Qiao et al. BMC Urology (2024) 24:222
https://doi.org/10.1186/s12894-024-01613-6

BMC Urology

RESEARCH Open Access

Identification and experimental validation of a sialylation–related long noncoding RNA signature for prognosis of bladder cancer

Yi Qiao[1], Xintao Tian[1], Shengxian Li[1*] and Haitao Niu[1*]

Abstract

Background The dysregulation of sialylation plays a pivotal role in cancer progression and metastasis, impacting various aspects of tumor behavior. This study aimed to investigate the prognostic significance of long non-coding RNAs (lncRNAs) in relation to sialylation. Additionally, we aimed to develop a signature of sialylation-related lncRNAs in the context of bladder cancer.

Methods This study used transcriptomic data and clinical information from the TCGA (the Cancer Genome Atlas) database to screen for sialylation-related lncRNAs and constructed a prognostic model. The relationships between these lncRNAs and biological pathways, immune cell infiltration, drug sensitivity, etc., were analyzed, and the expression of some lncRNAs was validated at the cellular level.

Results This study identified 6 prognostic lncRNAs related to sialylation and constructed a risk score model with high predictive accuracy and reliability. The survival period of patients in the high-risk group was significantly lower than that of the low-risk group, and it was related to various biological pathways and immune functions. In addition, this study found differences in the sensitivity of patients in different risk groups to chemotherapy drugs, providing a reference for personalized treatment.

Conclusion In this study, we examined the relationship between sialylation-related lncRNA and the prognosis of bladder cancer, providing new molecular markers and potential targets for diagnosis and treatment. Our research revealed correlations between sialylation-related lncRNA characteristics and clinicopathological features, potential mechanisms, somatic mutations, immune microenvironment, chemotherapy response, and predicted drug sensitivity in bladder cancer. Additionally, in vitro cellular studies were conducted to validate these findings and lay the groundwork for future clinical applications.

Keywords Bladder cancer, LncRNAs, Tumor microenvironment, Nomogram, Immunotherapy

*Correspondence:
Shengxian Li
lishengxian@qdu.edu.cn
Haitao Niu
niuht0532@126.com
[1]Department of Urology, The Affiliated Hospital of Qingdao University, Qingdao, China

Introduction

Bladder cancer is a highly widespread kind of malignancy that impacts the urinary tract [1]. Around 70–80% of patients are initially diagnosed with non-invasive bladder cancer (NMIBC) [2]. Despite notable advancements in the prognosis of NMIBC, a considerable proportion of patients, over 60%, nevertheless encounter recurrences [3, 4]. Moreover, more than 20% of these individuals eventually develop muscle-invasive bladder cancer (MIBC), a condition characterized by elevated mortality rates and a propensity for metastasis. This underscores the pressing necessity for timely and accurate categorization of patients in order to customize therapies. Immunotherapeutic approaches, including those involving immune checkpoint inhibitors (ICIs), exhibit considerable potential in enhancing patient care [5]. Hence, it is imperative to test the molecular pathways that underlie tumor progression in clinical contexts. This validation will facilitate the advancement of targeted therapy approaches aimed at enhancing the prognosis of bladder cancer.

Sialylation refers to the process of adding sialic acid units to oligosaccharides and glycoproteins at their terminal ends [6]. This modification plays a significant biological role in various processes such as embryonic development, neurodevelopment, reprogramming, and the manifestation of certain pathological conditions, including cancer, embryonic lethality, and anomalies in the immune system [7]. There is an increasing body of evidence that indicates the prevalent presence of abnormal sialylation in human malignancies [6]. This abnormal sialylation has been found to play a role in various aspects of cancer development, including oncogenesis, tumor cell dissociation, invasion, immune evasion, and resistance to therapy [7]. In addition, sialyltransferases(STs) have been identified as promising targets for the development of anticancer treatments [8]. Sialic acid-binding immunoglobulin-like lectins (Siglecs) are widely distributed on various tumor-infiltrating cells, including T cells, neutrophils, and natural killer (NK) cells, making them potential immune checkpoint targets with anti-tumor therapeutic potential [9–11]. For example, Siglec-6 expression is increased on circulating and urinary T cells in patients with non-muscle invasive bladder cancer and is associated with lower survival rates [12]. Additionally, Siglec-7 may regulate NK cell-mediated anti-tumor immunity in bladder cancer [13]. Sialyltransferases are crucial in the sialylation process, with significant research focusing on their role in cancer. For instance, ST3Gal.I is pivotal in the sialylation of the T antigen in bladder cancer [14]. Additionally, ST3Gal5 is linked to muscle invasion and poor prognosis in bladder cancer patients, where low levels of ST3Gal5 promote the occurrence and progression of the disease [15, 16]. Therefore, biomarkers related to sialylation molecules may be valuable for predicting the prognosis of bladder cancer patients.

Long non-coding RNAs (lncRNAs) are transcripts that exceed 200 nucleotides in length, playing significant roles in cancer development and progression through epigenetic modifications or translational regulation [17, 18]. The inherent stability of these biomarkers in the bloodstream, along with their resistance to destruction by nucleases, makes them highly promising for cancer detection and monitoring, particularly in comparison to other biomarkers such as circulating tumor cells (CTCs), cell-free DNA (cfDNA), circulating tumor DNA (ctDNA), and exosomes [19]. Nevertheless, there is a scarcity of research on sialylation-associated lncRNAs in cancer, despite the extensive use of mass spectrometry in studying abnormal sialylation. With the substantial number of unidentified lncRNAs, it is conceivable that numerous lncRNAs participate in sialylation and Siglec interactions, and their clinical relevance in bladder cancer warrants further exploration. Therefore, there is an urgent need to identify lncRNA biomarkers associated with sialylation for predicting prognosis and treatment response in bladder cancer patients.

Existing researches lack studies on the relationship between sialylation-related lncRNA and bladder cancer prognosis. Therefore, this study aimed to identify prognostic lncRNAs that are related with sialylation and Siglec functional pathways in bladder cancer. Besides, based on these sialylation-related prognostic lncRNAs, we constructed a risk scoring model with high predictive accuracy and reliability. Based on these sialylation-related prognostic lncRNAs, we constructed a risk scoring model with high predictive accuracy and reliability. Additionally, the study aimed to establish signatures of sialylation-associated lncRNAs. In order to accomplish this objective, a thorough examination was carried out encompassing the correlation between sialylation-associated lncRNA signatures, clinicopathological characteristics, underlying mechanisms, somatic mutations, the immune microenvironment, chemotherapeutic responses in bladder cancer, and anticipated drug sensitivity. Furthermore, we conducted in vitro cellular investigations to substantiate academic theories and lay the groundwork for prospective clinical applications.

Materials and methods

Data and tissue processing

The transcriptome data and clinical information for Bladder Urothelial Carcinoma (BLCA) were obtained from the Cancer Genome Atlas(TCGA, https://www.cancer.gov/tcga)database. This dataset included transcriptomic data for 413 bladder cancers and clinical information for the corresponding patients [20]. All data are in FPKM format, and a total of 413 samples were included

in the cohort. To prepare the data for further analysis, we annotated it using gene transfer format files obtained from Ensembl(https://asia.ensembl.org) and filtered for lncRNAs [21].

Obtaining sialylation-related long non-coding RNAs and developing a prognostic model

The identification of genes involved in the biological process of sialylation was accomplished by consulting the Molecular Signatures Database (MSigDB), which include STs, transporters, and neuraminidases. The association between lncRNAs and genes relevant to sialylation was assessed using the Pearson correlation test. A total of 2511 lncRNAs associated with sialylation were discovered based on a correlation coefficient ($|r| > 0.3$) and a significance level ($P<0.001$). The cohort of patients was randomly allocated into two groups, namely the training group and the validation group, with an equal distribution ratio of 50:50. The examination of lncRNAs associated with sialyation was conducted using one-way regression analysis. Only those sialyation-related lncRNAs that showed a significance level of $P<0.001$ in the univariate cox regression analysis were considered for inclusion in the lasso regression. Following this, lncRNAs associated with sialylation were discovered using lasso regression and subsequently incorporated into a multifactorial Cox regression model in order to provide risk ratings. We have successfully identified eight lncRNAs associated with prognosis-related sialylation. These lncRNAs will be utilized to develop a predictive risk score. We finally identified 8 prognosis-related sialyation-related lncRNAs to construct a prognostic risk score. Risk score for BLCA patients was calculated by the following formula: risk score=$\sum_{i=1}^{n} \beta_i \times (expression\ of\ LncRNA_i)$. The ROC curves were plotted using Kaplan-Meier method in "survival" ROC R package. Patients in the cohort were divided into high-risk and low-risk groups based on the median risk score.

Assessing the value of prognostic sialyation-related lncRNAs

We conducted Kaplan-Meier survival analysis for overall survival (OS) to compare the high-risk and low-risk groups, using the "survminer" and "survival" R packages. In order to examine the distribution of clinical factors among patients with BLCA in different groups, such as age, sex, grading, staging, and TMN staging, we utilized the chi-square test with the "ComplexHeatmap" R package. Furthermore, the Wilcoxon test was employed in the "limma" [22] and "ggpubr" R packages to evaluate the extent of variation in risk scores across patients with distinct genders, grades, tumor stages, or node stages. Finally, we assessed the feasibility of our model by performing univariate and multivariate Cox regression analyses on clinical parameters such as age, gender, stage, and risk score to predict the survival of BLCA patients, using the "survival" R package.

Functional enrichment analysis

We identified differentially expressed genes (DEGs) for mRNAs between the two risk groups using the "limma" package in R [22]. We applied a threshold of an absolute log2-fold change (FC)>1 and a corrected threshold $P<0.05$ for DEGs selection. Following this, we performed an examination of these DEGs utilizing the "clusterprofiler" R package. This study encompassed Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and Gene Ontology (GO) analyses, along with Gene Set Enrichment study (GSEA) [23].

Construction and validation of a nomogram

The correlation between clinical parameters and risk scores was evaluated using univariate and multivariate Cox regression analyses. A nomogram was constructed incorporating age, clinical grade, risk score, lymph node involvement, and pathological grade. We developed OS line plots using the R package "rms" to predict the probability of recurrence-free survival at 1, 3, and 5 years. In order to evaluate the precision of our model, we employed calibration curves to visually depict the disparity between our projected values and the observed results. Additionally, we conducted ROC analysis to evaluate the predictive power of the risk score. Principal Component Analysis (PCA) was conducted using the "scatterplot3D" R package to assess the distribution of patients across different groups.

Analysis of immune cell infiltration

In order to evaluate the impact of the signature on immune cells, we utilized CIBERSORT (Cell-type Identification by Estimating Relative Subsets of RNA Transcripts; https://cibersort.stanford.edu) to measure the abundance of tumor-infiltrating immune cells through linear support vector regression [24]. The abundance of lymphocyte profiles in the TCGA-BLCA dataset was determined using the "CIBERSORT" R package. Furthermore, we conducted an analysis of immune function, immune checkpoint differences, immune infiltration, and immune checkpoint correlations using R packages such as "limma" [22] ,"tidyverse" [25] and "GSVA" [26] .

Tumor mutational burden and drug sensitivity prediction

Tumor mutational burden (TMB) was visualized using the 'ggpubr' package. Survival curves were visualized using the "survival" and "survminer" packages. Then, We obtained drug sensitivity data from the Genomics of Drug Sensitivity in Cancer (GDSC) database (https://

www.cancerrxgene.org/) [27]. To analyze drug sensitivity, we utilized the R package "Oncopredict." [28].

Cell culture

We obtained two cell lines, SV-HUC-1 and T24, from the Cell Bank of the Chinese Academy of Sciences (Shanghai, China). These cells were cultured at 37 °C and 5% CO2 using DMEM medium supplemented with 10% FBS (Gibco, Grand Island, NY, USA) and 1% penicillin/streptomycin.

Validation of hub genes by quantitative real-time PCR

Total RNA was extracted with TRIzol (Takara), and cDNA for each group was synthesized using the PrimeScript RT kit (Takara). Subsequently, qPCR was conducted using a Roche LightCycler 480 II real-time PCR detection system (Roche, Basel, Switzerland). The primer sequences for the target genes are as follows: GRASLND-F/R (AGGATTCAGGGGATGCACAG/TGGGCTGAAGATGAGACGTT), ARHGAP5-AS1-F/R (CTCAAGAGCAAACCACCGTAC/ACATGTTCCTGCGAACGA), LINC01508-F/R (CTCTCTCGAACTTGAACTTGCC/GCTCCTCTTTCCTGGTGTCTC). To ensure accurate measurements, the expression levels of each gene were normalized to the expression level of GAPDH.

Results

Recognition of Prognostic sialylation-related LncRNAs and model development

A total of 120 messenger RNAs (mRNAs) related to the interaction between sialic acid and siglec were initially acquired from the MsigDB website. A total of 2511 potential lncRNAs associated with sialylation were found using the Pearson correlation test, employing a threshold of $|r| > 0.3$ and $P<0.001$.

In order to assess the prognostic relevance of the potential lncRNAs, we utilized 413 lncRNAs from TCGA with complete clinical data. Following a random allocation of the 413 patients into distinct training and validation groups, we conducted univariate Cox regression analysis, with a p-value threshold of 0.001. The present investigation resulted in the identification of 14 lncRNAs that are related with sialylation and exhibit prognostic importance. The further screening of these variables was conducted using lasso regression in order to proceed with the subsequent stage, which entailed multifactorial Cox regression analysis. (Fig. 1A, B).

Subsequently, a comprehensive investigation was conducted to identify and construct a model for a set of six lncRNAs that are connected with sialylation. The prognostic significance of these lncRNAs was determined using multifactorial Cox regression analysis. Several genes were identified as negative prognostic factors, including *LINC01508*, *AC012625.1*, *AL135905.1*, *GRASLND*, and *ARHGAP5-AS1*. Conversely, *AC104785.1* was determined to be a favorable prognosis factor (Fig. 2A).

These 6 sialylation-related LncRNAs were used to create a TIL-related LncRNA signature with the following risk score formula: Risk score = (0.097657 × expression level of *LINC01508*) + (-0.3505 × expression level of *AC104785.1*) + (0.230 × expression level of *AC012625.1*)

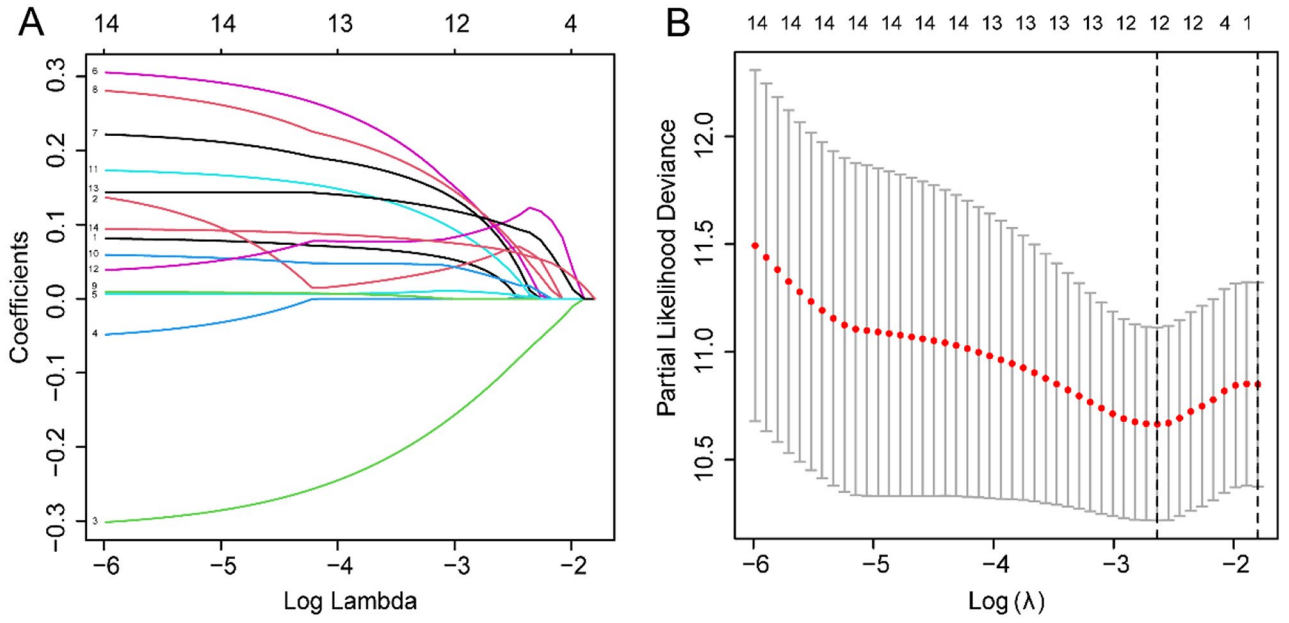

Fig. 1 Identifying 6 sialylation-related Long noncoding RNAs (lncRNAs)s. (**A**) Profiles of Lasso coefficients. (**B**) Cross-validation for tuning parameter selection

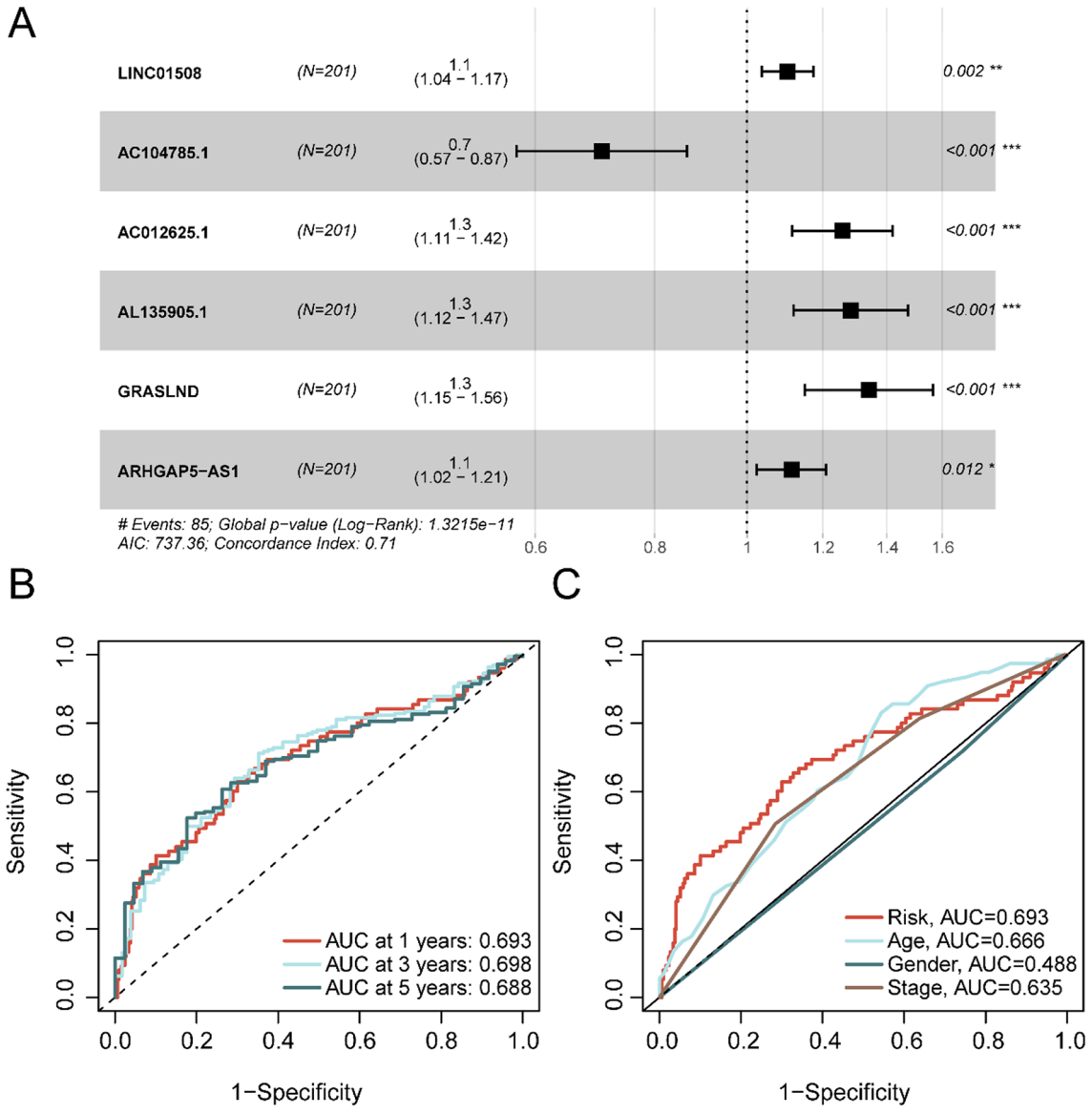

Fig. 2 Recognition of Prognostic sialylation-related lncRNAs and Model Development. (**A**) Multivariate Cox model of 6 sialylation-related lncRNAs. (**B**) ROC curves for 1, 3, and 5 years. (**C**) The 1-year ROC curves of the risk score and other clinical characteristics

+ (0.2500 × expression level of *AL135905.1*) + (0.2936 × expression level of *GRASLND*) + (0.1070 × expression level of *ARHGAP5-AS1*).

Subsequently, we evaluated the predicted survival values of the model by plotting ROC curves for 1, 3, and 5 years. The area under the curve (AUC) values were 0.693, 0.698, and 0.688 (Fig. 2B). The AUC values for the risk model are not sufficiently high at 1, 3, and 5 years. This could be due to factors such as sample size, feature variables, and the chosen machine learning algorithms. Notably, the AUC values of the 1-year curves were significantly higher than other clinical parameters, demonstrating that the model developed using the risk score reliably predicted survival (Fig. 2C).

Value assessment of prognostic sialylation-related LncRNAs

Patients were categorized as either high-risk or low-risk based on the median of the calculated risk scores in each group. In both the training group (Fig. 3, left), the test group (Fig. 3, middle), and the combined group (Fig. 3, right), the high-risk group exhibited higher mortality rates, indicating a less favorable prognosis for patients with high-risk scores ($p<0.001$). The risk scores and survival status of the training, test, and combination groups are displayed in Fig. 3(bottom).

The results shown in Fig. 4A highlight a significant difference between stage II and stage III ($p=0.00094$) and between stage II and stage IV ($p=0.00037$). Furthermore,

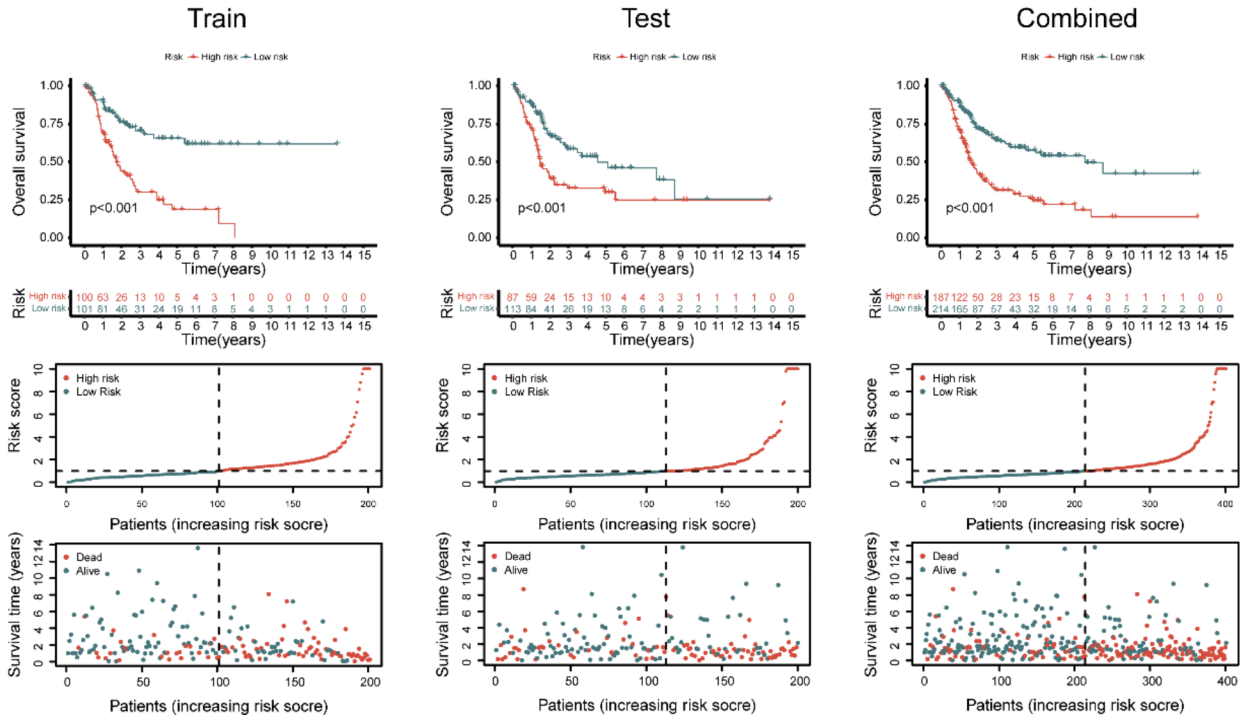

Fig. 3 Distribution of risk score (high or low) and status (dead or alive) and KM curves of OS in the training (left), testing (medium), and combined (right) sets

Fig. 4B reveals that risk scores were lower in stage N0 compared to stage N2 ($p=0.028$). Additionally, Fig. 4C demonstrates that T2 had significantly lower scores than T3 ($p=0.0018$) and also lower than T4 ($p=0.019$). Additionally, the prognostic value of the risk score was evaluated in patients from different stratified cohorts based on age, gender, and tumor stage. For this evaluation, we only included subgroups with 50 or more patients after stratification. The results indicated that the risk score effectively predicted prognosis in various cohorts, including age≤65, age>65, female, male, stage I/II, and stage III/IV ($p<0.001$)(Fig. 4D).

To explore the relationship between clinicopathologic characteristics and risk scores, we utilized the chi-squared (c2) test and Cox regression in the combined group. The c2 test indicated differences in grade, clinical stage, and T stage between the high-risk and low-risk groups (Fig. 5A). Univariate Cox regression analysis revealed significant associations of stage (HR 1.643, 95% CI: 1.255–2.152, $p<0.001$) and risk score (HR 1.084, 95% CI: 1.055–1.114, $p<0.001$) with the outcome (Fig. 5B). Furthermore, multivariate Cox regression analysis confirmed a similar association between stage (HR 1.637, 95% CI: 1.241–2.160, $p<0.001$) and risk score (HR 1.087, 95% CI: 1.059–1.116, $p<0.001$) and the outcome, underscoring the significant connection between risk scores and overall survival (Fig. 5C).

Potential mechanism analysis of the sialylation-related LncRNA signature

We conducted KEGG pathway, GSEA, and GO analyses to explore the underlying mechanisms by which the risk signature stratifies patient prognosis. Initially, we employed pathway analysis and GSEA to uncover the biological significance of the sialylation-related lncRNA signature. Through the use of the R package "edgeR," we identified a total of 612 DEGs with $|\log2(FC)| > 1$ and adjusted $P<0.05$ between the two risk groups in the combined data.

According to the GSEA analysis, the high-risk group exhibited enrichment in the top five pathways. These pathways were identified as cytokine-cytokine receptor interaction, extracellular matrix (ECM)-receptor interaction, focal adhesion, neuroactive ligand-receptor interaction, and regulation of actin cytoskeleton (Fig. 6A). Conversely, the top five enriched pathways in the low-risk group encompassed alpha-linolenic acid metabolism, drug metabolism - cytochrome P450, linoleic acid metabolism, primary immunodeficiency, and ribosome (Fig. 6B). In addition, the GO enrichment analysis conducted on the 612 DEGs indicated their participation in many biological processes, including epidermis formation, epidermal cell differentiation, and granulocyte chemotaxis (Fig. 6C).

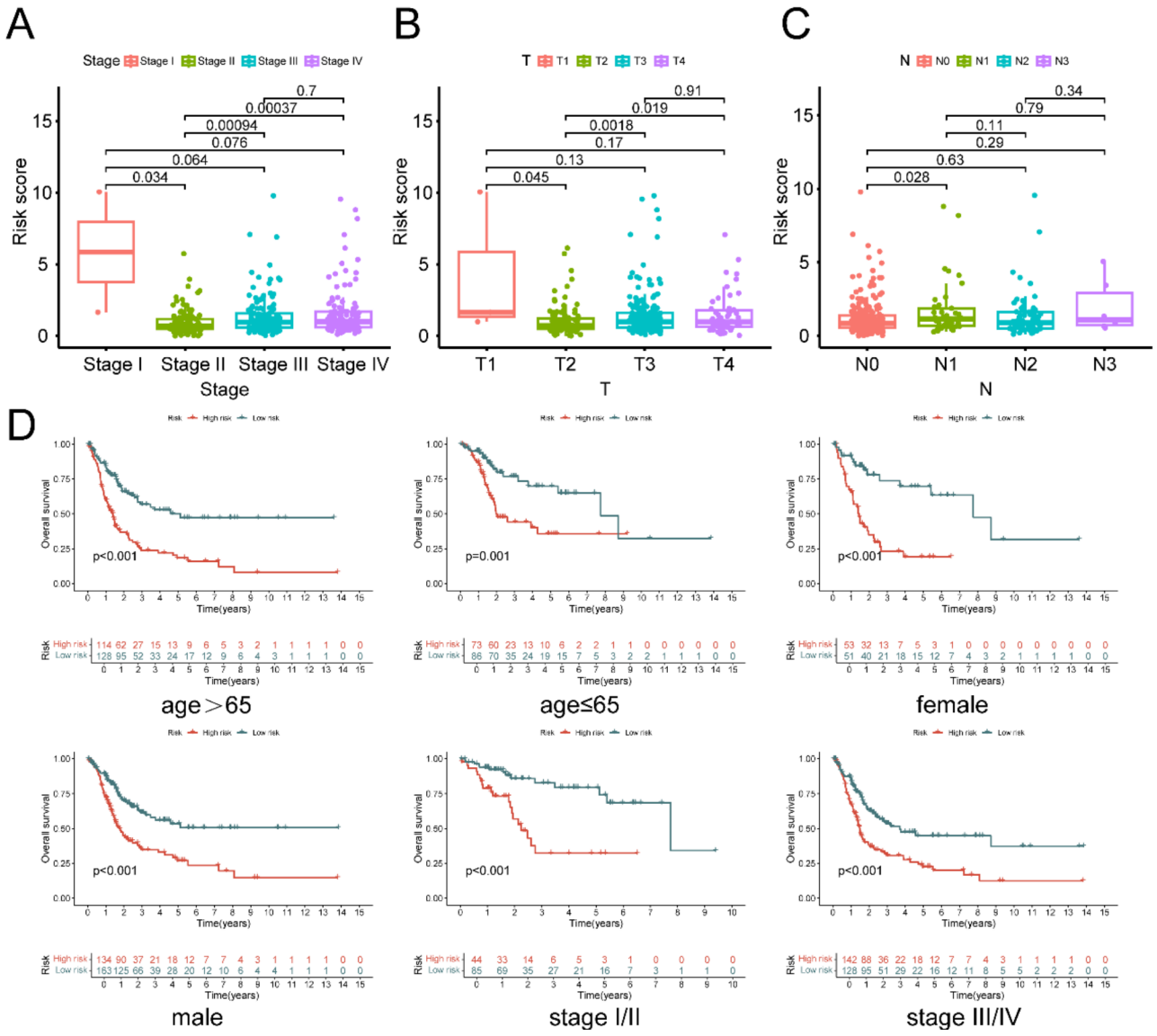

Fig. 4 Significant differences between between risk scores and clinical characteristics in TCGA-BLCA cohort. Differences between risk score and tumor stage (**A**), T stage (**B**), N stage (**C**). (**D**) Survival analysis of the sialylation-related lncRNA signature in different stratified cohorts

Construction and validation of a sialylation-related LncRNA prognostic model

The column line graphs present data on risk score, age, lymph node metastatic stage, and pathological grading. The column line plot demonstrates a clear association between the risk score and OS in patients with BLCA (Fig. 7A).

The nomogram's C-index was found to be 0.693, indicating the predictive accuracy of the model. Additionally, the calibration curves exhibited a high level of concordance between the projected probability and the observed outcomes. It is worth mentioning that the calibration curves for the OS at 1-year, 3-year, and 5-year intervals (Fig. 7E) exhibited a tight adherence to the 45-degree line. In addition, the nomograms exhibited area under the curve (AUC) values of 0.778, 0.743, and 0.748 for the 1-year, 3-year, and 5-year OS periods, as illustrated in Fig. 7B-D. The nomograms exhibit superior performance compared to individual clinical predictors, as indicated by their high AUC values. This suggests that the integration of many risk indicators can improve the prognostic accuracy for BLCA. PCA revealed the distribution of the two risk groups along two axes. This observation suggests that the riskscore-associated lncRNAs exhibit a more effective classification of BLCA patients into high-risk and low-risk groups (Fig. 7F), compared to sialylation-related lncRNAs (Fig. 7G), sialylation -related genes (Fig. 7H), and all genes (Fig. 7I). These findings highlight the superior discriminative ability of the riskscore-associated lncRNAs in the identification process.

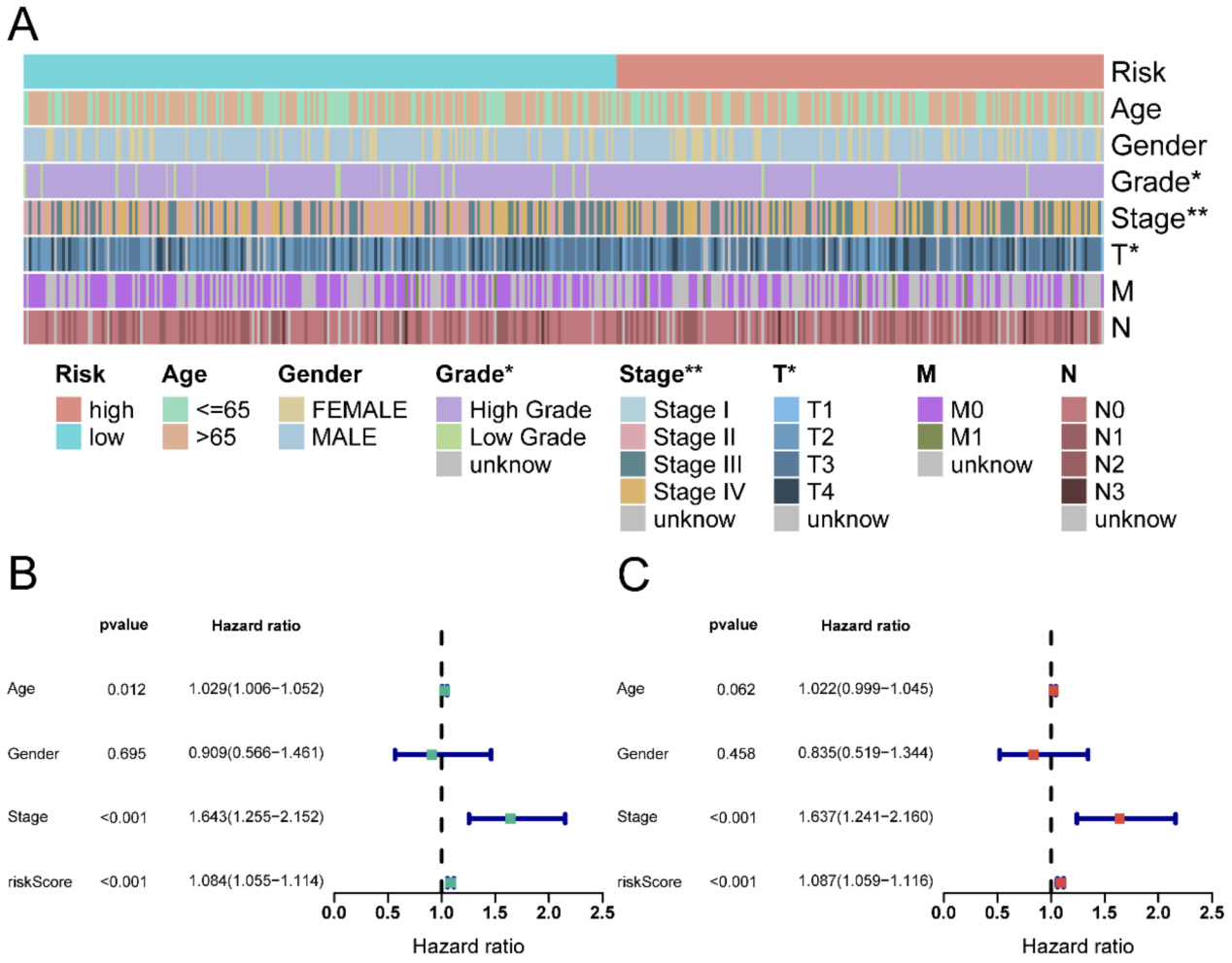

Fig. 5 Value assessment of prognostic sialylation-related lncRNAs. (**A**) The result of chi-squared test. (**B**) Univariate Cox regression analysis. (**C**) Multivariate Cox regression analysis

(F) Principal component analysis (PCA) of the risks-core-related lncRNAs. (G) PCA of the sialylation-related lncRNAs. (H) PCA of the sialylation -related genes. (I) PCA of the all genes.

Assessing the immune microenvironment of sialylation-related signature score

In order to delve deeper into the relationship between the immune process and risk scores, we conducted an assessment of lymphocyte lines using CIBERSORT. The results of the CIBERSORT analysis (Fig. 8A), indicate that the high-risk group demonstrated a greater prevalence of activated T cells CD4+memory resting Macrophages M0. Conversely, the low-risk group exhibited a larger prevalence of active Plasma cells, CD8+T cells, and regulatory T cells (Tregs). The correlation between lymphocyte populations and sialylation-related lncRNAs is depicted in Fig. 9A. Significantly, there were high associations observed between *AC104785.1* and Treg, T cells CD4+memory activated, Plasma cells, and Macrophages M1. Additionally, a robust link was found between *AC012625.1* and Plasma cells. Furthermore, there was a significant association observed between T cells follicular helper and Macrophage M1 with *GRASLND*. Similarly, T cells CD4 memory resting, T cells CD4 memory activated, and mast cells resting exhibited a high correlation with *ARHGAP5-AS1*.

Our analysis revealed significant variations in immune function between the high and low-risk groups, particularly for APC_co_inhibition, Macrophages, MHC_class_I, Parainflammation, TIL, and Treg, as demonstrated in Fig. 8B. In addition, our investigation involved the evaluation of a total of 79 genes related with immune checkpoint. Among these, 47 genes displayed differences between the high and low-risk groups, as depicted in Fig.S1. Furthermore, the analysis of Fig. 9B demonstrates a greater occurrence of C1 subtypes in the low-risk groups, while the high-risk groups exhibit a higher proportion of C2 subtypes ($P=0.001$).

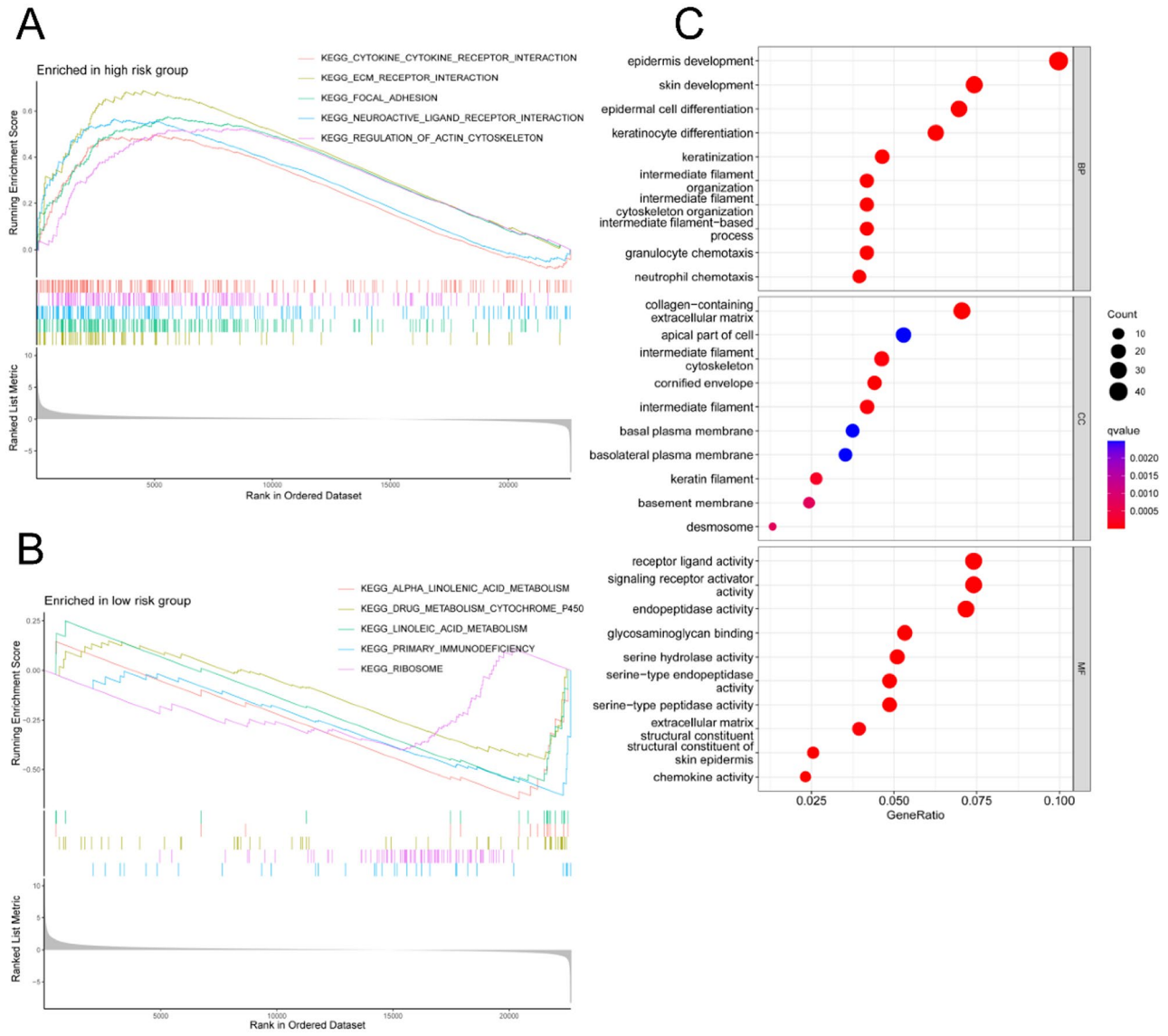

Fig. 6 Potential mechanism analysis of the sialylation-related LncRNA signature. (**A**) The top five enriched pathways in the low-risk group. (**B**) The top five enriched pathways in the low-risk group. (**C**) The result of the GO enrichment analysis

Tumor mutational burden and drug sensitivity analysis

Both in the high-risk (HR) and low-risk (LR) group, *TP53*, *TTN*, *KMT2D*, *MUC16*, *ARID1A* and *KDM6A* exhibited mutation rates exceeding 20% (Fig. 10A, B). Patients with low tumor mutation burden (TMB) had shorter OS compared to those with high TMB (Fig. 10C). Notably, when considering both risk score and TMB, it was evident that patients with low TMB and high risk exhibited the poorest prognosis (Fig. 10D).

In order to conduct a thorough evaluation of therapeutic efficacy in response to various chemotherapeutic drugs, we calculated the IC50. Our results showed that Sorafenib was less sensitive in the high-risk group, while it was more sensitive in the low-risk group compared to Cisplatin, Docetaxel, and Dasatinib. Therefore, our risk profiles can serve as a valuable tool in evaluating chemotherapeutic drug sensitivity (Fig. 10E-H).

Expression validation of sialylation-related LncRNAs signature

The expression of model-associated lncRNAs was confirmed at the cellular level through quantitative real-time polymerase chain reaction (qRT-PCR). In bladder cancer (BCa) cell lines, including *ARHGAP5-AS1*, *GRASLND*, and *LINC01508*, their expression levels were notably elevated when compared to human urinary epithelial SV-HUC-1 cell lines (Fig. 11A-C).

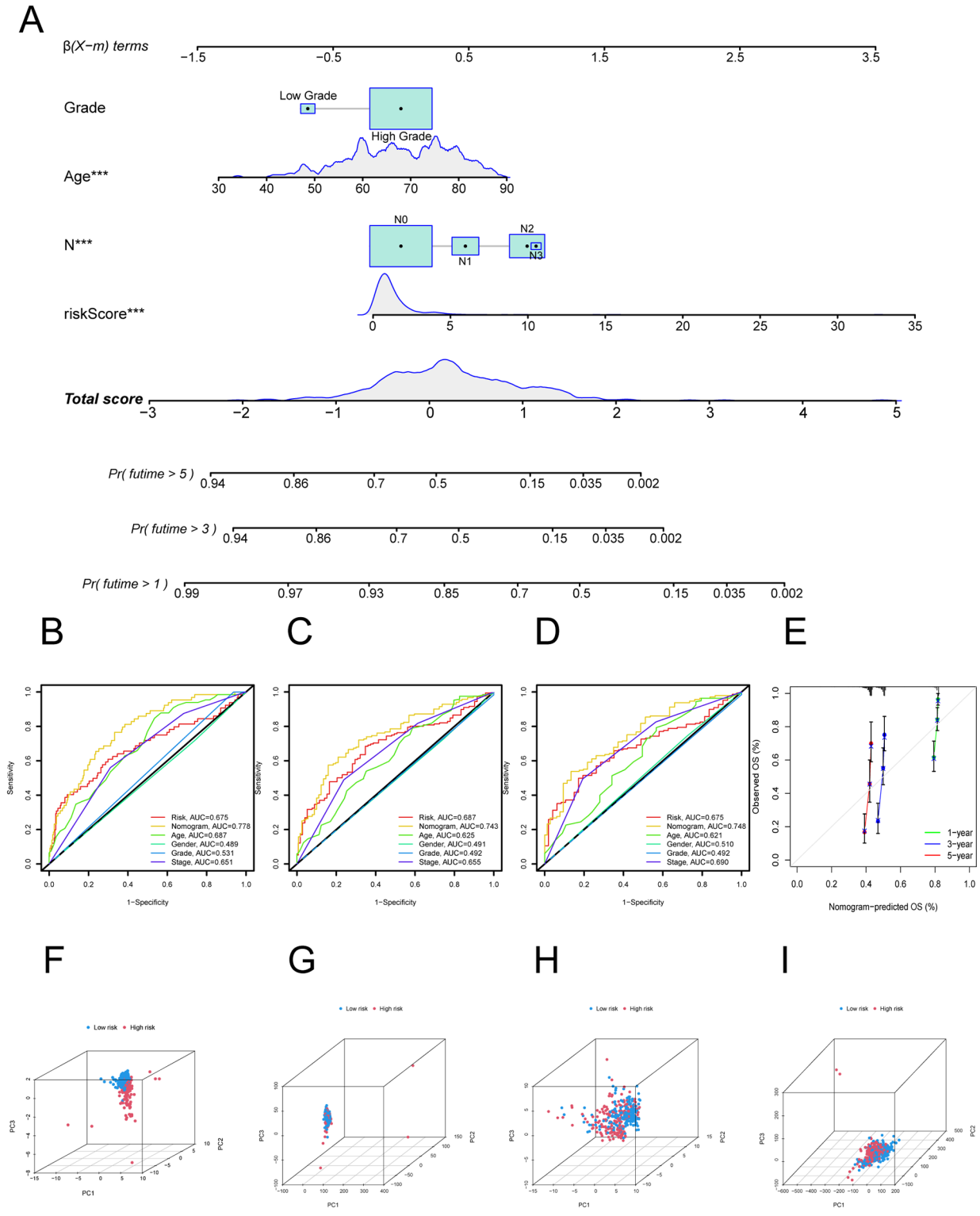

Fig. 7 Construction and validation of a sialylation-related LncRNA prognostic model. (**A**) Nomogram constructed by independent prognostic factors. (**B**-**D**) ROC curves show the predictive accuracy of the nomogram, risk score, and other clinical characteristics. (**E**) Calibration curves for the OS at 1-year, 3-year, and 5-year intervals

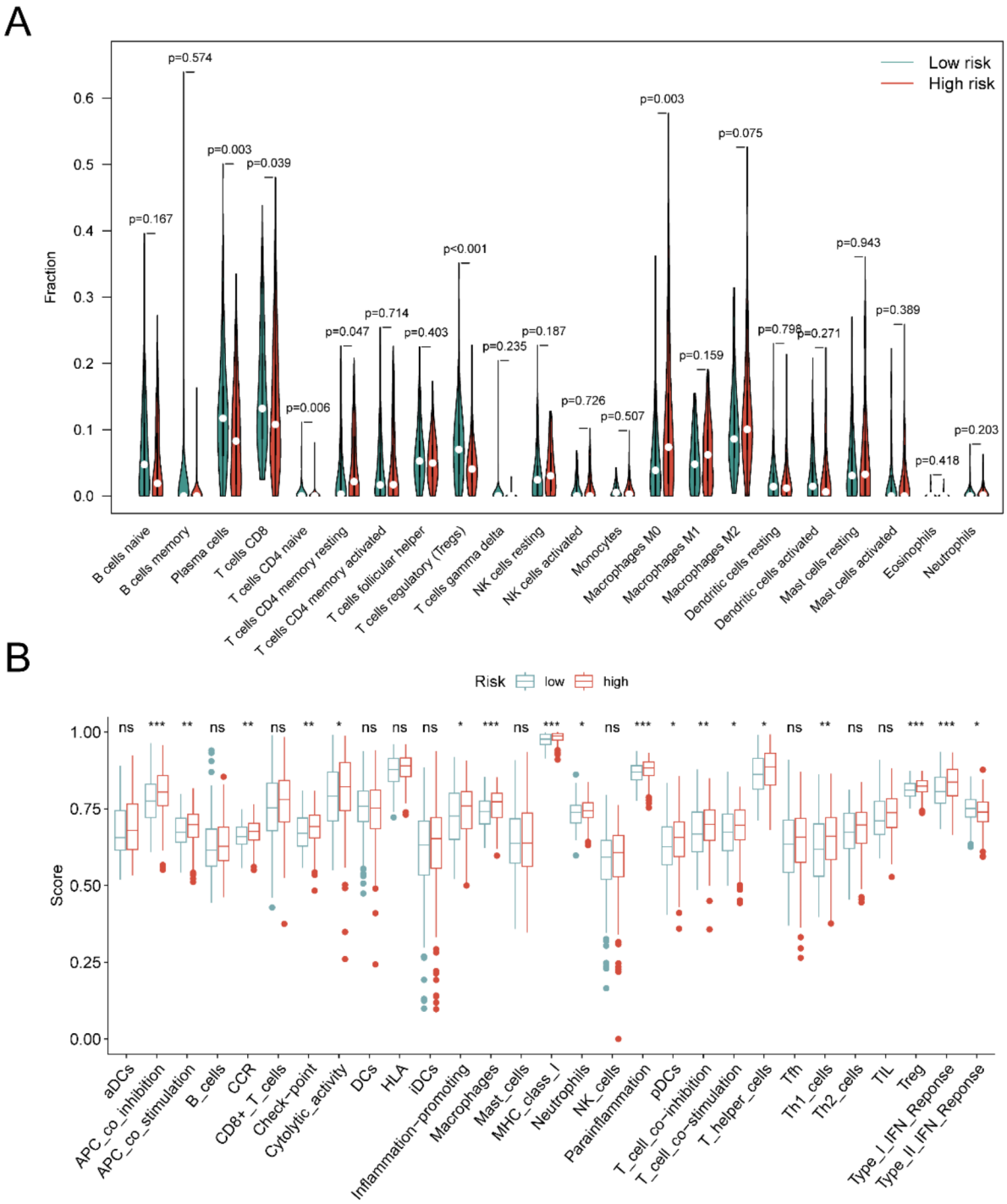

Fig. 8 Immune cell infiltration and immune function of the different risk groups. (**A**) Immune cell infiltration of the high and low-groups. (**B**) Immune function of the high and low-risk groups

Discussion

We followed a sequential approach, employing univariate Cox, lasso, and multivariate Cox regression analyses, to create a novel signal based on 6 lncRNAs. Our findings highlight the potential of this 6-lncRNA signature in distinguishing patient prognoses, offering a theoretical basis for clinical treatment strategies. Furthermore, we constructed nomograms that incorporate age, grading, N-staging, and risk scores, facilitating the prediction of 1-, 3-, and 5-year survival. These nomograms exhibited stronger predictive power compared to traditional clinical staging, hinting at the potential for optimizing

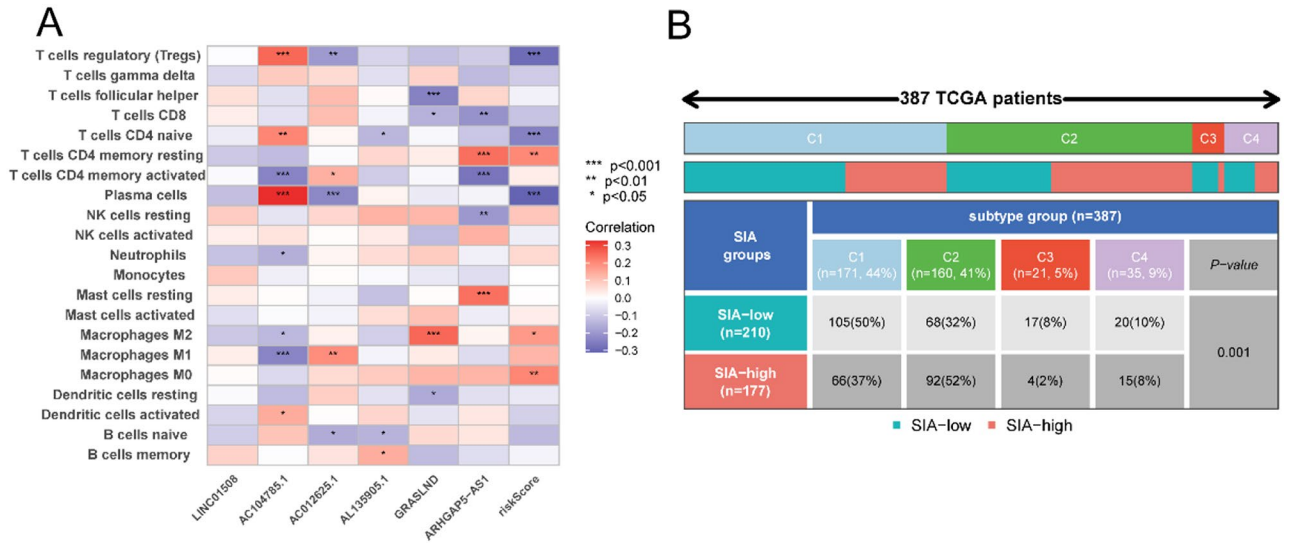

Fig. 9 Immunoscape of sialylation-related lncRNAs and risk grops. (**A**) The correlation between lymphocyte populations and sialylation-related lncRNAs. (**B**) Immune subtypes of the high and low-risk groups

current clinical staging classifications. The calibration plots bolstered this notion by demonstrating superior predictive performance for actual survival prognoses in BCa patients. This underscores the wide-ranging applications and practicality of salivary acid-associated lncRNAs in prognostic prediction. The results of our study emphasize the potential of the identified 6-lncRNA signature in effectively differentiating patient prognoses, hence providing a theoretical foundation for the development of clinical treatment options. In addition, we developed nomograms that integrate age, grading, N-staging, and risk scores, enabling the estimation of 1-, 3-, and 5-year survival rates. The nomograms demonstrated superior predictive capability in comparison to conventional clinical staging, suggesting the possibility of enhancing existing clinical staging categories. The hypothesis that the calibration plots provided evidence of better prediction accuracy for real survival prognoses in breast cancer patients was supported. This highlights the extensive range of uses and feasibility of lncRNAs related with salivary acid in prognostic prediction.

We observed a strong association between the high-risk group and several critical pathways through GSEA analysis. The high-risk group exhibited connections with pathways involving cytokine-cytokine receptor interactions, ECM-receptor interactions, focal adhesion, neuroactive ligand-receptor interactions, and regulation of the actin cytoskeleton. The results of this study are consistent with previous research suggesting that cytokines may contribute to the development, advancement, and control of tumors [29]. In addition, it is widely recognized that the modification of focal adhesion plays a crucial role in facilitating the contacts between tumor cells and ECM, ultimately leading to the promotion of tumor growth and metastasis [30]. This relationship exerts a significant impact on critical biological processes such as invasion, epithelial-mesenchymal transition (EMT), tumor angiogenesis, and stromal fibrosis [31–33]. Furthermore, the notion of EMT holds significant importance in numerous biological phenomena, with a special focus on its involvement in the invasion of tumors and the advancement of metastasis [34]. In various malignancies affecting human beings, neoplastic cells undergo EMT, a phenotypic alteration characterized by the loss of cellular polarity and intercellular adhesion [35]. Conversely, these organisms acquire migratory and invasive characteristics, hence facilitating their proliferation and metastasis [36]. The progression of bladder cancer has been extensively described. Meanwhile, numerous studies have consistently confirmed the significant importance of dynamic alterations in sialylation during EMT and their possible significance in the process of cancer metastasis [37]. Meanwhile, sialylation undergoes significant dynamic changes during EMT [38]. During EMT, the expression levels of sialylation-related enzymes, such as STs and neuraminidase (NEU), change markedly [39]. Existing research indicates that the expression of STs significantly increases in the early stages of EMT, leading to an increase in sialic acid content, which may be related to changes in glycosylation patterns on the cell surface [40]. Additionally, the distribution of sialic acid modifications on the cell membrane changes as EMT progresses. During EMT, sialic acid residues appear more frequently on the N-glycans of the cell surface, potentially affecting cell adhesion and signal transduction [41]. Based on existing research and our analysis, we hypothesize that sialylation mediates the dynamic changes in EMT through variations in expression levels, thereby influencing the growth

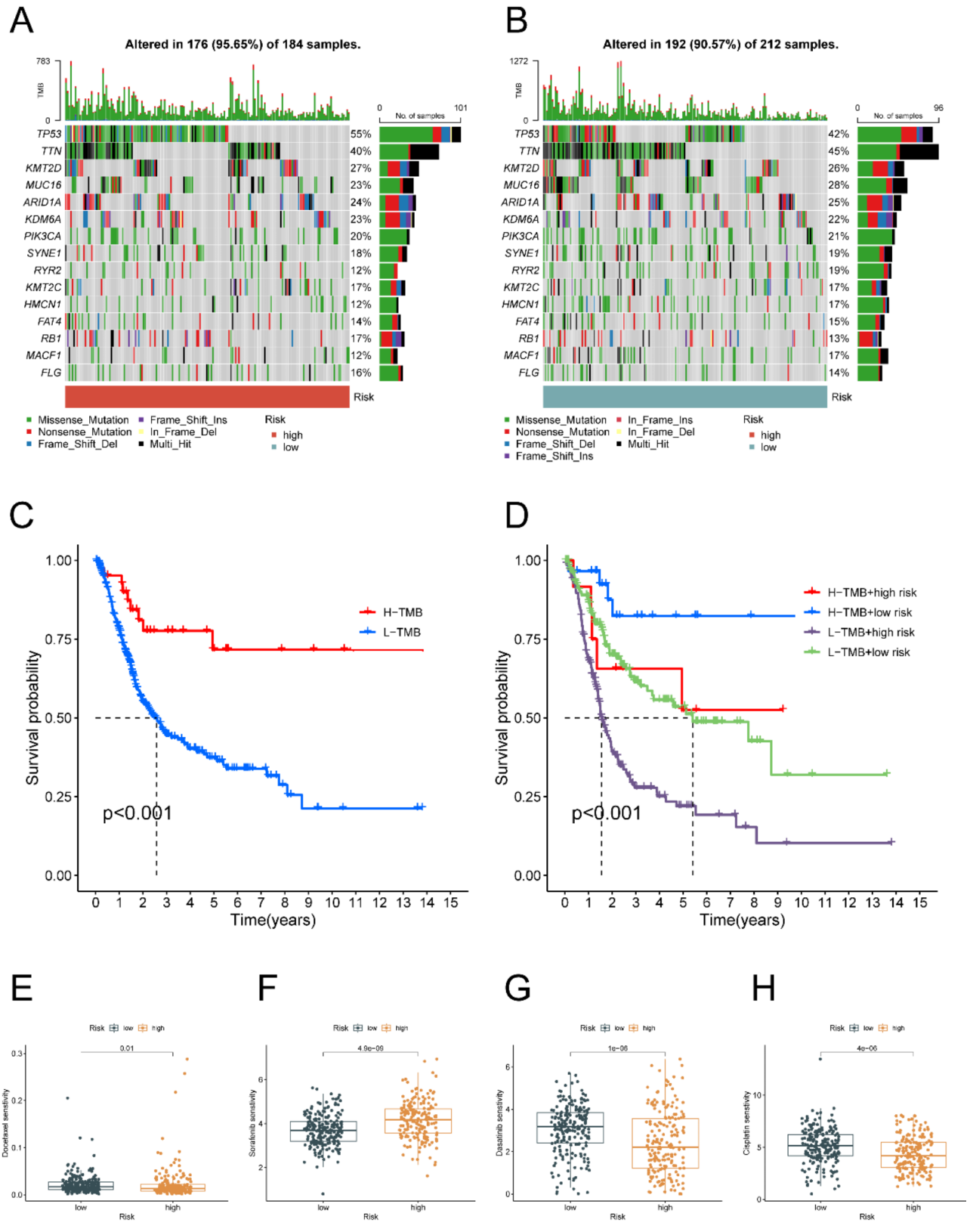

Fig. 10 Tumour mutational burden (TMB) and Drug susceptibility analysis. (**A**) The waterfall chart of the frequently mutated genes in the high-risk group. (**B**) The waterfall chart of the frequently mutated genes in the low-risk group. (**C**) Kaplan–Meier curves of high- and low-TMB on OS. (**D**) Kaplan–Meier curves of TMB and risk score on OS. exploration of the risk signature and drug sensitivity. (**E**) Docetaxel. (**F**) Sorafenib. (**G**) Dasatinib. (**H**) Cisplatin

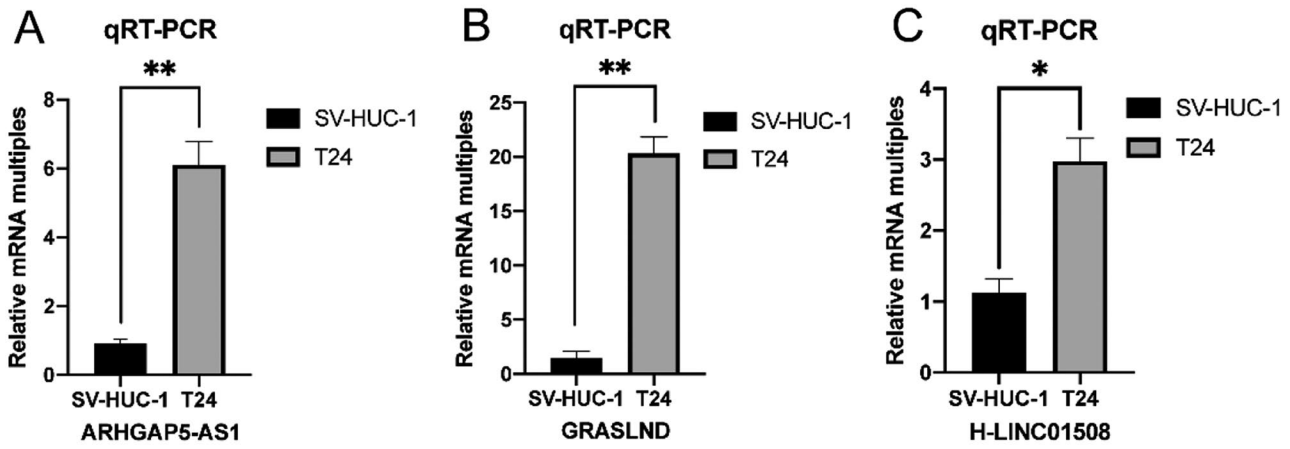

Fig. 11 Relative expression of lncRNAs. (**A**) *ARHGAP5-AS1*. (**B**) *GRASLND*. (**C**) *LINC01508*

and metastasis of bladder cancer. Furthermore, The observed results are consistent with the outcomes of our analysis on biological pathway enrichment. Specifically, we found a notable enrichment in pathways associated with the formation of the epidermis and the differentiation of epidermal cells. Based on the analysis of the high-risk group associated with sialylation in this study, we hypothesize that sialylation contributes to the growth, survival, invasion, and metastasis of bladder cancer through multiple pathways. First, sialylation may regulate the growth, survival, and metastasis of bladder cancer cells by affecting cytokine expression. Second, sialylation might promote the invasion and metastasis of bladder cancer cells by modifying the ECM or its receptors. Additionally, focal adhesions, which are complexes that connect the ECM to the actin cytoskeleton, play a crucial role in regulating cell movement and signal transduction. Proteins such as integrins and focal adhesion kinase (FAK) are important in cell migration, invasion, and anti-apoptosis, and these may also be targets through which sialylation regulates the progression of bladder cancer. These results provide clear direction for our subsequent in-depth mechanistic studies.

We conducted an in-depth analysis of the immune landscape within the risk model. Firstly, our research entailed an evaluation of the composition of immune cells in both high-risk groups utilizing CIBERSORT and ssGSEA methodologies. The findings of our study indicate that the group classified as high-risk demonstrated significantly elevated levels of M0 macrophage infiltration, as well as T-cells CD4 memory inactive, which aligns with previous investigations in this field. In contrast, the group classified as low-risk exhibited heightened amounts of plasma cells, T cells CD8, T cells CD4 naive, and Tregs, suggesting that individuals with low-risk profiles generally demonstrated a more prominent immune response to tumor development. Immunological checkpoints are a distinct group of receptors and ligands that regulate programmed cell death, so modulating immunological responses and safeguarding against autoimmune harm, thereby playing a role in maintaining immune homeostasis [42]. In the examination of the two risk-scoring cohorts, it was observed that the low-risk group exhibited elevated expression levels of some immune checkpoint genes, namely *CEACAM1*, *BTNL9*, *CD96*, *TNFRSF14*, and *ICOSLG*. Conversely, the high-risk group displayed heightened expression of other immune checkpoint genes. The observed variation in immune checkpoint expression indicates that individuals classified as low-risk exhibit reduced levels of immune checkpoint infiltration, which in turn facilitates immunological tolerance and has implications for patient prognosis. Additionally, a prevalence of immunological subtype C2 (characterized by IFN-γ dominance) was observed in the high-risk cohort, representing 54% of the total cases. The prevalence of this particular subtype was found to be considerably higher in the high-risk group in comparison to the low-risk group, while the opposite trend was observed for subtype C1. There is an inverse relationship between risk score and TMB, which is a biological marker for immune checkpoint inhibitors(ICIs). Previous studies have demonstrated that IFN-γ dominant subtypes exhibit a strong response to immune checkpoint inhibitors therapy and derive therapeutic benefits from immunotherapy interventions utilizing drugs such as anti-PD-1 and anti-CTLA-4 [43]. Immunotherapy based on ICIs is currently a prominent area of focus in the field of oncology treatment and serves as a pivotal component in the management of diverse malignancies.

However, this study has several limitations. First, the sialylation-related lncRNA features were developed and validated solely within the TCGA database, which limits the model's stability and applicability. Second, the sample size for the normal group was limited due to cell line variability, potentially explaining the inconsistency between qPCR validation results and TCGA database

differential analysis results. Additionally, the AUC values for the risk model constructed using six prognostically relevant sialylation-related lncRNAs did not meet our criteria for reliability at 1, 3, and 5 years. This could be due to factors such as sample size, feature variables, and the chosen machine learning algorithms. Lastly, while we proposed plausible mechanisms for sialylation regulation in bladder cancer based on our findings, experimental validation is still lacking. Given these limitations, we suggest future research directions. We plan to collect more clinical samples through multicenter collaborations, introduce more feature variables, and optimize the model using various machine learning algorithms, such as decision trees and their ensembles (e.g., random forest, gradient boosting trees), support vector machines (SVM), and ensemble learning methods (e.g., stacking, bagging, boosting) to select the optimal model, ensuring broad applicability and stability. We will collect tumor and normal tissue samples from clinical patients and design a series of experiments, including qPCR, in vitro cell experiments, and in vivo animal model experiments, to validate the specific mechanisms of sialylation regulation in bladder cancer. Specifically, we will investigate how sialylation-related lncRNAs affect the proliferation, migration, and invasion capabilities of bladder cancer cells and their roles in the EMT process. In summary, this study has successfully created and verified a robust integrated model for predicting the prognostic status of BCa patients, showcasing superior predictive capabilities. We conducted a comprehensive assessment of disparities in clinicopathological features, TMB, the immune microenvironment, and responses to chemotherapy between high-risk and low-risk groups. These findings significantly contribute to our comprehension of TIL characteristics, the identification of new targets for BCa immunotherapy, and lay a crucial foundation for future personalized treatment approaches.

Conclusion

This study explored the relationship between sialylation-related lncRNA and bladder cancer prognosis. Six sialylation-related lncRNAs associated with the prognosis of BLCA were identified. Based on these identified lncRNAs, we constructed a predictive model for BLCA prognosis, which had the potential to guide clinical decision-making in BLCA management. Our analysis explored various aspects of the connection between sialylation and BLCA, encompassing clinicopathological characteristics, underlying mechanisms, somatic mutations, the immune microenvironment, chemotherapeutic responses, and anticipated drug sensitivity. Additionally, we conducted in vitro cellular investigations to validate theoretical findings and lay the groundwork for future clinical applications. In the future, we plan to collect more clinical samples through multi-center collaborations, introduce more feature variables, and use various machine learning algorithms to optimize the model. Furthermore, we aim to verify the specific mechanisms of sialylation regulation in bladder cancer through in vitro cell experiments and in vivo animal model experiments.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12894-024-01613-6.

Supplementary Material 1

Author contributions

Conception and design: SL, HN. Data analysis and manuscript writing: YQ. Cellular experiment: XT.

Funding

This study received financial support from the National Natural Science Foundation of China (Grant 82404089); and Youth Research Fund of Affiliated Hospital of Qingdao University (QDFYQN2023117).

Data availability

Data is provided within the manuscript and supplementary information files.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 2 February 2024 / Accepted: 30 September 2024
Published online: 10 October 2024

References

1. van Hoogstraten LMC, Vrieling A, van der Heijden AG, Kogevinas M, Richters A, Kiemeney LA. Global trends in the epidemiology of bladder cancer: challenges for public health and clinical practice. Nat Rev Clin Oncol. 2023;20:287–304.
2. Lenis AT, Lec PM, Chamie K. Bladder Cancer: a review. JAMA. 2020;324:1980–91.
3. Babjuk M, Burger M, Capoun O, Cohen D, Compérat EM, Dominguez Escrig JL, et al. European Association of Urology Guidelines on non-muscle-invasive bladder Cancer (Ta, T1, and carcinoma in situ). Eur Urol. 2022;81:75–94.
4. Kirkali Z, Chan T, Manoharan M, Algaba F, Busch C, Cheng L, et al. Bladder cancer: epidemiology, staging and grading, and diagnosis. Urology. 2005;66(6 Suppl 1):4–34.
5. Sathianathen NJ, Regmi S, Gupta S, Konety BR. Immuno-Oncology approaches to salvage treatment for non-muscle invasive bladder Cancer. Urol Clin North Am. 2020;47:103–10.
6. Li F, Ding J. Sialylation is involved in cell fate decision during development, reprogramming and cancer progression. Protein Cell. 2019;10:550–65.
7. Büll C, Heise T, Adema GJ, Boltje TJ. Sialic Acid mimetics to target the Sialic Acid-Siglec Axis. Trends Biochem Sci. 2016;41:519–31.
8. Macauley MS, Crocker PR, Paulson JC. Siglec-mediated regulation of immune cell function in disease. Nat Rev Immunol. 2014;14:653–66.
9. Ikehara Y, Ikehara SK, Paulson JC. Negative regulation of T cell receptor signaling by Siglec-7 (p70/AIRM) and Siglec-9. J Biol Chem. 2004;279:43117–25.

10. Vuchkovska A, Glanville DG, Scurti GM, Nishimura MI, White P, Ulijasz AT, et al. Siglec-5 is an inhibitory immune checkpoint molecule for human T cells. Immunology. 2022;166:238–48.
11. Läubli H, Pearce OMT, Schwarz F, Siddiqui SS, Deng L, Stanczak MA, et al. Engagement of myelomonocytic siglecs by tumor-associated ligands modulates the innate immune response to cancer. Proc Natl Acad Sci U S A. 2014;111:14211–6.
12. Benmerzoug S, Chevalier MF, Verardo M, Nguyen S, Cesson V, Schneider AK, et al. Siglec-6 as a new potential Immune checkpoint for bladder Cancer patients. Eur Urol Focus. 2022;8:748–51.
13. Benmerzoug S, Chevalier MF, Villier L, Nguyen S, Cesson V, Schneider AK, et al. Siglec-7 May Limit Natural Killer Cell-mediated antitumor responses in bladder Cancer patients. Eur Urol Open Sci. 2021;34:79–82.
14. Videira PA, Correia M, Malagolini N, Crespo HJ, Ligeiro D, Calais FM, et al. ST3Gal.I sialyltransferase relevance in bladder cancer tissues and cell lines. BMC Cancer. 2009;9:357.
15. Jian Y, Chen Q, Al-Danakh A, Xu Z, Xu C, Sun X, et al. Identification and validation of sialyltransferase ST3Gal5 in bladder cancer through bioinformatics and experimental analysis. Int Immunopharmacol. 2024;138:112569.
16. Ouyang S, Liu J-H, Ni Z, Ding G-F, Wang Q-Z. Downregulation of ST3GAL5 is associated with muscle invasion, high grade and a poor prognosis in patients with bladder cancer. Oncol Lett. 2020;20:828–40.
17. Prensner JR, Chinnaiyan AM. The emergence of lncRNAs in cancer biology. Cancer Discov. 2011;1:391–407.
18. Zhang Y, Chen X, Lin J, Jin X. Biological functions and clinical significance of long noncoding RNAs in bladder cancer. Cell Death Discov. 2021;7:278.
19. Badowski C, He B, Garmire LX. Blood-derived lncRNAs as biomarkers for cancer diagnosis: the Good, the bad and the Beauty. NPJ Precis Oncol. 2022;6:40.
20. Mounir M, Lucchetta M, Silva TC, Olsen C, Bontempi G, Chen X, et al. New functionalities in the TCGAbiolinks package for the study and integration of cancer data from GDC and GTEx. PLoS Comput Biol. 2019;15:e1006701.
21. Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, et al. Ensembl 2022. Nucleic Acids Res. 2022;50:D988–95.
22. Me R, B P DW, Cw YH. L, W S, Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43.
23. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. Innov (Camb). 2021;2:100141.
24. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12:453–7.
25. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the Tidyverse. J Open Source Softw. 2019;4:1686.
26. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-Seq data. BMC Bioinformatics. 2013;14:7.
27. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2013;41 Database issue:D955–61.
28. Maeser D, Gruener RF, Huang RS. oncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. Brief Bioinform. 2021;22:bbab260.
29. Balkwill F, Mantovani A. Inflammation and cancer: back to Virchow? Lancet. 2001;357:539–45.
30. Huang J, Zhang L, Wan D, Zhou L, Zheng S, Lin S, et al. Extracellular matrix and its therapeutic potential for cancer treatment. Signal Transduct Target Ther. 2021;6:153.
31. Fan H, Zhao X, Sun S, Luo M, Guan J-L. Function of focal adhesion kinase scaffolding to mediate endophilin A2 phosphorylation promotes epithelial-mesenchymal transition and mammary cancer stem cell activities in vivo. J Biol Chem. 2013;288:3322–33.
32. Frisch SM, Schaller M, Cieply B. Mechanisms that link the oncogenic epithelial-mesenchymal transition to suppression of anoikis. J Cell Sci. 2013;126 Pt 1:21–9.
33. Shibue T, Brooks MW, Inan MF, Reinhardt F, Weinberg RA. The outgrowth of micrometastases is enabled by the formation of filopodium-like protrusions. Cancer Discov. 2012;2:706–21.
34. Nieto MA, Huang RY-J, Jackson RA, Thiery JP. EMT: 2016. Cell. 2016;166:21–45.
35. Pastushenko I, Blanpain C. EMT Transition States during Tumor Progression and Metastasis. Trends Cell Biol. 2019;29:212–26.
36. Rokavec M, Öner MG, Li H, Jackstadt R, Jiang L, Lodygin D, et al. IL-6R/STAT3/miR-34a feedback loop promotes EMT-mediated colorectal cancer invasion and metastasis. J Clin Invest. 2014;124:1853–67.
37. Du J, Hong S, Dong L, Cheng B, Lin L, Zhao B, et al. Dynamic sialylation in transforming growth Factor-β (TGF-β)-induced epithelial to mesenchymal transition. J Biol Chem. 2015;290:12000–13.
38. Varki A. Sialic acids in human health and disease. Trends Mol Med. 2008;14:351–60.
39. Schauer R. Achievements and challenges of sialic acid research. Glycoconj J. 2000;17:485–99.
40. Fuster MM, Esko JD. The sweet and sour of cancer: glycans as novel therapeutic targets. Nat Rev Cancer. 2005;5:526–42.
41. Varki A. Biological roles of glycans. Glycobiology. 2017;27:3–49.
42. Dyck L, Mills KHG. Immune checkpoints and their inhibition in cancer and infectious diseases. Eur J Immunol. 2017;47:765–79.
43. Sun H, Liu S-Y, Zhou J-Y, Xu J-T, Zhang H-K, Yan H-H, et al. Specific TP53 subtype as biomarker for immune checkpoint inhibitors in lung adenocarcinoma. EBioMedicine. 2020;60:102990.

Publisher's note