scientific reports

Check for updates

OPEN

Mechanism of action for Troxerutin targeting the sialylation-related gene EGLN3 for the treatment of LUAD

Yanan Wu[1], Yanlei Ge[2,3,4], Junqing Gan[2,3,4], Ye Jin[3,4,5], Yishuang Cui[1], Xuan Zheng[1], Xuemin Yao[1] & Guogui Sun[2,3,4,6]✉

Studies have demonstrated that sialylation changes play a vital part in lung adenocarcinoma (LUAD), yet the specific mechanism is uncertain. Hence, in the present research, we screened sialylation-related biomarkers in LUAD using the bioinformatic strategy, predicted the drugs and performed relevant experiments to explore their role in regulating LUAD. The TCGA-LUAD, GSE31210, and GSE13213 datasets were combined to form LUAD ensemble. The sialylation-related genes (SRGs) linked with LUAD prognosis were determined by univariate Cox regression analysis, and their expressions and mutations in LUAD were analyzed in GSCA database. Then, depending on the consistent clustering of prognostic SRGs, LUAD patients were divided into sialylation-related subgroups, followed by the investigation of survival, immunity, and clinical characteristics in the subgroups. LASSO regression analysis was further employed to identify prognostic gene signatures and to build a sialylation-related model to predict the prognosis of LUAD patients. The gene signature were validated using RT-qPCR and used for predicting target medicines using molecular docking to further investigate the potential therapies for LUAD patients. A total of 26 SRGs in LUAD ensemble were associated with prognosis, and LUAD samples were classified into two sialylation-related subgroups based on these SRGs. Intergroup comparisons revealed that patients in Cluster A had greater survival rates, as well as higher immune infiltration. The risk prognostic model built based on 6 prognostic gene signature was able to effectively predict the survival of LUAD patients. Finally, the experimental findings indicated that Troxerutin exhibits a strong binding energy to the sialylation-related gene EGLN3, which could greatly reduce the growth of LUAD by inhibiting the expression of EGLN3, thus limiting the capacity of LUAD cells in the proliferation, migration, and invasion. Troxerutin could target and lower the expression of sialylation-related gene EGLN3, reducing LUAD cells' ability to proliferate, migrate, and invade, making it an essential reference for LUAD prevention and treatment.

Keywords Lung adenocarcinoma, Sialylation, Troxerutin, EGLN3

Lung cancer is one of the most common malignant tumors in the world, with lung adenocarcinoma (LUAD) accounting for the majority of cases. Patients typically have symptoms such as chest pain, cough, and blood in sputum[1–3], and clinical treatment consists primarily of surgery, radiotherapy, and chemotherapy[4]. Although diagnostic and therapeutic methods for LUAD have been continually developed, patients with LUAD tend to be diagnosed at a late stage, with a 5-year survival rate of just approximately 15%[5]. Currently, changes and fusions of different genes, including KRAS, MET, and BRAF, have been extensively examined in the etiology of LUAD, with the oncogenic activating mutation of the EGFR gene on chromosome 7 being the most well-documented[6]. Targeted immunotherapy focusing on several biomarkers has become increasingly essential in LUAD. One study demonstrated that the immune checkpoint inhibitor anti-PD-1/PD-L1 supported 18 patients with advanced

[1]School of Public Health, North China University of Science and Technology, Tangshan 063210, Hebei, China. [2]Department of Radiotherapy and Chemotherapy, Affiliated Hospital of North China University of Science and Technology, Tangshan 063000, Hebei, China. [3]Department of Hebei Key Laboratory of Medical-Industrial Intergration Precision Medicine, Tangshan 063000, Hebei, China. [4]Department of Tangshan Key Laboratory of Medical-Industrial Intergration Precision Medicine, Tangshan 063000, Hebei, China. [5]Clinical Medicine School, North China University of Science and Technology, Tangshan 063000, Hebei, China. [6]School of Pharmacy, North China University of Science and Technology, Tangshan 063210, Hebei, China. ✉email: guogui_sun2021@sina.com

LUAD along with ground-glass nodules[3], bringing hope to LUAD patients worldwide. However, most patients acquire resistance to targeted therapy, and the prognosis remains poor[7]. As a result, it is critical to uncover novel prognostic and pathogenic features of LUAD.

Sialylation is a type of glycosylation that is primarily regulated by salivary acid transferase activity[8]. A predictive model for LUAD constructed from sialylation-associated lncRNAs has recently been created[8]. Sialyl LewisX was discovered to be highly differently expressed in LUAD, with tumors that heavily expressed sialyl LewisX having a greater frequency of vascular invasion[9]. Sialyl Lewis X-i antigen has been employed as a diagnostic test for LUAD[10], however, its precise involvement in LUAD remains uncertain. Therefore, it is critical to continue looking into the mechanism of sialylation-related genes (SRGs) in LUAD, as well as their predictive characteristics.

This study applied bioinformatic tools, including differential expression analysis, consensus clustering, enrichment analysis, and immune infiltration, according to LUAD-related datasets in The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) public databases, to establish sialylation-related prognostic characteristics in LUAD. In the meantime, sialylation-related biomarkers as well as potential drugs targeting SRGs were determined by a variety of experimental techniques utilizing LUAD cell lines and clinical samples, with the goal of offering crucial references for the diagnosis and management of LUAD.

Materials and methods

Data source

The expression matrix and clinical information of 524 LUAD patients from the UCSC-XENA database (TCGA-LUAD, https://xenabrowser.net/datapages/), 226 LUAD patients in GSE31210 on the GPL570 platform and 117 LUAD patients in GSE13213 on the GPL6480 platform from the Gene Expression Omnibus (GEO, (http://www.ncbi.nlm.nih.gov/geo/) database were downloaded. Meanwhile, 109 sialylation-related genes (SRGs) were sourced from the MSigDB database (https://www.gsea-msigdb.org/gsea/msigdb)[11].

Merging of datasets

To acquire more precise data results, the TCGA-LUAD, GSE31210, and GSE13213 datasets were pooled to form the LUAD ensemble. Following that, batch effects were removed using limma and sva[12], and principal component analysis (PCA) plots before and after merging were created using FactoMineR and factoextra, respectively[13].

Selection of SRGs related to LUAD prognosis

Univariate Cox regression analysis[14,15] was performed to identify SRGs associated with LUAD prognosis in the LUAD ensemble dataset ($HR \neq 1$, $P < 0.05$), followed by the correlation analysis among prognostic SRGs. The results were presented using ggplot2[16]. Afterwards, the expressions of prognostic SRGs in LUAD and paracancer tissues were investigated utilizing the GSCA database (http://bioinfo.life.hust.edu.cn/GSCA/#/). Meanwhile, the SNV mutation of 567 LUAD samples were evaluated, and the mutation types of top 10 prognostic SRGs were analyzed.

Construction of sialylation-related subgroups

ConsensusClusterPlus[17] was utilized to classify LUAD patients into sialylation-related subgroups based on prognostic SRGs. Next, survival analysis utilizing survival and survminer was carried out to investigate the differences in survival between subgroups. Heatmaps were created by pheatmap[15] to illustrate the correlations between sialylation-related subgroups, prognostic SRGs, and clinical characteristics in LUAD ensemble dataset. Differential expression analysis was studied between sialylation-related subgroups to locate differentially expressed genes (DEGs) ($|\log FC| > 1$, $P < 0.05$). GO and KEGG enrichment analysis was then conducted with clusterprofiler to investigate the functional pathways implicated in DEGs[18].

Immunological and functional enrichment of sialylation-related subgroups

To explore the differences in the involvement of functional pathways among sialylation-related subgroups, the HALLMARK, KEGG and Reactome pathway from the MsigDB database served as background gene sets, and the corresponding enrichment score was calculated using GSVA package[19]. Besides, to compare immune regulation between two subgroups, the ESTIMATE software[20] was implemented to determine immune, stromal, and estimate score. Also, the ssGSEA was used to analyze the infiltration of 23 immune cells between subgroups.

Risk prognosis modeling

After removing the LUAD samples lacking survival data, 500 LUAD cases from TCGA were used to build the risk score model. First, the least absolute shrinkage and selection operator (LASSO) regression analysis was performed to identify prognostic SRG signature using glmnet package[21]. The formula $\left(\text{Risk score} = \sum_{i=1}^{n} (\text{coefi} * \text{Xi})\right)$ was used to determine risk scores of LUAD patients. LUAD patients were then split into high and low risk groups by the median risk score value. Subsequently, survival between risk groups were examined[22]. The receiver operating characteristic (ROC) curve of the risk model was then plotted with pROC[23]. Univariate and multivariate Cox regression analysis was employed to discover independent prognostic factors. Nomograms were created utilizing independent prognostic factors to predict 1, 3, and 5-year survival rates of LUAD patients, and the predictive efficacy was confirmed with calibration curves.

Drug screening

A simulated screening was conducted using 17,580 molecular small molecules from the Taoist Natural Products Library. First, the 3D structure was converted using LigPrep from the Schrödinger package[24]. PDB files for

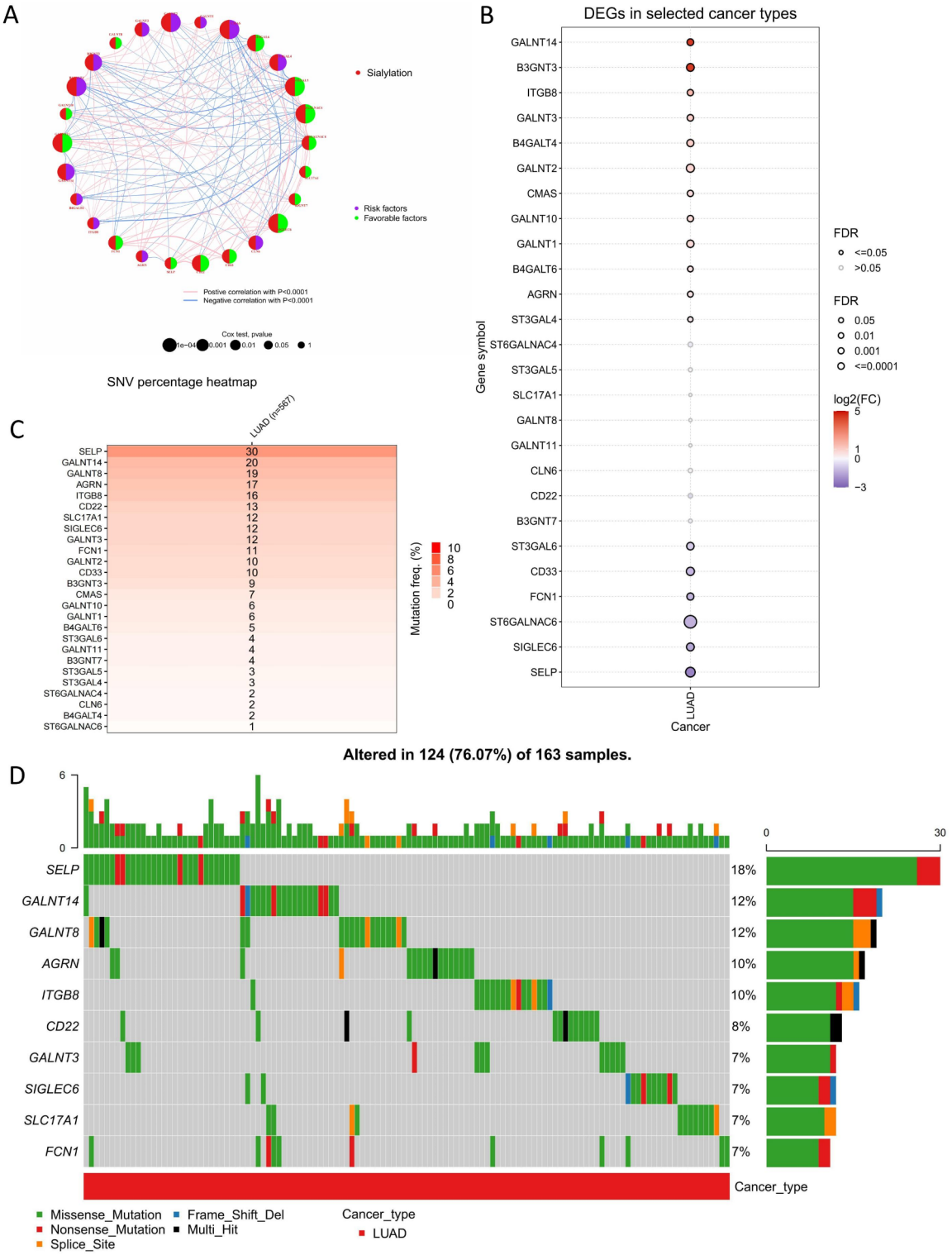


Fig. 1. Identification and analyses of sialylation-related genes (SRGs) associated with LUAD survival. (**A**) The correlation network comprising 26 SRGs associated with LUAD prognosis is presented. The blue and red lines signify negative and positive correlations between the two SRGs, respectively. The purple and green colors within the circles denote that the SRGs function as risk factors and favorable factors, respectively. The magnitude of the circles corresponds to the p-value. (**B**) The expression differences of 26 prognosis-related SRGs in LUAD and adjacent normal tissues according to the GSCA database. The solid circle represents a significant difference (FDR ≤ 0.05), while the dotted circle represents no significant difference between LUAD and adjacent normal tissues (FDR > 0.05). (**C**) The single nucleotide variation (SNV) mutation frequency heatmap of 26 SRGs in LUAD. (**D**) The top 10 mutation profile of prognosis-SRGs. The X-axis is the sample, the Y-axis is the mutated gene, and different colors represent different mutation types.

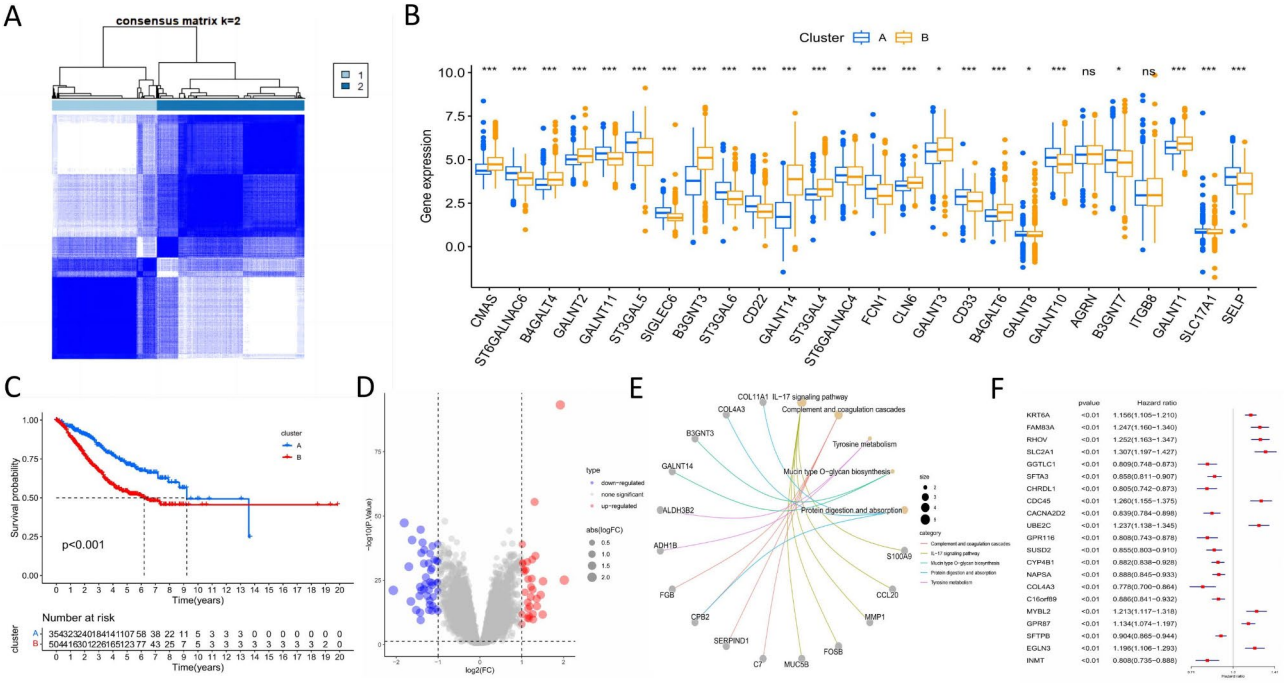


Fig. 2. Consensus clustering based on 26 prognosis-related SRGs. (**A**) The optimal consensus clustering matrix k = 2. (**B**) The expressions of 26 SRGs between sialylation-related subgroups Cluster A and Cluster B. (**C**) Kaplan-Meier curves showing the significant survival difference between sialylation-related subgroups Cluster A and Cluster B. (**D**) The volcano plot showing the differential expressed genes (DEGs) between sialylation-related subgroups Cluster A and Cluster B. (**E**) Chord diagram showing the KEGG enrichment results of DEGs. (**F**) Forest plot showing the results of univariate Cox analysis, in which 21 DEGs were found to be associated with LUAD prognosis. $^{*}P < 0.05$, $^{***}P < 0.001$, $^{ns}P > 0.05$.

EGLN3 protein were retrieved from the AlphaFold database. As a result, the Protein Preparation Wizard in the Maestro 13.0 package[25] was hydrogenated and dehydrogenated. Docking was conducted utilizing Glide's three accuracy algorithms (HTVS, SP, and XP).

Cell culture

Human lung cancer cell lines A549, PC-9, and 1299, as well as normal human bronchial epithelial cell line Beas-2b, were bought from eallbio. The 4 cell lines were cultured at 37 °C in an incubator containing 95% air and 5% CO_2, with 1299 and PC-9 cultured in 1640 medium (03.4007 C, eallbio, Beijing), and A549 and Beas-2b cultured in F12 and DMEM medium, respectively. The 4 cell lines were supplemented with 10% FBS (164210-50, Pricella, Wuhan) and 1% PS (03.1.2000a, eallbio, Beijing). When the cell fusion reached around 80%, they were washed twice with PBS (03.15017 C, eallbio, Beijing). Following that, 2 ml of trypsin (03.13005 A, eallbio, Beijing) was added for digestion, which was stopped by adding 2 ml of full culture solution after 5–10 min. The cell suspension was centrifuged for 5 min at 1500 rpm before being counted and placed on a plate. Culture conditions for Troxerutin intervention were as above.

IC50 test for Troxerutin intervention

A549 and PC-9 cells were seeded in 96-well plates at 2,000 cells per well and incubated overnight for cell attachment. Different concentrations (699.1 μm/ml and 365.1 μm/ml) of Troxerutin and 0.1% DMSO were added to the designated wells as Troxerutin-treated and control groups, respectively. The cell plates were incubated for two days. After incubation, 10 μL CCK-8 reagent (ZP328, ZOMANBIO, Beijing) was added to each well and incubated at 37 °C for 2 h. The light absorbance of each sample (OD value) was measured at 450 nm using a microplate reader.

Real-time reverse transcription polymerase chain reaction (RT-qPCR)

Blood samples were taken from 50 cases each of LUAD patients and healthy individuals, as well as from 46 cancerous tissue and 46 paraneoplastic tissue, respectively. All experiments involving human tissues were approved by the Ethics Committee of Tangshan People's Hospital (Approval No.: RMYY-LLKS-2021-103) and performed in accordance with the Declaration of Helsinki. All methods were carried out with approved protocols, relevant guidelines and regulations. Informed consents were obtained from all participants in the current study. RNA was extracted from cell lines, cell lines after EGLN3 knockdown, plasma samples and paraffin samples using Universal Total RNA Kit (R013, GeneBetter, Beijing), whole blood total RNA extraction kit (R213, GeneBetter, China) and FlashPure FFPETotal RNAMini Kit (R310, GeneBetter, Beijing) according to the instructions of

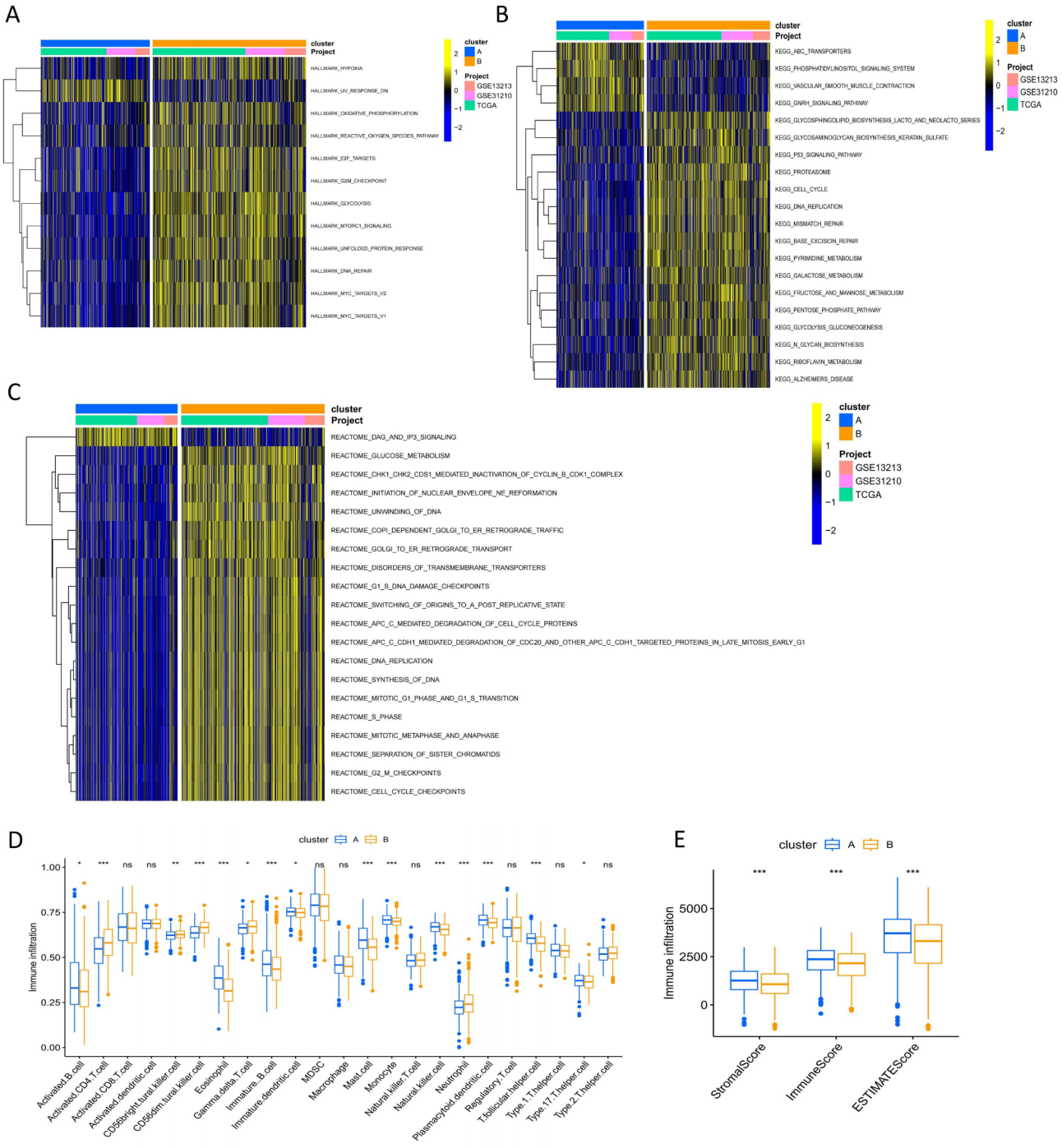


Fig. 3. Functional explorations in sialylation-related subgroups. GSEA analyses were performed to identify Hallmarks (**A**), KEGG pathways (**B**) and Reactome pathways (**C**) enriched in the sialylation-related subgroups. (**D**) The abundances of 23 immune cells were compared between sialylation-related subgroups. (**E**) Differences in immune, stromal, and estimate scores between sialylation-related subgroups. $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$, $^{ns}P>0.05$.

manufacturers. PrimeScript™ RT Master Mix (RR0036A, TARKRA, Beijing) was applied to configure the reverse transcription reaction system. The cDNA was gathered through executing reverse transcription at 37 °C for 15 min, 85 °C for 5 s, and cycling at 4 °C. Finally, RT-qPCR reactions were completed using TB Green® Premix Ex Taq™ II (RR820A, TARKRA, Beijing) on a QuantStudio 3-7500 equipment, and relative gene expression was determined using the $2^{-\Delta\Delta Ct}$ method. Table S1 displays primer information.

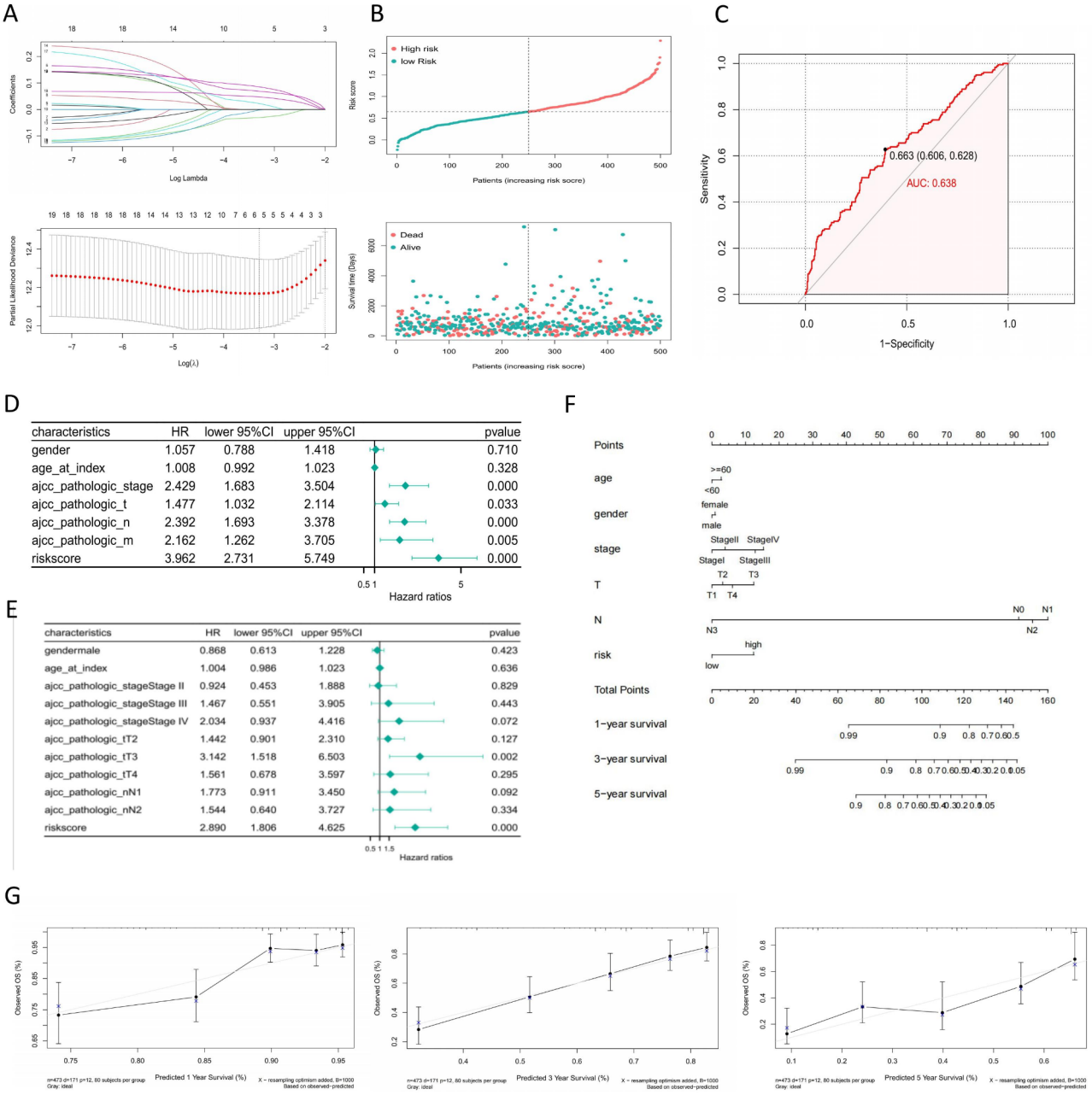
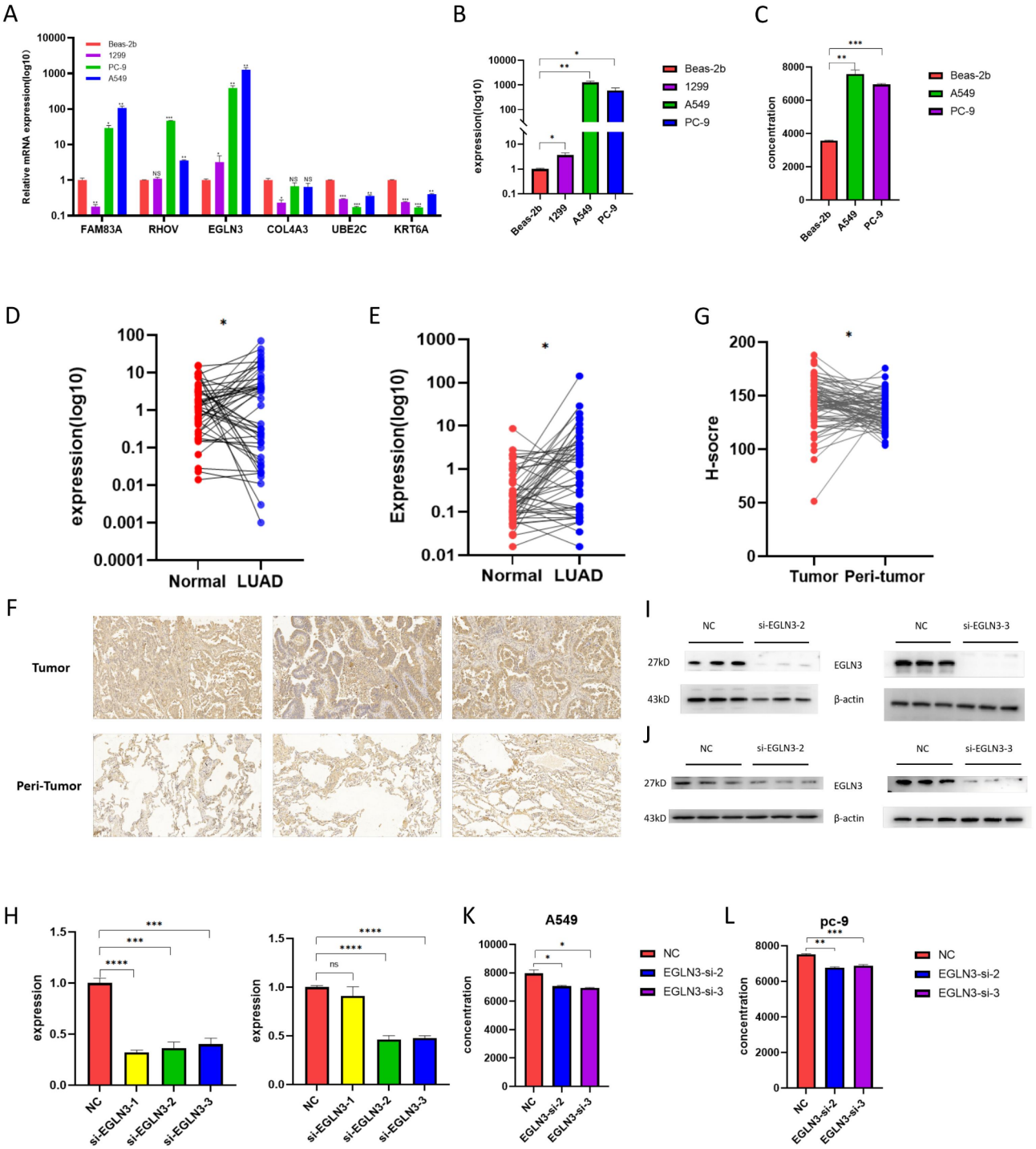


Fig. 4. Construction of the LUAD prognostic model. (**A**) LASSO regression analysis was performed to determine prognostic signatures. (**B**) LUAD patients were divided into low- and high-risk score groups according to the median value. The survival status distribution of LUAD patients in low- and high-risk score groups were shown in the dot plot. (**C**) The receiver operating characteristic (ROC) curve for the risk model. D-E. Univariate (**D**) and multivariate (**E**) Cox regression analyses were conducted to identify independent prognostic factors. (**F**). The nomogram was built based on independent prognostic factors to predict 1, 3, and 5-year survival of LUAD patients. (**G**) Calibration curves of the nomogram for predicting 1, 3, and 5-year survival of LUAD patients.

Immunohistochemical staining and clinical relevance

Paraffin specimens from 71 cancerous tissues and 71 paraneoplastic tissues were collected independently, and EGLN3 was detected with a rabbit two-step assay kit (PV-6001, ZSGB-BIO, China). Briefly, paraffin slices were dewaxed and hydrated before being repaired for antigens and blocked for endogenous peroxidase. Then, primary (Dilution concentration 1:800) and secondary antibodies were applied in the order specified, followed by color development and sealing. After panoramic scanning with digital sections, the number of positive cells was counted, and H-scores were determined. To investigate the relationship between H-scores and clinical characteristics of LUAD, the LUAD sample was separated into high ($n = 36$) and low ($n = 35$) scoring groups with



respect to H-score. Then, the clinical characteristics of LUAD patients were compared between groups, including age, gender, tumor diameter, pathological grading, clinical stage, and lymph node metastasis ($P < 0.05$).

Examination of Sialic acid

Sialic acid was tested in culture supernatants before and after EGLN3 knockdown using an ELISA kit (YPJ1802, UpingBio, Hangzhou). In brief, the pre-coated plate was filled with 50 µl of liquid for each of the following: blank wells (without any reagents or samples), standard wells ($n = 2$), quality control wells, and sample wells. Upon a 60 min incubation period at 37 °C, 50 µl of biotinylated antigen was added to each well, and the plate was then cleaned. After 30 min of incubation at 37 °C, 50 µl of enzyme-labeled avidin was added to each well, and the plate was next cleaned. Once the color developer and termination solution had been added, the absorbance of each well was lastly measured at 450 nm using an Microplate reader. Plotting the calibration curve allowed us to determine the quantity of sialic acid was present in the samples.

◀**Fig. 5**. The expression and the effect of EGLN3 on sialic acid in LUAD. (**A**) RT-qPCR was conducted to detect the expressions of 6 prognostic gene signatures in LUAD cells (A549, PC-9, 1299) and Beas-2b control cells, respectively. (**B**) RT-qPCR was utilized to detect EGLN3 expression in LUAD cells (A549, PC-9, 1299) and Beas-2b control cells, respectively. (**C**) The levels of sialic acid in the supernatants of LUAD cells (A549, PC-9) and Beas-2b control cells were detected using ELISA. (**D**) The expression levels of EGLN3 in LUAD cancer tissues ($N=46$) and adjacent normal tissues ($N=46$) were detected by RT-qPCR. (**E**) The expression levels of EGLN3 in blood samples from LUAD patients ($N=50$) and healthy controls ($N=50$) were detected by RT-qPCR. (**F**) Representative immunohistochemical images showing the expressions of EGLN3 in LUAD cancer tissues ($N=71$) and adjacent normal tissues ($N=71$). (**G**) The quantification of immunohistochemical results showing the elevated expression of EGLN3 in LUAD cancer tissues ($N=71$) compared with adjacent normal tissues ($N=71$). (**H**) Three small interfering RNA (si-EGLN3-1, si-EGLN3-2, si-EGLN3-3) used for EGLN3 knock-down and one control siRNA (NC) were transfected into A549 and PC-9 cells. RT-qPCR was performed to detect the EGLN3 knock-down efficiency in A549 (left) and PC-9 (right) cells. (**I**) Western blot was performed to detect the expressions of EGLN3 in A549 cells after transfecting si-EGLN3-2 or si-EGLN3-3. (**J**) Western blot was performed to detect the expressions of EGLN3 in PC-9 cells after transfecting si-EGLN3-2 or si-EGLN3-3. (**K**) The level of sialic acid determined by ELISA after EGLN3 knock-down in A549 cells. (**L**) The level of sialic acid determined by ELISA after EGLN3 knock-down in PC-9 cells. Data are presented as mean ± S.D. $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$, $^{****}P<0.001$. All the cellular experiments were performed three times independently.

Cell transfection

The A549 and PC-9 cell lines were seeded onto 6-well plates ($2*10^5$/well) and cultivated for 18–24 h, which were transfected after they achieved 60–80% cell density. Two hours before transfection, the medium was replaced with fresh complete media. Afterwards, siRNAs targeting EGLN3 as well as negative control (NC) siRNAs was transfected into cells using NEOfect and OPTI-MEM culture media according to the instructions of NEOfect (TF20141402, Beijing). Cells and cell supernatants were collected for functional and protein tests after 24 and 48 h of cultivation, respectively.

Western blot for protein measurement

After cell transfection or drug intervention, $1*10^6$ cells from NC, si-EGLN3-2, si-EGLN3-3, PC-9, A549, PC-9-Troxerutin and A549-Troxerutin groups were extracted using lysate (P0013, Beyotime, Shanghai) correspondingly, followed by protein electrophoresis. Subsequently, EGLN3 primary antibody (27875-1-AP, proteintech, Wuhan) (Dilution 1:1000), β-actin primary antibody (66009-1-Ig, proteintech, Wuhan) (Dilution 1:10,000), rabbit secondary antibody (074-1806, KPL, America) (Dilution 1:10000), and mouse secondary antibody (074-1506, KPL, America) (Dilution 1:10,000) were employed. Finally, an ECL luminescent solution (P1050, Applygen Technologies Inc, America) was applied for imaging.

Cell proliferation

In 96-well plates, the NC, si-EGLN3-2 and si-EGLN3-3 groups were labeled and cells were injected in the wells ($n=3$, $2*10^3$/well). After the first, second, third, and fourth days, the media was removed and replaced with the produced CCK-8 cell culture medium, which was incubated for two hours. Later, the absorbance was measured at 450 nm. The PC-9, A549, PC-9-Troxerutin and A549-Troxerutin group (Troxerutin intervention for 48 h) were both detected in the same manner.

Cell scratching

In 6-well plates, the NC, si-EGLN3-2, and si-EGLN3-3 groups were labeled, and the matching cells were injected ($n=3$, $3*10^5$/well). When the cell density reached 80-90%, a vertical line was painted on the bottom of the well plate with a pistol tip. After being washed twice with PBS, the medium was replaced with serum-free 1% PS to continue the culture, and photographs were taken at 0 and 24 h. The PC-9, A549, PC-9-Troxerutin and A549-Troxerutin groups were tested the same way.

Cell invasion

The matrix gel was mixed with Opti-MEM in a 1:4 ratio and added to the Transwell's upper chamber. Following two hours of incubation, 200ul and 600ul of cell suspension from the NC, si-EGLN3-2, and si-EGLN3-3 groups, as well as PC-9, A549, PC-9-Troxerutin and A549-Troxerutin groups, were introduced to the upper and lower chambers, respectively (20% complete medium in the lower chamber). After 24 h, the upper chambers were preserved in methanol and stained with 0.1% crystalline violet for 10 min. Finally, the entire filter membrane was removed, the sealed film was photographed, and the results were calculated and analyzed with Image J software.

Cell clone formation

Cells were collected after transfection or before and after Troxerutin intervention and inoculated in six-well plates ($2*10^3$/well). Around days 7–15, the cells were fixed with methanol, stained with 1% crystal violet, and photographed.

Cell cycle and apoptosis

The cells from the PC-9, A549, PC-9-Troxerutin, and A549-Troxerutin groups were separated, centrifuged at 300 g for 5 min, and then twice cleaned with PBS. The cell cycle test was then performed according the

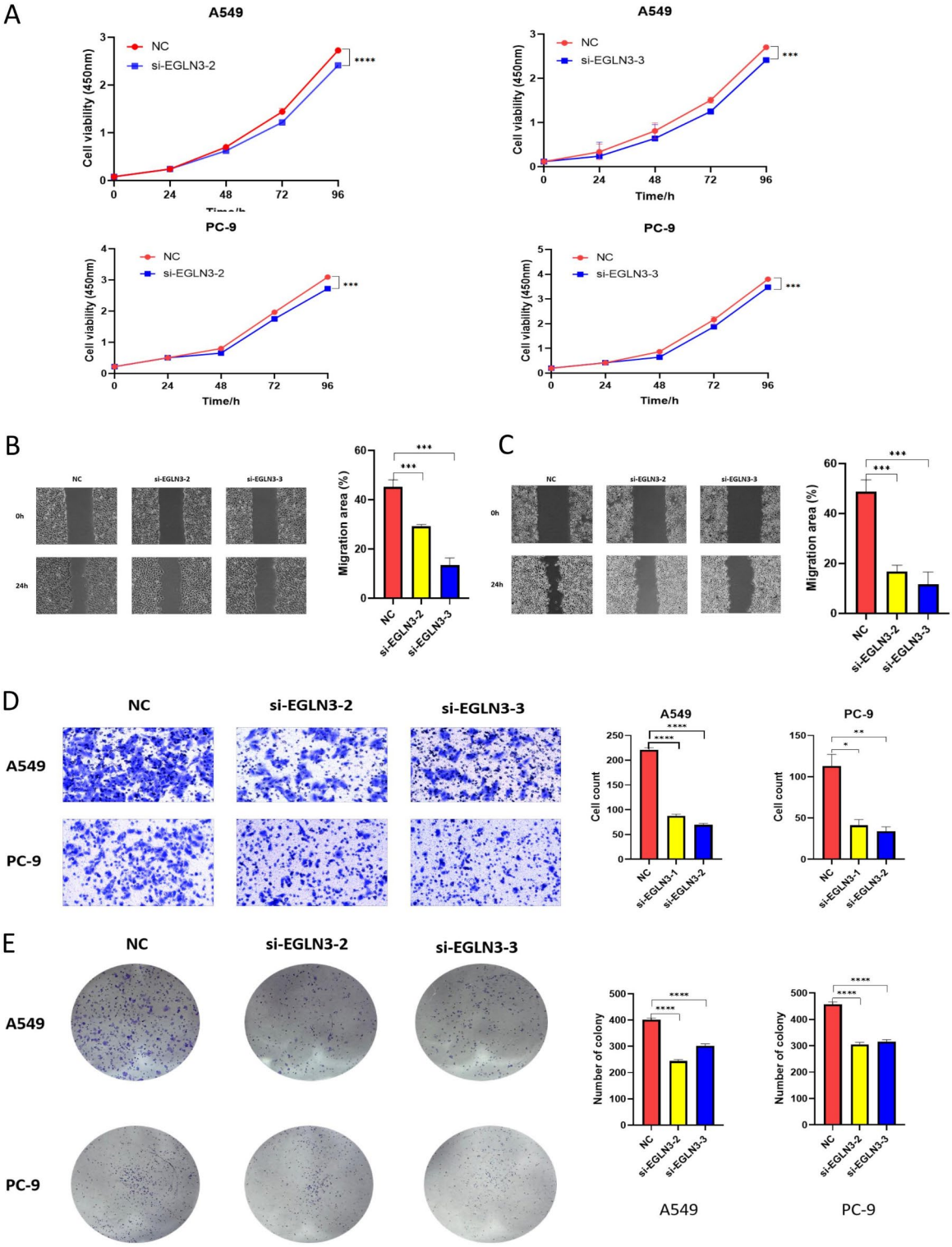


Fig. 6. The effect of EGLN3 knock-down on the behaviors of LUAD cells. (**A**) Knock-down of EGLN3 significantly inhibited the proliferation of A549 and PC-9 cells. (**B**, **C**) Knock-down of EGLN3 significantly inhibited the migration ability of A549 (**B**) and PC-9 cells (**C**). (**D**) Transwell experiment revealed that knock-down of EGLN3 significantly inhibited the invasive ability of A549 and PC-9 cells. (**E**) Knock-down of EGLN3 significantly inhibited the clone formation ability of A549 and PC-9 cells. Data are presented as mean ± S.D. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.001$. All the cellular experiments were performed three times independently.

manufacturer's instructions (Cell Cycle and Apoptosis Detection Kit (C1052, Biyun Tian). In brief, 70% ethanol was applied to each group of cell precipitation at 4 °C for 2 h of fixation, followed by centrifugation for 5 min at 300 g to remove the supernatant, and the process was repeated. Later on, each batch of cells received 0.5 ml of propidium iodide staining solution, which was incubated at 37℃ for 30 min under light protection before testing. Apoptosis was then detected, the Annexin V-FITC binding solution and propidium iodide staining solution were then added to each batch of samples in accordance with the instructions for the Annexin V-FITC/PI Apoptosis Kit (C1062L, Biyun Tian). The samples were detected by flow cytometry (Attune™ NxT, Thermo Fisher Scientific) after being incubated at room temperature for 10–20 min and kept out of the light.

Statistical analysis

All statistical analyses were executed by R software. The IC50 value is the concentration of Troxerutin, which leads to a 50% reduction in cell viability of LUAD cells. Calculations of the IC_{50}, statistical analysis, and data plotting were performed using Graphpad Prism software. A p-value < 0.05 was set to statistical significance.

Results

The survival and expression profiles were different between sialylation-related subgroups

The flowchart of the study is shown in Figure S1. To gain more precise results, we merged three LUAD datasets into the LUAD ensemble dataset, which consisted of 867 samples (Figures S2). Univariate cox regression analysis revealed that 26 SRGs in LUAD ensemble dataset were associated with LUAD prognosis (Fig. 1A). Among them, the expressions of 12 SRGs (e.g., GALNT14, B3GNT3, and ITGB8) were significantly elevated, while 6 SPRs (e.g., ST6GALNAC6, SIGLEC6, and SELP) were reduced in LUAD (Fig. 1B). Further anlaysis showed that SELP had the highest SNV mutation frequency, followed by GALNT14 (Fig. 1C), and the most common mutation type was missense mutation among the top 10 mutated SRGs (Fig. 1D) in LUAD.

To investigate the relationship between LUAD and 26 SRGs, the LUAD ensemble dataset was separated into two sialylation-related subgroups, Cluster A and B based on the expressions of 26 SRGs (Fig. 2A). The expressions of 24 SRGs were significantly different between two subgroups (Fig. 2B). Further survival analysis demonstrated that Cluster A had a greater survival rate than Cluster B (Fig. 2C), yet the clinical characteristics of patients in the two groups were not substantially different (Figure S3). Furthermore3, a total of 77 DEGs between Cluster A and B were identified (Fig. 2D), which were primarily involved in the IL-17 signaling pathway and the complement and coagulation cascades pathways (Fig. 2E). Among these DEGs, 21 DEGs were identified as being associated with LUAD prognosis through univariate Cox analysis (Fig. 2F).

The sialylation-related subgourps had different tumor microenvironments

To discover more about the biological processes involved in sialylation-related subgroups, functional enrichment of HALLMARK, KEGG, and Reactome pathways was conducted. The findings suggested that Cluster A was primarily involved in the phosphatidylinositol signaling system, vascular smooth muscle contraction, and GnRH signaling pathway, whereas Cluster B was mainly involved in glucose metabolism, pentose phosphate pathway, and pyrimidine metabolism (Fig. 3A,C).

Immune infiltration indicated that mast cells, follicular helper T cells, eosinophils, and type 17 T helper cells were more abundant in Cluster A, while neutrophils and activated CD4 T cells were higher in Cluster B (Fig. 3D). Meanwhile, the stromal, immune, and ESTIMATE scores were remarkably higher in Cluster A, implying that LUAD patients in Cluster A might have higher levels of stromal and immune cell infiltration, promoting anti-tumor actions (Fig. 3E).

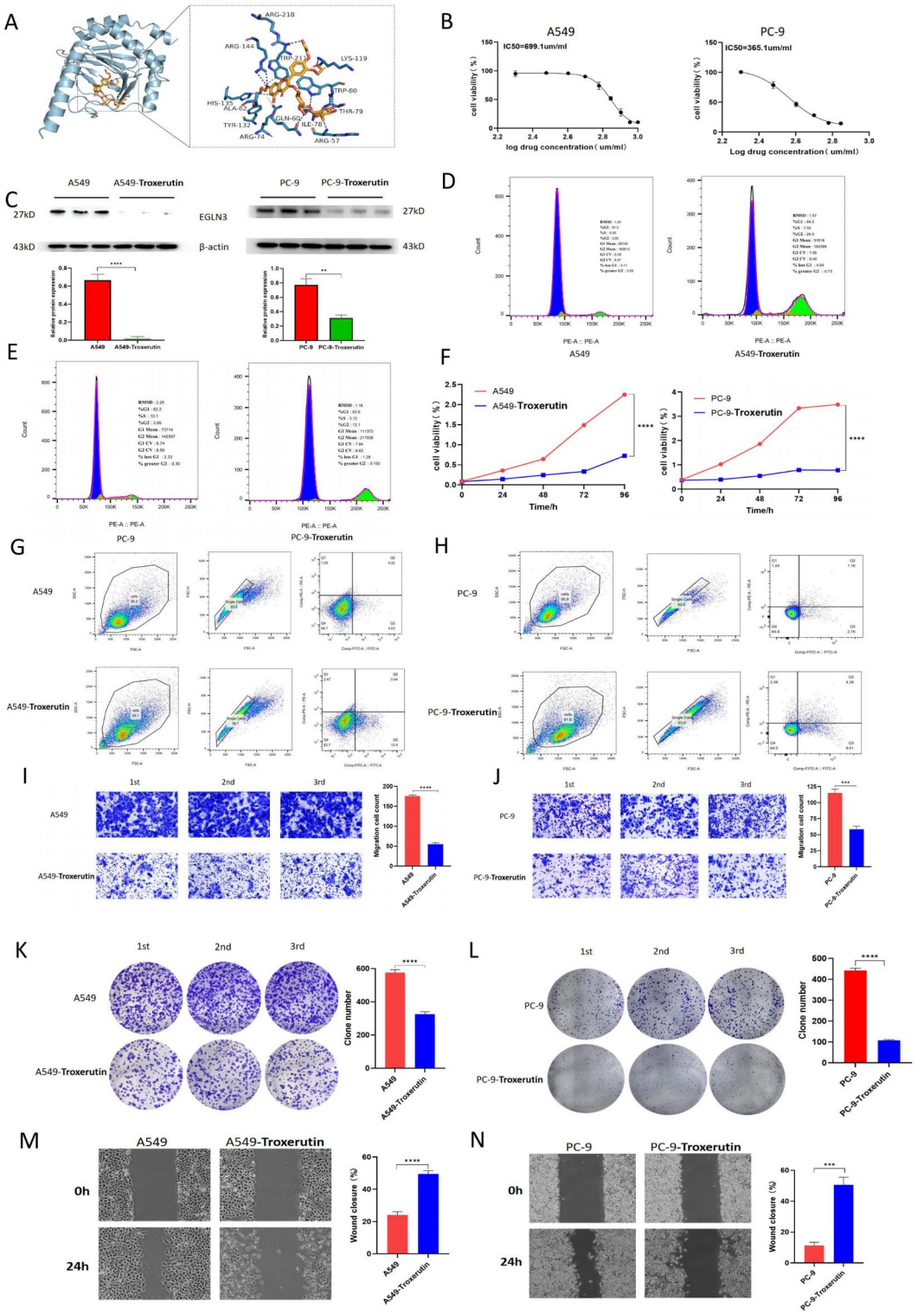
The model constructed with prognostic SRG signature performed effectively in predicting the prognosis of LUAD

To build a risk prognostic model for LUAD, 6 prognostic SRG signature were determined using LASSO regression analysis (Fig. 4A). The prognostic SRG signature and risk coefficients were next utilized to create risk scores, and patients with LUAD were divided into high- and low-risk groups according to the median risk score (Fig. 4B). Moreover, the risk curve indicated that an elevated risk score corresponded to a greater number of dead patients (Fig. 4B), and the AUC value under ROC curve exceeded 0.6, showing that the risk model performed well (Fig. 4C). The inclusion of clinical characteristics and risk score in univariate and multivariate cox analysis revealed that T3 and risk score were independent prognostic factors (Fig. 4D,E). The nomogram was then developed to predict LUAD patients' survival at 1, 3, and 5 years (Fig. 4F). The slope of the calibration curve was close to 1, suggesting that the nomogram model was accurate (Fig. 4G).

Knock-down of EGLN3 inhibited the proliferation, migration and invision of LUAD cells

By determining the expressions of 6 prognostic SRG signature, we found that EGLN3 expression was higher than the other 5 SRGs and was significantly elevated in LUAD cells (Fig. 5A,B). ELISA results confirmed that the level of sialic acid was also increased in LUAD cells (Fig. 5C). Consistently, EGLN3 expression was significantly higher in cancer tissues and blood from LUAD patients than in paraneoplastic tissues and healthy blood (Fig. 5D,E). Meanwhile, immunohistochemical results further confirmed that the protein level of EGLN3 was much higher in cancer tissues compared with adjacent normal tissues (Fig. 5F,G). Furthermore, we divided LUAD patients into high- and low-EGLN3 expression groups and found significant differences in the distribution of pathological grading, clinical stage, and lymph node metastases between two groups (Table S2). These findings indicate that EGLN3 was an important player in LUAD.

To further study EGLN3's effect on LUAD, knock-down experiments were performed in A549 and PC-9 cells using siRNA technology. RT-qPCR results proved that si-EGLN3-2 and si-EGLN3-3 were more effectively in silencing EGLN3 (Fig. 5H), and western blotting yielded similar findings (Fig. 5I,J, Figure S4A-S4H).



ELISA results indicated that knockdown of EGLN3 also decreased the level of sialic in LUAD cells (Fig. 5K,L). Furthermore, the proliferative ability of A549 and PC-9 cells was considerably inhibited in si-EGLN3-2 and si-EGLN3-3 compared to NC group (Fig. 6A). Meanwhile, EGLN3 knock-down greatly reduced A549 and PC-9 cells' ability to migrate, invade, and clone formation (Fig. 6B,E).

The development of LUAD could be restricted by Troxerutin's ability to target and inhibit EGLN3 expression.

Troxerutin, a medication targeting EGLN3 for the treatment of LUAD, was identified with a binding energy of -9.439 KJ/mol (Fig. 7A). Meanwhile, Troxerutin has been shown in the literature to have considerable therapeutic efficacy in breast and stomach malignancies, as well as in animal non-small cell lung cancer[26–28]. Subsequently, the IC50 value of Troxerutin in LUAD cells were determined. We found that the IC50 of A549

◀**Fig. 7**. The effect of Troxerutin on EGLN3 expression and LUAD cell behaviors. (**A**) Cartoon representation showing the structure of interaction between EGLN3 (blue) and Troxerutin (orange). The hydrogen-bonding interactions are shown as dashed black lines. (**B**) The viability of A549 (left) and PC-9 (right) cells treated with varied concentrations of Troxerutin were determined by CCK-8 to determine the IC50 value of Troxerutin. (**C**) The effect of Troxerutin intervention on EGLN3 expression are examined by western blot in A549 (left) and PC-9 (right) cells, respectively. (D-E) Flowcytometry analysis of the cell cycle phase distribution in A549 (**D**) and PC-9 (**E**) cells after treatment with Troxerutin or vehicle control. (**F**) The cell viability analysis by CCK-8 in A549 (left) and PC-9 (right) cells after treatment with Troxerutin or vehicle control. (G-H) Flow cytometry was conducted to detect apoptosis after Troxerutin intervention in A549 (**G**) and PC-9 cells (**H**), respectively. (**I**, **J**) Troxerutin treatment significantly inhibited the invasive ability of A549 (**I**) and PC-9 cells (**J**). (**K**, **L**) Troxerutin treatment significantly inhibited the clone formation of A549 (**K**) and PC-9 cells (**L**). (**M**, **N**) Troxerutin treatment significantly inhibited the migration of A549 (**M**) and PC-9 (**N**) cells. Data are presented as mean ± S.D. $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.001$. All the cellular experiments were performed three times independently.

and PC-9 are 699.1 μm/ml and 365.1 μm/ml, respectively, indicating that its inhibitory effect on PC-9 cells was stronger (Fig. 7B). LUAD cells treated with Troxerutin had lower EGLN3 protein levels than the control group, indicating that Troxerutin might target and decrease EGLN3 expression in LUAD (Fig. 7C, Figure S3I-S3L). Furthermore, the G2 phase ratio (Fig. 7D,E) and viability (Fig. 7F) of LUAD cells was dramatically reduced following Troxerutin treatment, indicating that Troxerutin could inhibit cell proliferation. Additionally, we observed a large rise in apoptotic cells (Fig. 7G,H), as well as a significant reduction in LUAD cells' ability to invade, migrate, and clone formation (Fig. 7I,N) in Troxerutin treatment groups.

Discussion

LUAD is a rather frequent malignancy with a dismal prognosis. Sialylation-related proteins have an important role in the development of LUAD, however the specific process is unknown. The current study's findings revealed that among sialylation-related subgroups determined by SRGs connected with LUAD prognosis, LUAD patients in group A had a better survival rate, demonstrating that SRGs had a strong connection with LUAD prognosis. Meanwhile, the expression of several immune cells differed dramatically between groups. RHOV, FAM83A, EGLN3, COL4A3, UBE2C, and KRT6A were recognized as LUAD-related prognostic genes, and a risk prognostic model was developed to reliably predict the survival status of LUAD patients.

The current work detected 6 prognostic genes (RHOV, FAM83A, EGLN3, COL4A3, UBE2C, and KRT6A) related to sialylation leveraging the TCGA-LUAD dataset. RHOV has been shown to enhance neural crest development during vertebrate embryogenesis, as well as LUAD cell proliferation and metastasis via the JNK/c-Jun pathway[29]. RhoA and RhoC, members of the Rho family, have been linked to lung cancer development, progression, and migration[30]. Besides, FAM83A may enhance the prognosis of LUAD by controlling B-cell and dendritic cell infiltration[31]. Meanwhile, Dong et al. developed a predictive model for LUAD employing five genes, including COL4A3, that accurately predicts the prognosis of LUAD patients[32]. Cai and Sun et al. have also confirmed UBE2C and KRT6A as LUAD prognostic biomarkers and a new target for cancer immunotherapy[33,34]. These findings were extremely compatible with the findings of the current investigation, implying that sialylation-related prognostic genes RHOV, FAM83A, SFTPB, COL4A3, UBE2C, and KRT6A might serve as prospective biomarkers and prognostic targets for LUAD in the future.

EGLN3, also known as PHD3, is a member of the EGL-9 (EGLN) family of cryptic nematode genes in Hidradenitis elegans. It regulates cell signaling, cell metabolism, cell cycle, apoptosis, and cancer cell migration and is essential for the growth of a variety of cancers, including lung cancer[35,36], which is highly consistent with our findings. It has been established that EGLN3 is involved in the control of hypoxia-inducible factors, which are substantially expressed in lung squamous cell carcinoma[37]. Another study determined that reducing host EGLN3 hydroxylase activity reduces tumor growth in lung large cell carcinoma by improving anti-tumor immunity and lowering angiogenesis[35]. Our investigations proved that EGLN3 was highly overexpressed in cancer tissues, blood from LUAD patients, and the A549, PC-9, and 1299 cell lines. A549 and PC-9 cells' proliferation, migration, invasion, and clonogenic ability were all reduced following siRNA knockdown.

Troxerutin is a semi-synthetic flavonoid produced from rutin (rutin) that has antioxidant, anti-inflammatory, hypolipidemic, and nephroprotective effects[38]. Troxerutin has been demonstrated in studies to lower the expression of proliferating cell nuclear antigen, cell cycle protein D1, P13k, Akt, and mTOR. Furthermore, Troxerutin inhibits the inflammatory response of lung cancer cells[27] to combat cancer. Moreover, Troxerutin suppresses STAT3/NF-κB and Bcl-2 signaling pathways, resulting in therapeutic effects in gastric cancer[39]. The current study observed that Troxerutin dramatically reduced the proliferation, migration, invasion, and clonogenic potential of A549 and PC-9 cells, while down-regulating EGLN3 expression. Interestingly, the sialic acid content in the culture supernatants of A549 and PC-9 cells was considerably reduced following EGLN3 knockdown and Troxerutin treatment, suggesting the crucial regulatory role of sialylation in LUAD. As a result, we anticipated that Troxerutin would inhibit the development of LUAD by down-regulating EGLN3 expression in A549 and PC-9 cells, thereby interfering with sialic acid synthesis as well as their ability to proliferate, migrate, and invade.

There are some limitations in our study. Although we identified key prognostic SRGs and build the prognostic model for LUAD patients, the accuracy and utility of the model need to be validated and evaluated in the real world by collecting large LUAD samples, as well as their clinical information. Secondly, we analyzed the function of EGLN3 by performing relevant experiments in the vitro system. However, the in vitro system cannot fully

mimic the complex in vivo environment, and the LUAD cell lines cannot replace the primary LUAD cells from patients. Thus, using patient-derived cells or organoids and performing EGLN3 knockout experiments in vivo using the mouse model is essential to enhance the clinical relevance of the study. Likewise, we evaluated the Troxerutin intervention in the LUAD cells, however, Troxerutin may have off-targets effects, such as unintended modulation of normal tissues, thus affecting treatment efficacy or safety. We need to carry out more cautious in vitro and in vivo experiments in the future, such as pharmacokinetic and pharmacodynamic studies and drug delivery systems to optimize its distribution and to minimize the risk of off-target effects.

In general, six sialylation-related prognostic genes, RHOV, FAM83A, SFTPB, COL4A3, UBE2C, and KRT6A, were located with bioinformatics methods, and a prognostic model was developed to effectively predict the risk of LUAD patients. Meanwhile, EGLN3 was confirmed to be a sialylation-associated biomarker via multiple experimental methods, and Troxerutin intervention could down-regulate EGLN3 expression and sialic acid production while inhibiting lung cancer cells' ability to proliferate, migrate, and invade. This finding will provide new insights and references for future diagnosis and focused treatment of LUAD patients.

Data availability

Any data and R scripts in this study can be obtained from the corresponding author upon reasonable request.

Received: 8 October 2024; Accepted: 25 February 2025
Published online: 18 March 2025

References

1. Shu, J., Jiang, J. & Zhao, G. Identification of novel gene signature for lung adenocarcinoma by machine learning to predict immunotherapy and prognosis. *Front. Immunol.* **14**, 1177847 (2023).
2. Wei, X., Li, X., Hu, S., Cheng, J. & Cai, R. Regulation of ferroptosis in lung adenocarcinoma. *Int. J. Mol. Sci.* **24**(19), 14614 (2023).
3. Zhao, C., Xiao, R., Jin, H. & Li, X. The immune microenvironment of lung adenocarcinoma featured with ground-glass nodules. *Thorac. Cancer* **15**(19), 1459–1470 (2024).
4. Puri, S., Shafique, M. & Gray, J. E. Immune checkpoint inhibitors in early-stage and locally advanced non-small cell lung cancer. *Curr. Treat. Options Oncol.* **19**(8), 39 (2018).
5. Song, Y., Kelava, L. & Kiss, I. MiRNAs in lung adenocarcinoma: Role, diagnosis, prognosis, and therapy. *Int. J. Mol. Sci.* **24**(17), 13302 (2023).
6. Jones, D. R., Wu, Y. L., Tsuboi, M. & Herbst, R. S. Targeted therapies for resectable lung adenocarcinoma: ADAURA opens for thoracic oncologic surgeons. *J. Thorac. Cardiovasc. Surg.* **162**(1), 288–292 (2021).
7. Song, P. et al. Integrated analysis of single-cell and bulk RNA-sequencing identifies a signature based on B cell marker genes to predict prognosis and immunotherapy response in lung adenocarcinoma. *Cancer Immunol. Immunother.* **71**(10), 2341–2354 (2022).
8. Wang, B., Hou, C., Yu, X., Liu, J. & Wang, J. The prognostic value of sialylation-related long non-coding RNAs in lung adenocarcinoma. *Sci. Rep.* **14**(1), 8879 (2024).
9. Ogawa, J., Sano, A., Koide, S. & Shohtsu, A. Relation between recurrence and expression of proliferating cell nuclear antigen, sialyl LewisX, and sialyl Lewis(a) in lung cancer. *J. Thorac. Cardiovasc. Surg.* **108**(2), 329–336 (1994).
10. Satoh, H., Ishikawa, H., Yamashita, Y. T., Ohtsuka, M. & Sekizawa, K. Serum sialyl Lewis X-i antigen in lung adenocarcinoma and idiopathic pulmonary fibrosis. *Thorax* **57**(3), 263–266 (2002).
11. Zhou, M. et al. Characterization of sialylation-related long noncoding RNAs to develop a novel signature for predicting prognosis, immune landscape, and chemotherapy response in colorectal cancer. *Front. Immunol.* **13**, 994874 (2022).
12. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**(7), e47 (2015).
13. Kuncman, Ł et al. The kinetics of FMS-related tyrosine kinase 3 ligand (Flt-3L) during chemoradiotherapy suggests a potential gain from the earlier initiation of immunotherapy. *Cancers (Basel)* **14**(16), 3844 (2022).
14. Ramsay, I. S. et al. Model selection and prediction of outcomes in recent onset schizophrenia patients who undergo cognitive training. *Schizophr. Res. Cogn.* **11**, 1–5 (2018).
15. Shi, Y. et al. Crosstalk of ferroptosis regulators and tumor immunity in pancreatic adenocarcinoma: novel perspective to mRNA vaccines and personalized immunotherapy. *Apoptosis* **28**(9–10), 1423–1435 (2023).
16. Gustavsson, E. K., Zhang, D., Reynolds, R. H., Garcia-Ruiz, S. & Ryten, M. ggtranscript: An R package for the visualization and interpretation of transcript isoforms using ggplot2. *Bioinformatics* **38**(15), 3844–3846 (2022).
17. Wilkerson, M. D. & Hayes, D. N. ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26**(12), 1572–1573 (2010).
18. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *Omics* **16**(5), 284–287 (2012).
19. Wang, B. et al. Identification and validation of chromatin regulator-related signatures as a novel prognostic model for low-grade gliomas using translational bioinformatics. *Life Sci.* **336**, 122312 (2024).
20. Wu, F. et al. Molecular subtyping reveals immune alterations in IDH wild-type lower-grade diffuse glioma. *J. Pathol.* **251**(3), 272–283 (2020).
21. Li, Y., Lu, F. & Yin, Y. Applying logistic LASSO regression for the diagnosis of atypical Crohn's disease. *Sci. Rep.* **12**(1), 11340 (2022).
22. Cheng, Q., Chen, X., Wu, H. & Du, Y. Three hematologic/immune system-specific expressed genes are considered as the potential biomarkers for the diagnosis of early rheumatoid arthritis through bioinformatics analysis. *J. Transl. Med.* **19**(1), 18 (2021).
23. Robin, X. et al. pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* **12**, 77 (2011).
24. Sheikh, I. A., Jiffri, E. H., Ashraf, G. M., Kamal, M. A. & Beg, M. A. Structural studies on inhibitory mechanisms of antibiotic, corticosteroid and catecholamine molecules on lactoperoxidase. *Life Sci.* **207**, 412–419 (2018).
25. Rolta, R. et al. In silico screening of hundred phytocompounds of ten medicinal plants as potential inhibitors of nucleocapsid phosphoprotein of COVID-19: An approach to prevent virus assembly. *J. Biomol. Struct. Dyn.* **39**(18), 7017–7034 (2021).
26. Saranya, T. et al. Enhanced apoptogenesis and oncogene regulatory mechanism of troxerutin in triple negative breast cancer cells. *Toxicol. Res. (Camb).* **9**(3), 230–238 (2020).
27. Yu, J. et al. Anticancer effect of troxerutin in human non-small-cell lung cancer cell A549 and inhibition of tumor formation in BALB/c nude mice. *J. Environ. Pathol. Toxicol. Oncol.* **40**(3), 25–35 (2021).
28. Li, Y., Gu, J. & Yu, Q. Efficacy of mecobalamin tablets combined with troxerutin in the treatment of NSCLC chemotherapy-induced peripheral neuropathy. *Evid. Based Complement Alternat. Med.* **2022**, 7946934 (2022).
29. Zhang, D. et al. RHOV promotes lung adenocarcinoma cell growth and metastasis through JNK/c-Jun pathway. *Int. J. Biol. Sci.* **17**(10), 2622–2632 (2021).

30. Richter, L., Oberländer, V. & Schmidt, G. RhoA/C inhibits proliferation by inducing the synthesis of GPRC5A. *Sci. Rep.* **10**(1), 12532 (2020).
31. Zhang, M. et al. AURKA and FAM83A are prognostic biomarkers and correlated with tumor-infiltrating lymphocytes in smoking related lung adenocarcinoma. *J. Cancer* **12**(6), 1742–1754 (2021).
32. Dong, L. et al. A five-collagen-based risk model in lung adenocarcinoma: prognostic significance and immune landscape. *Front. Oncol.* **13**, 1180723 (2023).
33. Cai, D., Tian, F., Wu, M., Tu, J. & Wang, Y. UBE2C is a diagnosis and therapeutic biomarker involved in immune infiltration of cancers including lung adenocarcinoma. *J. Cancer* **15**(6), 1701–1717 (2024).
34. Sun, K. et al. Regulation of early diagnosis and prognostic markers of lung adenocarcinoma in immunity and hypoxia. *Sci. Rep.* **13**(1), 6459 (2023).
35. Jin, Y. et al. Inactivation of EGLN3 hydroxylase facilitates Erk3 degradation via autophagy and impedes lung cancer growth. *Oncogene* **41**(12), 1752–1766 (2022).
36. German, N. J. et al. PHD3 loss in cancer enables metabolic reliance on fatty acid oxidation via deactivation of ACC2. *Mol. Cell* **63**(6), 1006–1020 (2016).
37. Amatschek, S. et al. Tissue-wide expression profiling using cDNA subtraction and microarrays to identify tumor-specific genes. *Cancer Res.* **64**(3), 844–856 (2004).
38. Arab, H. H. et al. Targeting inflammation, autophagy, and apoptosis by troxerutin attenuates methotrexate-induced renal injury in rats. *Int. Immunopharmacol.* **103**, 108284 (2022).
39. Xu, G. Y. & Tang, X. J. Troxerutin (TXN) potentiated 5-Fluorouracil (5-Fu) treatment of human gastric cancer through suppressing STAT3/NF-κB and Bcl-2 signaling pathways. *Biomed. Pharmacother.* **92**, 95–107 (2017).

Acknowledgements
We are grateful to the contributors to the public databases used in this study and all the authors of the study.

Author contributions
Author Contributions: Conceptualization, G.S.; methodology, Y.W. and J.G.; software, Y.W. and Y.J.; validation, Y.W. and Y.C.; formal analysis, Y.W.; investigation, X.Y.; resources, Z.X.; data curation, Y.G.; writing—original draft preparation, Y.W.; writing—review and editing, Y.W.; visualization, Y.W.; supervision, G.S.; project administration, G.S.; funding acquisition, G.S. All authors have read and agreed to the published version of the manuscript.

Funding
This research was funded by Natural Science Foundation of Hebei Province (Grant No.H2023209083); Hebei Province Innovation Capability Enhancement Plan Project (Grant No.235A2403D). Hebei Province Higher Education Scientific Research Special Task Project (Grant No.JZX2024009). Hebei Provincial Department of Education Hebei Experimental Teaching and Teaching Laboratory Construction Project (Grant No.81). High level Research and Innovation Team Construction Plan of School of Public Health, North China University of Science and Technology (Grant No.KYTD202309). Tangshan Municipal Science and Technology Plan Project (Grant No.24150218 C).

Declarations

Competing interests
The authors declare no competing interests.

Ethical approval
Tissues and blood samples used in the current study were obtained from Tangshan People's Hospital. Specifically, patients were admitted to the Pathology Department of Tangshan People's Hospital between December 2021 and December 2023. Patients diagnosed with LUAD were enrolled. All patients were diagnosed based on histological examination results. Patients with active systemic infections or autoimmune diseases were excluded. The cancer tissues and adjacent normal tissues used in the current study were then obtained during intraoperative. The paracancerous tissues, located more than 5.0 cm away from the tumor, were confirmed as normal controls through H&E stains. Healthy donors were recruited as healthy controls from the physical examination center of Tangshan People's Hospital. All experiments involving human tissues were approved by the Ethics Committee of Tangshan People's Hospital (Approval No.: RMYY-LLKS-2021-103, Approval date: 13 December 2021) and performed in accordance with the Declaration of Helsinki. All methods were carried out with approved protocols, relevant guidelines and regulations. Informed consents were obtained from all participants in the current study.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-92028-2.

Correspondence and requests for materials should be addressed to G.S.