteristics of STAD patients using data extracted from the TCGA. We found that GXYLT2 expression was not affected by age or gender;

however, it was significantly correlated with primary tumor size and grade (III versus II), but not metastasis (Fig. 8A). Heatmaps of demographic and clinicopathological characteristics of patients are shown in Figure 8B. Kaplan–Meier survival analysis revealed that high GXYLT2 expression was linked to low OS and progression-free survival (PFS) in STAD patients (Fig. 8C).

3.10. Prognostic nomograms for predicting OS of STAD patients

To establish prognostic models for predicting the OS of STAD patients, we first conducted univariate and multivariate OS analyses on predictive factors, including GXYLT2 expression level, age, pathological TNM stage (I, II, III, IV), and radiation

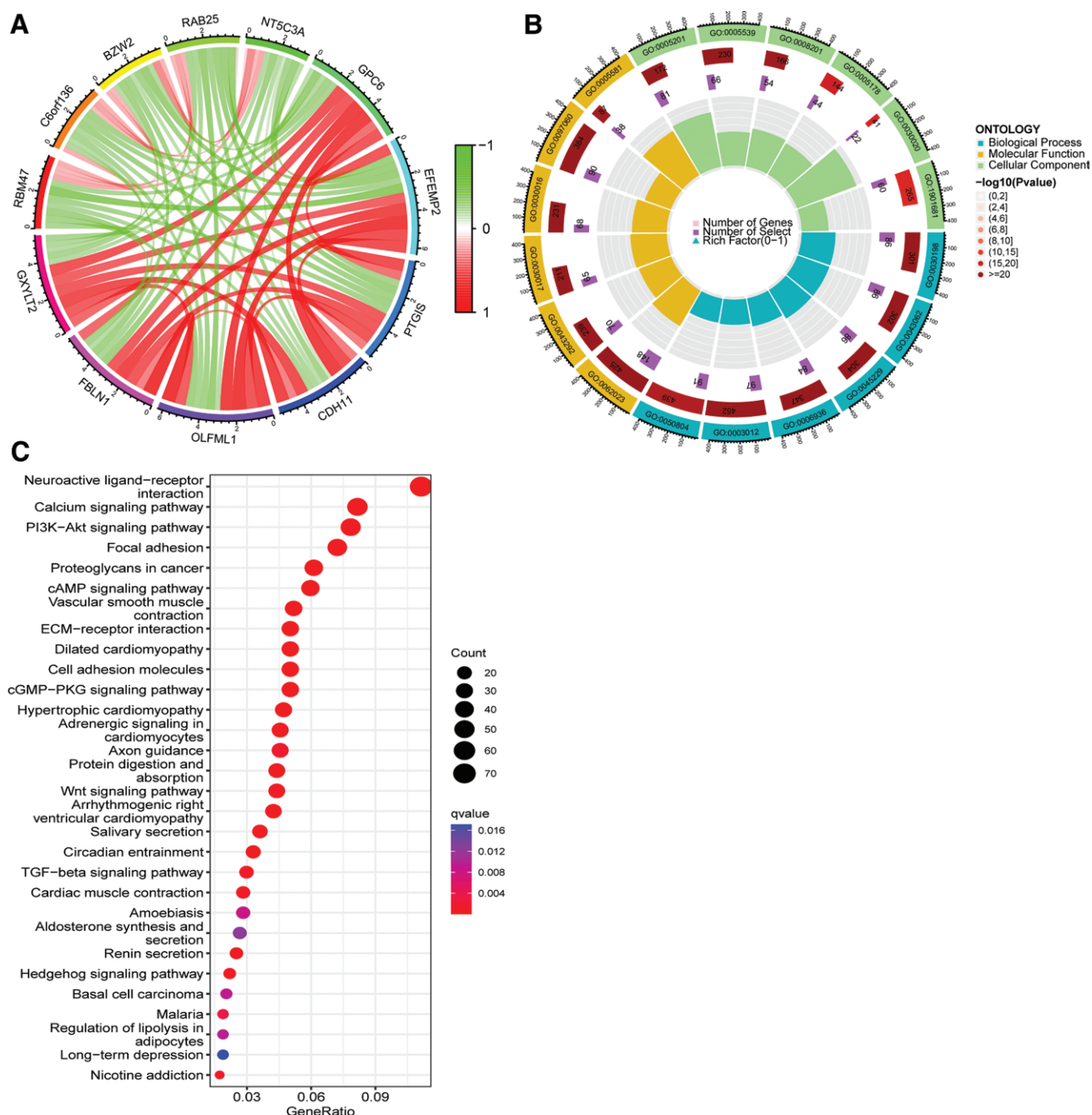


Figure 11. GXYLT2 co-expression analysis for functional classification in STAD. (A) Genes co-expressed with GXYLT2 as identified from STAD expression data extracted from TCGA. (B and C) GO and KEGG enrichment analysis of genes co-expressed with GXYLT2 in STAD. GO = Gene Ontology, GXYLT2 = Glucosyltransferase 2, KEGG = Kyoto Encyclopedia of Genes and Genomes, TCGA = The Cancer Genome Atlas.

therapy in a training cohort of patients (Fig. 9A). Based on the results from the training cohort, we created nomograms for predicting the 1-, 3-, and 5-year OS, and all 3 prognostic nomograms demonstrated good performance with regard to their predictive accuracy in a validation cohort of patients (Fig. 9B).

3.11. GXYLT2 association with immune and molecular subtypes of STAD

Next, we analyzed the association between GXYLT2 and the immune and molecular features of STAD using the TISIDB database. The results showed that GXYLT2 was significantly correlated with STAD molecular subtypes (Fig. 10A). Across TCGA cancer types, 6 immune subtypes were defined: C1, wound healing; C2, IFN- γ dominant; C3, inflammatory; C4, lymphocyte depleted; C5, immunologically quiet; C6, TGF- β dominant.^[25] GXYLT2 expression was found to be associated with STAD immune subtypes (Fig. 10B).

3.12. GO and KEGG analysis of genes co-expressed with GXYLT2 in STAD

To probe the possible function of GXYLT2 in STAD, we performed a co-expression analysis based on STAD RNA sequencing data from the TCGA database (Fig. 11A). Co-expression networks were constructed using GO and KEGG enrichment analysis (Fig. 11B and C). The results showed that genes co-expressed with GXYLT2 were mainly involved in neuroactive ligand-receptor and extracellular matrix (ECM)-receptor interactions, as well as the calcium, PI3K-AKT-mTOR, cAMP, cGMP-PKG, Wnt, and TGF- β signaling pathways. In terms of functions related to tumor growth and tumor immunity, GXYLT2 was associated with genes involved in epithelial-mesenchymal transition, ECM production and degradation, angiogenesis, and apoptosis, as well as tumor inflammation. Notably, GXYLT2 demonstrated a negative association with genes involved in cellular responses to hypoxia or reactive oxygen species, DNA replication and repair, G2M checkpoint, and the P53 and MYC pathways (Figure S8, Supplemental Digital Content, <http://links.lww.com/MD/K388>).

4. Discussion

Although GXYLT2 has been implicated in tumor cell proliferation and migration via upregulation of Notch,^[6] its prognostic significance and oncogenic function in most types of cancers remain largely unknown. In the present study, analysis of GXYLT2 expression across 33 TCGA tumor types using data from public cancer genome databases revealed significant GXYLT2 upregulation in a number of major cancer types. We found that GXYLT2 expression was correlated with disease progression, mortality, TMB, MSI, and tumor immune cell infiltration in certain cancer types including STAD. This pan-cancer analysis suggested that GXYLT2 may play a role in tumorigenesis and tumor immunity, and may serve as a prognostic marker and potential immunotherapeutic target for STAD and some other cancer types.

Previous studies have shown that GXYLT2 regulates Notch,^[1,2] a receptor in a highly conserved signaling pathway that can play either an oncogenic or anti-tumor role in cancer cells.^[3] While most research has focused on the role of Notch in cancer cell biology, recent studies have indicated that Notch also regulates the crosstalk between the different components of the TME through paracrine and juxtacrine signaling.^[26] In this study, we found that GXYLT2 was upregulated in most tumor tissues including STAD, but downregulated in some other types of cancer, suggesting that, similar to Notch, GXYLT2 could function as either a tumor promoter or suppressor, depending on the specific type of cancer. STAD is a leading cause of cancer-related

deaths worldwide.^[27] Most STAD patients are diagnosed at a late stage, thus, finding new biomarkers for early diagnosis is critical.^[28] In this study, sequencing data from public databases shows that GXYLT2 was overexpressed in human STAD, which was confirmed by our IHC experiments. We also found that high GXYLT2 expression was strongly correlated with low OS, DSS, and PFS, as well as short DFI in STAD patients. In addition, GXYLT2 expression in late stage STAD tumors was higher than that in early stage tumors. These results strongly support the use of GXYLT2 as a diagnostic/prognostic marker for STAD.

TMB and MSI have emerged as genomic biomarkers for identifying cancer patients who are likely to benefit from the use of immune checkpoint inhibitors.^[29–32] Across the 33 TCGA tumor types, GXYLT2 expression showed a mostly negative association with TMB and MSI. However, positive correlations with immune scores were detected in more than 10 tumors, including STAD. Zhao et al^[4] reported that GXYLT2 expression in STAD was positively associated with the infiltration of M2 macrophages, monocytes, and active NK cells, but negatively correlated with the infiltration of activated CD4⁺ T cells and neutrophils. Taken together, these results implicate GXYLT2 as an important regulator of tumor immunity in STAD and other types of cancer.

GO and KEGG pathway analyses of genes co-expressed with GXYLT2 in STAD revealed a complex interactive network of signaling pathways involved in processes related to tumor growth and tumor immunity, such as epithelial-mesenchymal transition, ECM production and degradation, angiogenesis, apoptosis, and tumor inflammation. Finally, we observed that patients with increased GXYLT2 expression were more likely to respond favorably to immunotherapy (data not shown). These findings support that GXYLT2 may regulate both cancer cells and other cell components (e.g. immune cells) of the TME and TIME.

This study has a few limitations. Firstly, the data used in this study were almost exclusively extracted from public databases, which may or may not have had proper controls in terms of patient age, gender, and other factors that may have affected the reported results. Secondly, bioinformatic findings from this study were not confirmed by either *in vitro* or *in vivo* experiments.

This study was the first pan-cancer analysis of the prognostic significance and oncogenic role of GXYLT2. The results showed that GXYLT2 expression was correlated with disease stage, patient prognosis, and tumor immune infiltration in a number of major cancers including STAD. Co-expression analysis suggested that GXYLT2 may play a role in tumorigenesis and tumor immunity. Thus, GXYLT2 may serve as a prognostic marker and potential immunotherapeutic target for STAD and some other types of cancer.

Acknowledgments

The authors would like to thank the donors of STAD and normal tissue specimens that were used in the IHC experiments.

Author contributions

Formal analysis: Yi-Bei Song, Wen-Guang Bao, Deng-He Liu, Li-Qiang Wei, Dong Lan, Hui-Min He.

Investigation: Yi-Bei Song, Wen-Guang Bao, Shu-Ting Yang, Chun-Yu Lin, Hong-Jun Li.

Methodology: Xue-Jing Miao, Chun-Yu Lin.

Resources: Shu-Ting Yang.

Writing – original draft: Yi-Bei Song, Wen-Guang Bao.

References

- [1] Sethi MK, Buettner FF, Krylov VB, et al. Identification of glycosyltransferase 8 family members as xylosyltransferases acting on O-glycosylated notch epidermal growth factor repeats. *J Biol Chem.* 2010;285:1582–6.
- [2] Lee TV, Sethi MK, Leonardi J, et al. Negative regulation of notch signaling by xylose. *PLoS Genet.* 2013;9:e1003547.

- [3] Aster JC, Pear WS, Blacklow SC. The varied roles of notch in cancer. *Annu Rev Pathol.* 2017;12:245–75.
- [4] Zhao Y, Hu S, Zhang J, et al. Glucoside xylosyltransferase 2 as a diagnostic and prognostic marker in gastric cancer via comprehensive analysis. *Bioengineered.* 2021;12:5641–54.
- [5] Wu S, Qiu S, Chen W, et al. Prognostic signature *gxytl2* is correlated with immune infiltration in bladder cancer. *Dis Markers.* 2022;2022:5081413.
- [6] Cui Q, Xing J, Gu Y, et al. *GXYLT2* accelerates cell growth and migration by regulating the Notch pathway in human cancer cells. *Exp Cell Res.* 2019;376:1–10.
- [7] Weinstein JN, Collisson EA, Mills GB, et al.; Cancer Genome Atlas Research Network. The cancer genome atlas pan-cancer analysis project. *Nat Genet.* 2013;45:1113–20.
- [8] Li Y, Roberts ND, Wala JA, et al.; PCAWG Structural Variation Working Group. Patterns of somatic structural variation in human cancer genomes. *Nature.* 2020;578:112–21.
- [9] Cortés-Ciriano I, Lee JJ, Xi R, et al.; PCAWG Structural Variation Working Group. Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. *Nat Genet.* 2020;52:331–41.
- [10] Cheng X, Wang X, Nie K, et al. Systematic pan-cancer analysis identifies *trem2* as an immunological and prognostic biomarker. *Front Immunol.* 2021;12:646523.
- [11] Ju M, Bi J, Wei Q, et al. Pan-cancer analysis of *NLRP3* inflammasome with potential implications in prognosis and immunotherapy in human cancer. *Brief Bioinform.* 2021;22:bbaa345.
- [12] Rhodes DR, Yu J, Shanker K, et al. ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia.* 2004;6:1–6.
- [13] Tomczak K, Czerwińska P, Wiznerowicz M. The cancer genome atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol.* 2015;19:A68–77.
- [14] GTEx Consortium. The genotype-tissue expression (GTEx) project. *Nat Genet.* 2013;45:580–5.
- [15] Li T, Fan J, Wang B, et al. TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res.* 2017;77:e108–10.
- [16] Chen B, Khodadoust MS, Liu CL, et al. Profiling tumor infiltrating immune cells with *cibersort*. *Methods Mol Biol.* 2018;1711:243–59.
- [17] Aran D, Hu Z, Butte AJ. *xCell*: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* 2017;18:220.
- [18] Finotello F, Mayer C, Plattner C, et al. Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. *Genome Med.* 2019;11:34.
- [19] Racle J, de Jonge K, Baumgaertner P, et al. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *eLife.* 2017;6:e26476.
- [20] Becht E, Giraldo NA, Lacroix L, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* 2016;17:218.
- [21] Wang C, Li X, Ren L, et al. Gankyrin as potential biomarker for colorectal cancer with occult liver metastases. *Front Oncol.* 2021;11:656852.
- [22] Binnewies M, Roberts EW, Kersten K, et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat Med.* 2018;24:541–50.
- [23] Bussard KM, Mutkus L, Stumpf K, et al. Tumor-associated stromal cells as key contributors to the tumor microenvironment. *Breast Cancer Res.* 2006;18:1–11.
- [24] Desbois M, Wang Y. Cancer-associated fibroblasts: key players in shaping the tumor immune microenvironment. *Immunol Rev.* 2021;302:241–58.
- [25] Thorsson V, Gibbs DL, Brown SD, et al.; Cancer Genome Atlas Research Network. The immune landscape of cancer. *Immunity.* 2018;48:812–830.e14.
- [26] Nowell CS, Radtke F. Notch as a tumour suppressor. *Nat Rev Cancer.* 2017;17:145–59.
- [27] Ferlay J, Colombet M, Soerjomataram I, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer.* 2019;144:1941–53.
- [28] Norwood DA, Montalvan EE, Dominguez RL, et al. Gastric cancer: emerging trends in prevention, diagnosis, and treatment. *Gastroenterol Clin North Am.* 2022;51:501–18.
- [29] McNamara MG, Jacobs T, Lamarca A, et al. Impact of high tumor mutational burden in solid tumors and challenges for biomarker application. *Cancer Treat Rev.* 2020;89:102084.
- [30] Yarchoan M, Hopkins A, Jaffee EM. Tumor mutational burden and response rate to PD-1 inhibition. *N Engl J Med.* 2017;377:2500–1.
- [31] Picard E, Verschoor CP, Ma GW, et al. Relationships between immune landscapes, genetic subtypes and responses to immunotherapy in colorectal cancer. *Front Immunol.* 2020;11:369.
- [32] Palmeri M, Mehnert J, Silk AW, et al. Real-world application of tumor mutational burden-high (TMB-high) and microsatellite instability (MSI) confirms their utility as immunotherapy biomarkers. *ESMO Open.* 2022;7:100336.