Original Article

Multi-omics analysis identified serum *B4GALT1* as a prognostic factor for small cell lung cancer

Chenyang Wang[1#], Xiaoyue Zhu[1#], Yiheng Shi[2], Guirong Wang[2], Teng Ma[1]

[1]Cancer Research Center, Beijing Chest Hospital, Capital Medical University, Beijing Tuberculosis and Thoracic Tumor Research Institute, Beijing, China; [2]Department of Clinical Laboratory, Beijing Chest Hospital, Capital Medical University, Beijing Tuberculosis and Thoracic Tumor Research Institute, Beijing, China

Contributions: (I) Conception and design: T Ma; (II) Administrative support: T Ma; (III) Provision of study materials or patients: Y Shi, G Wang; (IV) Collection and assembly of data: C Wang; (V) Data analysis and interpretation: C Wang, X Zhu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Teng Ma, MD. Cancer Research Center, Beijing Chest Hospital, Capital Medical University, Beijing Tuberculosis and Thoracic Tumor Research Institute, No. 9 Courtyard, Beiguan Street, Tongzhou District, Beijing 101149, China. Email: Mateng82913@163.com.

Background: Small cell lung cancer (SCLC) is an aggressive neuroendocrine tumor characterized by rapid progression, early metastasis, and high mortality, with limited effective long-term treatment options. *B4GALT1*, a β-1,4-galactosyltransferase, has been implicated in the malignant progression of various cancers, but its specific role and underlying mechanisms in SCLC remain largely unexplored. We conducted a multi-omics analysis and clinical sample study to explore the function of *B4GALT1* in SCLC.

Methods: This study comprehensively investigated the expression pattern, functional significance, and clinical relevance of *B4GALT1* in SCLC. We conducted multi-omics analyses, including single-cell data processing, InferCNV analysis, and immune infiltration analysis, to explore the association between *B4GALT1* and the immune microenvironment of SCLC and patient survival. To determine *B4GALT1* as a potential circulating biomarker, quantitative data-independent acquisition (DIA) proteomics analysis was performed on serum samples from SCLC patients and healthy controls. Enzyme-linked immunosorbent assay (ELISA) was used to further verify the differential expression of serum *B4GALT1* in a larger cohort of SCLC patients, to evaluate its diagnostic, prognostic, and treatment response predictive value.

Results: Multi-omics analysis revealed that *B4GALT1* expression was significantly associated with patient survival. The expression of *B4GALT1* positively correlated with macrophage infiltration in the tumor and negatively correlated with $CD4^+$ T cells in the tumor. There was a negative correlation in inactivated naïve B cells, eosinophils, and CD4 naïve T cells, while it showed a positive correlation in dendritic cells, M0/M1/M2 macrophages, natural killer (NK) cells, CD8 T cells, follicular helper T cells, and regulatory T cells. ELISA results showed that serum protein *B4GALT1* expression was higher in patients with SCLC than in healthy controls. Elevated serum B4GALT1 protein levels correlated with poor treatment outcomes in patients with SCLC undergoing chemoradiotherapy.

Conclusions: Our findings establish *B4GALT1* as a critical prognostic, diagnostic, and predictive biomarker in SCLC, with its expression closely linked to the tumor immune microenvironment and treatment response. Targeting *B4GALT1* or its related pathways may represent a novel therapeutic strategy, and serum *B4GALT1* holds promise as a liquid biopsy marker for SCLC patient stratification, monitoring, and guiding treatment decisions.

Keywords: *B4GALT1*; small cell lung cancer (SCLC); liquid chromatography-tandem mass spectrometry (LC-MS/MS); biomarker; programmed death-1 (*PD-1*)

Submitted Dec 12, 2025. Accepted for publication Jan 28, 2026. Published online Mar 20, 2026.
doi: 10.21037/jtd-2025-1-2610
View this article at: https://dx.doi.org/10.21037/jtd-2025-1-2610

Introduction

Small cell lung cancer (SCLC) is the second most common malignant tumor of the chest, accounting for 10–15% of all lung cancers (1). It is characterized by its high invasiveness, rapid proliferation, and tendency to metastasize. One-third of patients present with limited-stage SCLC (LS-SCLC), whereas two-thirds present with extensive-stage SCLC (ES-SCLC) (2). Current treatments for LS-SCLC include chemotherapy, early intervention hyperfractionated thoracic radiotherapy, and prophylactic cranial irradiation (PCI) following chemotherapy (3). However, the optimal dose and schedule for thoracic radiotherapy remain unclear. The primary treatment for ES-SCLC is chemotherapy with or without PCI; however, the value of radiotherapy remains unclear and warrants further investigation. Moreover, SCLC exhibits an initial response to chemotherapy and radiotherapy; however, treatment resistance typically develops. Most patients relapse after first-line therapy, with only approximately 10% remaining disease-free after 2 years, resulting in a maximum 5-year overall survival rate of 10% (4). Currently, the primary challenge in small-cell lung cancer research is the scarcity of available tumor samples, and most existing studies have been conducted almost exclusively at the transcriptomic level. Proteins are the executors of cellular life activities, and secretion-mediated intercellular communication plays a pivotal role in regulating metabolic homeostasis and in the pathological progression of metabolic diseases (5,6). Thus, a liquid biopsy may be an effective tool.

Data independent acquisition (DIA) proteomics quantitative analysis employs a data-independent acquisition mode, utilizing liquid chromatography-mass spectrometry (LC-MS/MS) to analyze proteolytic peptides from proteins. Combined with a spectral library constructed through traditional data-dependent acquisition, this approach enables the identification and quantification of proteomes. It yields information on peptides, proteins, their origins, and changes in protein expression levels. DIA is a non-targeted proteomic quantification technology that provides comprehensive proteomic information without being limited to specific disease types. It is widely applied in research areas such as disease biomarkers, pathogenesis, disease classification, disease monitoring, drug target discovery, and drug efficacy evaluation (7).

Uridine diphosphate (UDP) is an important nucleotide. The most well-known function of UDP is to act as a carrier or donor for activated sugar molecules (such as galactose, glucose, N-acetylglucosamine, N-acetylgalactosamine, glucuronic acid, etc.). The B4GALT1 (β-1,4-galactosyltransferase 1) gene is located in the p13.3 region of human chromosome 9. The enzyme that it encodes is the founding member of the β-1,4-galactosyltransferase family and plays a core role in the extension of N- and O-linked glycan chains on proteins. This enzyme catalyzes the transfer of galactose from the UDP-galactose donor to the terminal N-acetylglucosamine (GlcNAc) of glycoproteins or glycolipid acceptors in the Golgi apparatus in the form of a β-1,4 glycosidic bond, generating an N-acetyllactosamine

Highlight box

Key findings

- Multi-omics analysis revealed that *B4GALT1* expression was significantly associated with patient survival. The expression of *B4GALT1* positively correlated with macrophage infiltration in the tumor and negatively correlated with $CD4^+$ T cells in the tumor. There was a negative correlation in inactivated naïve B cells, eosinophils, and CD4 naïve T cells, while it showed a positive correlation in dendritic cells, M0/M1/M2 macrophages, natural killer (NK) cells, CD8 T cells, follicular helper T cells, and regulatory T cells. Serum protein B4GALT1 expression was higher in patients with small cell lung cancer (SCLC) than in healthy controls. Elevated serum B4GALT1 protein levels correlated with poor treatment outcomes in patients with SCLC undergoing chemoradiotherapy.

What is known and what is new?

- *B4GALT1* is a key component of the lactose synthase complex and catalyzes the synthesis of lactose from glucose and UDP-galactose in the mammary gland. In non-mammary tissues, it mainly catalyzes the extension of N-glycans and O-glycans, transferring galactose to the N-acetylglucosamine (GlcNAc) residues to form the Galβ1-4GlcNAc structure.
- Multi-omics analysis revealed that *B4GALT1* expression was significantly associated with patient survival. The expression of *B4GALT1* positively correlated with macrophage infiltration in the tumor and negatively correlated with $CD4^+$ T cells in the tumor. Serum protein B4GALT1 expression was higher in patients with SCLC than in healthy controls. Elevated serum B4GALT1 protein levels correlated with poor treatment outcomes in patients with SCLC undergoing chemoradiotherapy.

What is the implication, and what should change now?

- Our findings establish *B4GALT1* as a critical prognostic, diagnostic, and predictive biomarker in SCLC, with its expression closely linked to the tumor immune microenvironment and treatment response. Targeting *B4GALT1* or its related pathways may represent a novel therapeutic strategy, and serum *B4GALT1* holds promise as a liquid biopsy marker for SCLC patient stratification, monitoring, and guiding treatment decisions.

(LacNAc) structural unit (8-10). This reaction is crucial for constructing complex sugar chain structures of various biologically active glycoconjugates and is widely involved in physiological processes such as cell adhesion, immune regulation, and signal transduction. Notably, the function of *B4GALT1* is context-specific (11). In breast tissue, it combines with α-lactalbumin to form a lactose synthase complex, changing the substrate specificity from N-acetylglucosamine to glucose and is thus responsible for the biosynthesis of lactose. Given its extensive and crucial functions, defects in *B4GALT1* are closely related to human diseases. Biallelic loss-of-function mutations in this gene lead to autosomal recessive congenital disorders of glycosylation (*B4GALT1*-CDG, formerly known as CDG-IId), with clinical manifestations, including severe neurodevelopmental abnormalities and multiple congenital malformations. Moreover, numerous studies have shown that *B4GALT1* is dysregulated in various cancer cells and promotes tumor invasion and metastasis by altering the cell surface glycocalyx structure, making it a potential biomarker and therapeutic target in cancer biology (8,12-16). Specifically, a recent study by Cui *et al.* demonstrated that in lung adenocarcinoma (LUAD), high *B4GALT1* expression was significantly associated with poor prognosis in patients. Their research further revealed that *B4GALT1* mediated N-glycosylation of programmed death-ligand 1 (*PD-L1*) protein, thus preventing *PD-L1* degradation (17). The important point is that inhibiting *B4GALT1* can increase the number and activity of $CD8^+$ T cells, and enhance the anti-tumor immune effect of anti-*PD-1* therapy in the body. The expression of *B4GALT1* can promote the growth and metastasis of non-small cell lung cancer. These findings collectively emphasize the multi-faceted role of *B4GALT1* in promoting the malignant development of lung cancer (18). Given the aggressive nature, high mortality rate, and limited effective long-term treatment options of SCLC, understanding how *B4GALT1* functions in the development and progression of SCLC, especially its role in immunity, is of particular interest and requires further research. Therefore, this study aims to investigate the expression pattern, functional significance, and clinical relevance of *B4GALT1* in SCLC. To achieve this, we combined transcriptomics and proteomics, including conducting DIA proteomics quantitative analysis on serum samples from SCLC patients, in order to determine that *B4GALT1* is a potential circulating biomarker. We also identified *B4GALT1* as a promising therapeutic target or diagnostic/prognostic biomarker for patients with SCLC. We present this article in accordance with the REMARK reporting checklist (available at https://jtd.amegroups.com/article/view/10.21037/jtd-2025-1-2610/rc).

Methods

Study design and collection of clinical samples

Patients who visited Beijing Chest Hospital in 2020–2025 and who had a clear pathological diagnosis of SCLC were collected. As healthy individuals, serum samples were collected from 24 participants. These participants underwent health check-ups at Beijing Chest Hospital from 2020 to 2024. None of them had a history of malignant tumors before or during the health check-ups. Serum was separated by centrifugation at 3,000 ×g for 10 min at 4 ℃ and immediately stored at –80 ℃. This study was conducted in accordance with the Declaration of Helsinki and its subsequent amendments. The study was approved by the Ethics Committee of the Beijing Chest Hospital, Capital Medical University (approval number: 2020-18).

The therapeutic effects of all patients were evaluated by the clinical doctors.

The information of the samples sent for sequencing is in the attachment (Lable free-patients). The sample information for the enzyme-linked immunosorbent assay (ELISA) test is provided in the attachment (Elisa-patients).

Inclusion and exclusion criteria

Inclusion criteria

For healthy controls, eligible participants were required to have no clinically significant diseases or health conditions, as assessed through physical examinations and medical histories. Additional criteria included no history of drug dependence or abuse, no history of severe mental disorders (including anxiety, depression, or other mental illnesses), and normal blood routine test results. Female participants were required to confirm that they were neither pregnant or breastfeeding at the time of enrollment.

For SCLC patients, eligible samples were primary tumors diagnosed as I–IV stage tumors. Serum samples were immediately frozen and stored after collection. All tumor samples underwent pathological assessment and were reviewed by at least two independent expert pathologists. The diagnosis of SCLC was confirmed through H&E staining and immunohistochemical detection for chromogranin A, synaptophysin, CD56, and Ki67. The patient case reports were examined, and it was determined

whether the patient samples had received treatment or what type of treatment they had received based on the sample collection time. The patient samples were stored at –80 ℃. All efficacy was evaluated based on the clinical reports of the follow-up physicians, and the efficacy was assessed by the physicians.

Exclusion criteria

For both healthy controls and SCLC patients, individuals with a history of serious or chronic diseases (such as diabetes, hypertension, heart disease, etc.), which may affect the research results were excluded. Participants currently receiving medication or other treatments (such as psychotherapy, physical therapy, etc.) were also excluded. The participant voluntarily withdraws or refuses to participate in the research. For healthy controls, additional exclusion criteria included abnormal blood routine test results and any prior or current history of malignant tumors.

DIA

Total protein extraction

To extract low-abundance proteins from blood using the EasyPept Deep Low-Abundance Enrichment Kit (Thermo Fisher Scientific, Waltham, MA, USA), first use a pipette to gently mix the magnetic beads until they are evenly dispersed. Take an appropriate amount of the magnetic bead suspension and place it on a magnetic rod to stand still for magnetic separation. Discard the supernatant. Add the magnetic bead wash solution to the sample and shake to clean the beads. Place it on the magnetic rod to stand still, then perform magnetic separation and discard the supernatant. Resuspend the beads in an appropriate amount of magnetic bead wash solution. Add 100 µL of plasma to the resuspended beads. Place the mixture on a thermostatic mixer and incubate at 37 ℃ with shaking for 1 hour (1,000 rpm). After incubation, add the magnetic bead wash solution and shake to clean for 5 minutes. Repeat the washing step three times. The low-abundance proteins enriched on the magnetic beads can now be collected.

Protein digestion

Add a reducing alkylation reagent to the magnetic beads and heat at 95 ℃ for 5 minutes. After the sample returns to room temperature, add the enzyme and digest at 37 ℃ for 2 hours (37 ℃, 1,000 rpm). After digestion, add the stop reagent to mix thoroughly and terminate the enzymatic reaction. Centrifuge at 20,000 g for 1 minute and collect the supernatant for subsequent desalting.

Peptide desalting and quantification

After activating, removing impurities, and equilibrating the Cartridge desalting column, load the entire sample into the column and centrifuge at 700 g for 1 minute, repeating the operation once. Wash the desalting column, repeating the operation once. Transfer the desalting column to a new collection plate and elute the sample twice. Concentrate the sample using a vacuum freeze concentrator. After reconstituting the sample with loading buffer, perform peptide quantification using a NANODROP ONE (Thermo Fisher Scientific).

DIA mass detection

The peptides were examined using a VanquishNeo linked to an Orbitrap Astral mass spectrometer (Thermo Fisher Scientific) at Majorbio Bio-Pharm Technology Co. Ltd. (Shanghai, China), according to peptide quantification results. Briefly, the ES906 column (150 µm × 15 cm, Thermo Fisher Scientific) was used with solvent A (water with 2% ACN and 0.1% formic acid) and solvent B (water with 80% ACN and 0.1% formic acid). The peptides were eluted using the 60 SPD gradient at the flow rate of 500 nL/min.

Protein identification

DIA raw data were processed using Spectronaut software (version 18) with the following parameters: protein FDR ≤1%, peptide FDR ≤1%, peptide confidence ≥99%, and XIC width ≤75 ppm. Quantification was performed using 6 peptides per protein and 3 fragment ions per peptide, with shared and modified peptides excluded. Peak areas were summed to generate quantitative values, and only proteins with at least one unique peptide were retained for identification (19-21).

Single-cell data download and processing

This study downloaded single-cell RNA-seq data for SCLC from the GEO database, with dataset numbers GSE138474 and GSE164404, comprising a total of 5 SCLC tissue samples. To eliminate batch effects caused by technical variations between samples, the raw data underwent correction processing first. The SelectIntegrationFeatures function from the Seurat package (version 4.3.0.1) was used to identify 3,000 highly variable genes (HVGs) as the feature set for integration analysis. Subsequently, the

FindIntegrationAnchors function was employed to identify cross-dataset cell correspondences (anchors) based on the mutual nearest neighbor (MNN) algorithm, establishing mapping relationships between batches. Finally, these anchors were used to integrate the expression matrices, yielding the corrected data.

Unsupervised clustering of cells was performed using a graph-based clustering method (FindClusters function) to identify distinct cell subpopulations. To visually represent cell distribution, the RunUMAP function was employed for Unified Manifold Approximation and Projection (UMAP) dimensionality reduction, followed by visualization in a two-dimensional space. Furthermore, the FindAllMarkers function was employed to identify marker genes specific to each cell subpopulation, and differential expression analysis was performed using the FindMarkers function. The screening criteria for differentially expressed genes were: P value <0.05 and $\log_2$FoldChange >0.

InferCNV analysis

To detect copy number variation (CNV) levels in pulmonary neuroendocrine cells (PNECs), the "infercnv" package (version: 1.24.0) was employed for PNEC analysis. Following the official documentation, the analysis retrieved the raw count matrix, cell type annotation file, and gene/chromosome position annotation file. Since no normal samples were available as a control group, T cells were selected as the standard normal cells. The denoising threshold was set to 2.5, with all other parameters retained at their default values.

Immune infiltration analysis of bulk RNA-seq data

Bulk RNA-seq data for SCLC were downloaded from the GEO public database, with dataset numbers GSE6044, GSE40275, and GSE60052. These datasets collectively encompassed 102 SCLC samples and 55 normal tissue samples. Data preprocessing employed the normalizeBetweenArrays function for normalization to eliminate technical batch effects and enhance comparability between samples. Based on the normalized expression data, an immune infiltration analysis was subsequently performed. The estimateScore function from the ESTIMATE package (version: 1.0.13) was utilized to calculate the immune score, stromal score, and overall tumor purity for each sample. This enabled a systematic assessment of the overall composition of immune and stromal components within the tumor microenvironment (TME) and facilitated comparisons of differences between tumor and normal tissues. Additionally, the CIBERSORT algorithm was employed to calculate the relative infiltration levels of immune cells within tumor tissues. This analysis aids in revealing the distribution patterns of different immune cell types within the TME.

Verification of biomarker candidates by ELISA

The level of serum *B4GALT1* was measured using commercially available kits according to the manufacturer's instructions (RayBiotech, Peachtree Corners, GA, USA, Catalog #: ELH-*B4GALT1*-1). Briefly, 50 µL of assay diluent, human recombinant protein standards, and diluted serum samples were added sequentially to appropriate wells. Fifty microliters of detection antibody cocktail were added to each well and incubated at room temperature for 1 h. Following three washes with washing buffer to remove unbound antibody, 100 µL of TMB substrate was added and incubated for 10 min. The reaction was terminated by addition of 100 µL stop solution, and absorbance was measured immediately at 450 nm.

Correlation analysis of characteristic genes

Based on single-cell RNA-seq and bulk RNA-seq data, the expression levels of *PD-L1* and *B4GALT1* in PNEC were evaluated, as well as the expression distribution of *PD-1* and *B4GALT1* in different immune cell subtypes. To systematically assess the correlation between gene expression and cell types, the following analysis process was adopted: all gene expression levels were standardized, then the genes were grouped according to the preset set, and their expression values were integrated in a weighted average manner. On this basis, the Spearman rank correlation analysis method was used to calculate the correlation coefficients between each gene and different immune cell subtypes. To identify statistically significant results, the screening threshold was set at: P<0.05. The final results were visualized in the form of scatter plots to reveal the potential synergistic effects of the target genes in the tumor development process.

Statistical analysis

Bioinformatic analysis for proteomics data was performed using the Majorbio Cloud platform. Differential expression

analysis identified differentially expressed proteins (DEPs) with thresholds of fold change (FC) >1.2 or <0.83 and P<0.05. Functional annotation and enrichment analyses were conducted using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases on the identified DEPs.

For all other statistical analyses, the mean values of variables between two groups were compared using t-test (for normally distributed data with homogeneity of variance), Welch t-test (for normally distributed data with unequal variance), or Wilcoxon rank sum test (for non-normally distributed data). Multi-group comparisons were analyzed using one-way analysis of variance (ANOVA) (for normally distributed data with equal variances), Welch's one-way ANOVA (for normally distributed data with unequal variances), or Kruskal-Wallis test (for non-normally distributed data). Overall survival analysis involved proportional hazards assumption tests and Cox regression, with patients divided into high- and low-expression groups based on median gene expression values. Receiver operating characteristic (ROC) analysis and ROC tests were employed for diagnostic accuracy assessment. Statistical significance was defined as two-sided P<0.05. All statistical analyses were performed using R software (version 4.2.2) with the following packages: survival (version 3.5-7), survminer (version 0.4.9), ggplot2 (version 3.4.4), pROC (version 1.18.0), stats (version 4.2.1), and car (version 3.1-0).

Results

Serum proteomics in SCLC

We employed DIA quantitative proteomics technology to analyze 20 serum samples (including 5 healthy control samples and 15 SCLC samples). Quality control results revealed 11,363 peptide segments and 1,222 proteins that underwent proteomic analysis. First, we calculated correlations between biological replicates within different sample groups and clustered the results. The samples from the same group showed a high correlation (*Figure 1A*). Volcano plots revealed 246 proteins with statistically significant differential relative expression (FC ≥1.2 or FC ≤1, P<0.05) (*Figure 1B*). GO enrichment analysis of the DEPs revealed major enrichment in chromatin remodeling, organization of protein-DNA complexes, chromosome structure, and pore complexes (*Figure 1C*). KEGG enrichment analysis further indicated that these proteins were predominantly enriched in alcoholic liver disease, various N-glycan biosynthesis pathways, and metabolic pathways (*Figure 1D*). Reactome analysis showed enrichment of SLIT and ROBO expression regulation, formation of free 40S subunits, and apoptosis-induced DNA cleavage (*Figure 1E*).

Multi-omics analysis

We analyzed publicly available data from a previously published SCLC cohort (22), comprising 112 tumors and matched specimens analyzed using mass spectrometry-based proteomics (OMIX database: OMIX002489; https://ngdc.cncb.ac.cn/omix). We extracted log-transformed expression values and performed Kaplan-Meier plotter survival analysis on DEPs within the dataset, dividing the patient cohort into two groups based on median protein expression levels. Subsequently, we integrated the analysis of publicly available data from Atay's study (22). Their study primarily conducted a comprehensive transcriptomic analysis of three gene expression datasets (GSE30219, GSE43346, and GSE149507), identifying 1,734 upregulated genes and 2,907 downregulated genes. Integrating proteins, transcriptomes, and our own serum proteomics analyses identified *B4GALT1* as a gene associated with survival in SCLC patients (*Figure 2A*). Patients with high B4GALT1 protein expression had shorter survival times than those with low protein expression.

We also analyzed the study by Stewart *et al.* (23). Using classic marker annotations, the cells were classified as B cells, endothelial cells, fibroblasts, PNEC, and T cells, and the immune cells were further divided into subgroups (*Figure 2B*). We grouped the tumor cells based on the expression level of *B4GALT1* and conducted a high-low classification and differential analysis (*Figure 2C*). We found that the immune microenvironment of small cell lung cancer was suppressed (*Figure 2D*,*2E*). At the same time, we also analyzed the direct relationship between *B4GALT1* and different immune cells, and found that the expression of *B4GALT1* was mainly positively correlated with macrophages and negatively correlated with CD4 cells (*Figure 2F*). We then performed GO and KEGG enrichment analysis on the obtained differential genes, and found that the differentially expressed genes were mainly enriched in cytoplasmic translation, ribosome, ribosomal subunit, structural constituent of ribosome, protein processing in endoplasmic reticulum, and chemical

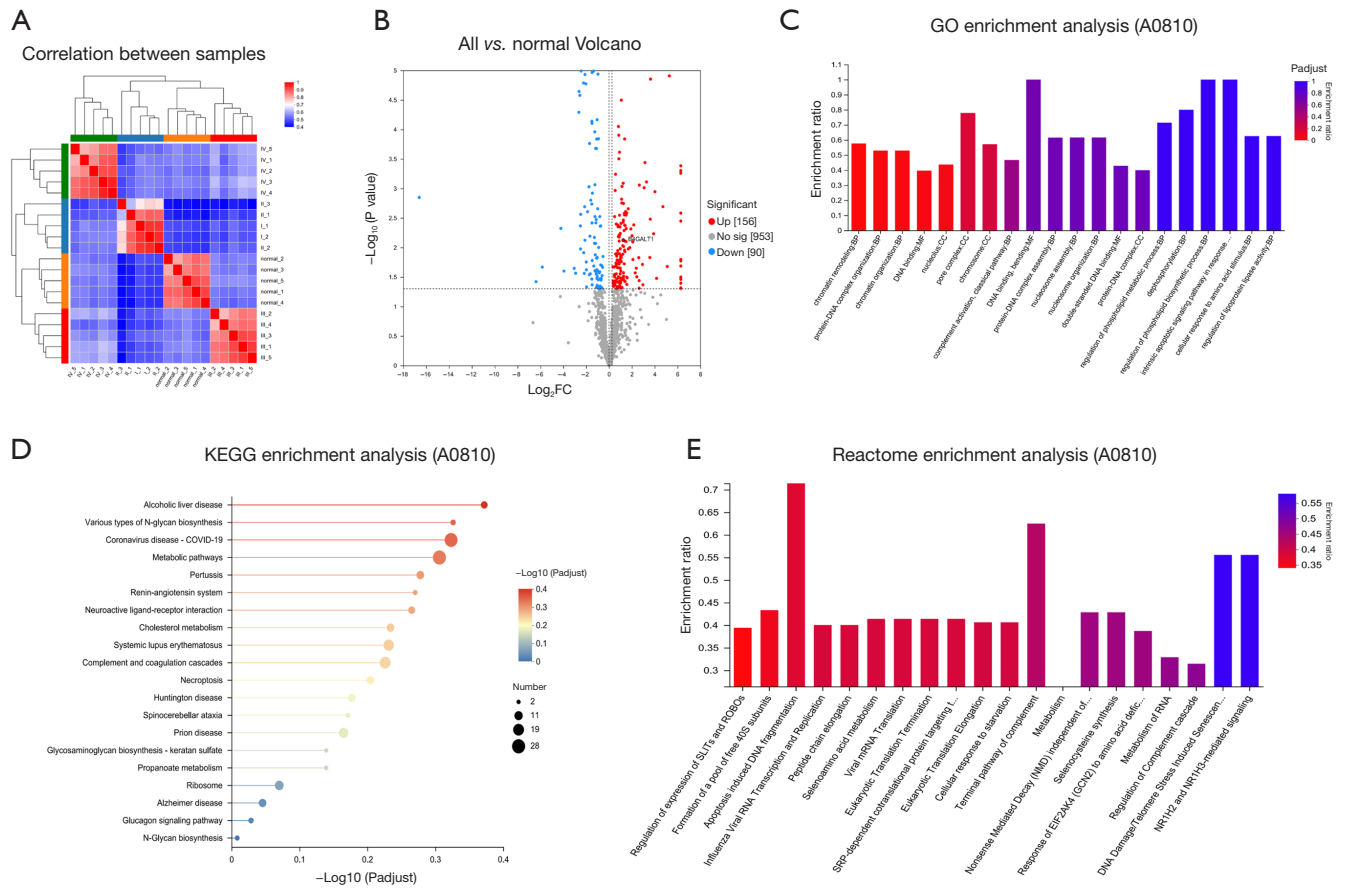

Figure 1 Serum proteomics in small cell lung cancer. (A) Sample correlation heatmap. (B) Volcano plot showing the differences in protein expression. (C,D) GO and KEGG enrichment analysis. (E) Reactome enrichment analysis. BP, biological process; CC, cellular component; FC, fold change; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; MF, molecular function.

carcinogenesis-reactive oxygen species (*Figure 2G-2J*).

Correlation analysis between B4GALT1 and the SCLC immune microenvironment

Previously, it was found that the expression of *B4GALT1* and *PD-L1* was correlated in LUAD, and that they were significantly associated with the occurrence, development, invasion, and clinical prognosis of LUAD (17,23). In PNEC, we analyzed the relationship between *PD-L1* and *B4GALT1* and simultaneously analyzed the relationship between *B4GALT1* and *PD-1* in different immune cells (*Figure 3*). The expression of *B4GALT1* and *PD-1* was negatively correlated in naive B cells (Spearman r=−0.359, P<0.001), eosinophils (Spearman r=−0.37, P<0.001), and CD4 naive cells (Spearman r=−0.31, P=0.002), suggesting that *B4GALT1* may inhibit the upregulation of immune checkpoints in these cells and maintain their activity. The expressions of *B4GALT1* and *PD-1* were positively correlated in DC cells (Spearman r=0.248, P=0.01), M0 macrophages (Spearman r=0.22, P=0.03), M1 macrophages (Spearman r=0.489, P<0.001), M2 macrophages (Spearman r=0.355, P<0.001), natural killer (NK) cells (Spearman r=0.281, P=0.004), CD8 T cells (Spearman r=0.397, P<0.001), follicular helper T cells (Spearman r=0.316, P=0.001), and Tregs (Spearman r=0.36, P<0.001).

ELISA validation of the B4GALT1 value in serum of SCLC patients

We collected serum samples from 82 patients with SCLC and 24 healthy individuals at Beijing Chest Hospital during the period from 2020 to 2025, as well as their clinical information. Patients were categorized into the following

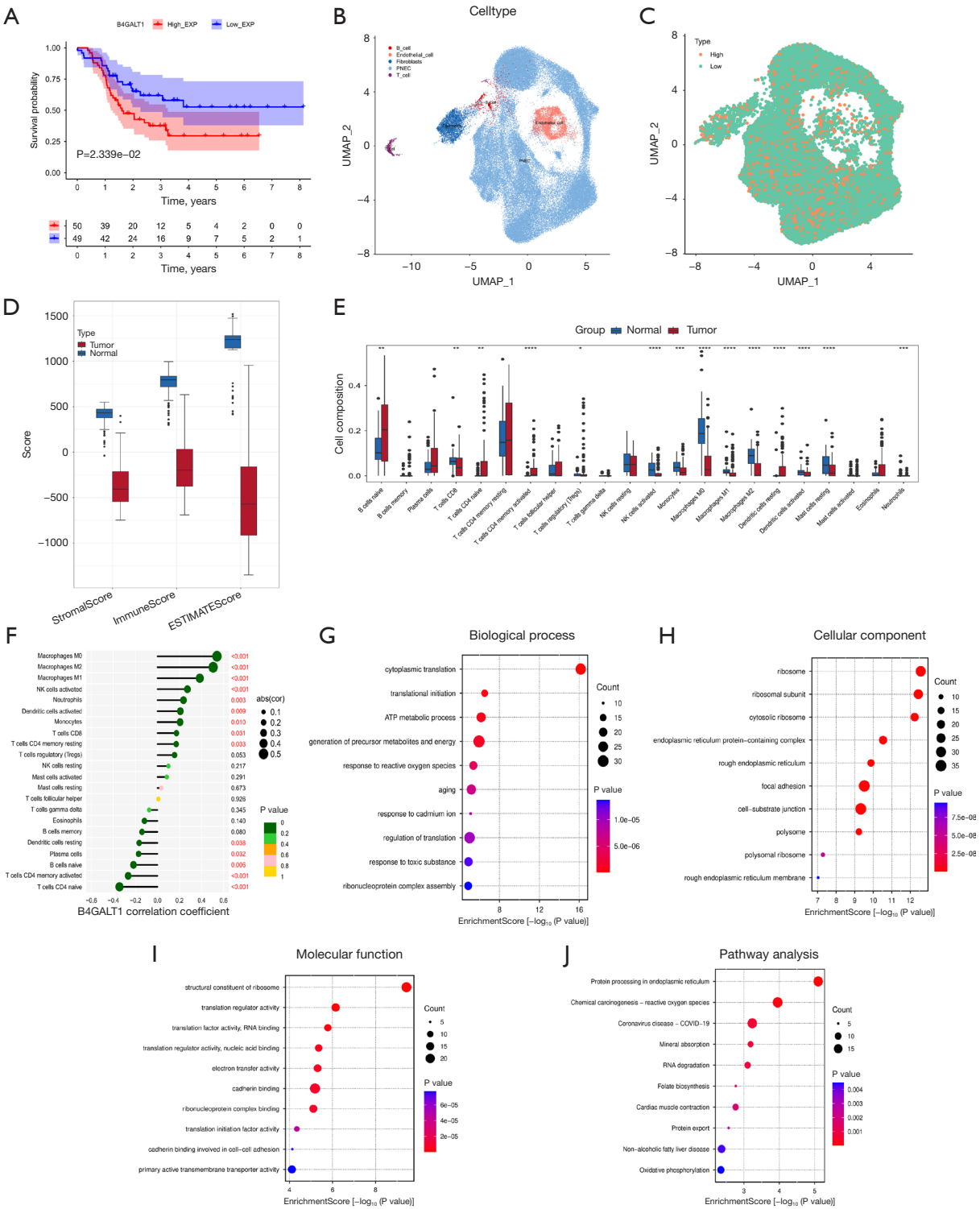

Figure 2 Multi-omics analysis. (A) Kaplan-Meier plotter survival analysis. (B) Cell clustering in each group. (C) The tumor cells were grouped according to the median value of *B4GALT1* expression (high: above median; low: below median). (D,E) Comparison of normal and tumor samples using CIBERSORT and ESTIMATE analyses for immune cell infiltration. (F) Correlation between *B4GALT1* expression and immune cell infiltration. (G-J) GO and KEGG enrichment analysis. ns, P>0.05; *, P<0.05; **, P<0.01; ***, P<0.001; ****, P<0.0001. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

groups: healthy controls (n=24), untreated SCLC group (n=17), chemotherapy group (n=17), chemoradiotherapy group (n=23), chemotherapy-immunotherapy group (n=18), and other groups (chemoradiotherapy combined with targeted therapy, n=7). The total number of patients included in the therapy group was 65. Detection and comparison of B4GALT1 protein expression in patient serum revealed higher levels in patients with SCLC than in healthy controls, with treated patients also exhibiting significantly elevated expression compared to healthy controls (healthy controls: 13.981 ng/mL; SCLC: 45.354 ng/mL; therapy: 40.987 ng/mL) (*Figure 4A*). Serum B4GALT1 levels were elevated in all treatment groups compared to healthy controls, but no statistically significant differences existed between the treatment groups (chemotherapy group: 37.634 ng/mL; chemoradiotherapy group: 43.752 ng/mL; chemotherapy immunotherapy group: 39.964 ng/mL; others: 42.676 ng/mL) (*Figure 4B*).

ROC analysis was performed to determine the diagnostic accuracy of serum B4GALT1 for disease diagnosis. The area under the curve for B4GALT1 was 0.995 (95% confidence interval: 0.983–1.000) (*Figure 4C*). The cut-off value for B4GALT1 was 29.207 ng/mL. After the cutoff, the sensitivity and specificity of B4GALT1 were 1 and 0.95833, respectively. We analyzed whether serum B4GALT1 protein levels in SCLC patients correlated with treatment efficacy. The results suggest that high serum B4GALT1 protein expression is associated with poor response to chemoradiotherapy in patients with SCLC (*Figure 4D*).

Discussion

In summary, our study demonstrates that serum B4GALT1 is significantly elevated in patients with SCLC compared to healthy controls, and higher levels of B4GALT1 are closely associated with poor clinical outcomes and diminished responses to chemoradiotherapy. Multi-omics and immune infiltration analyses further elucidate the complex role of B4GALT1 in the tumor immune microenvironment. The negative correlation between *B4GALT1* expression and *PD-1* expression observed in immature B cells, eosinophils, and CD4 immature cells suggests that *B4GALT1* may play a crucial role in maintaining the activity of these immune cells by inhibiting the upregulation of immune checkpoints. Conversely, the positive correlation between *B4GALT1* and *PD-1* in functional immune cell populations such as macrophages, NK cells, and $CD8^+$ T cells indicates that it may enhance the expression of immune checkpoints through post-translational modifications. These findings not only highlight the relevance of B4GALT1 as a potential biomarker for evaluating the tumor immune microenvironment, but also provide a promising approach for regulating the *PD-1/PD-L1* axis through targeting *B4GALT1* to optimize immunotherapy strategies. These findings, combined with previous reports in lung adenocarcinoma and other cancers, support the pivotal role of *B4GALT1* in promoting tumor progression via glycosylation-mediated modulation of immune checkpoints such as PD-1/PD-L1, ultimately influencing immune evasion and therapy resistance (17).

Serum B4GALT1 has emerged as a sensitive and easy-to-use biomarker, suitable for liquid biopsy. It can provide timely stratification, monitoring, and treatment guidance for patients with SCLC (especially those who cannot undergo tissue sampling). Our research results advocate incorporating serum-based B4GALT1 measurement into clinical practice, not only for diagnosis and prognosis prediction, but also for optimizing immunotherapy strategies - possibly by targeting *B4GALT1* to enhance the anti-tumor immune response.

Next, mechanistic studies and large-scale clinical validations are required to further clarify the therapeutic potential of B4GALT1 and explore its broader applicability in different lung cancer subtypes. Overall, this work lays the foundation for developing precision medicine based on B4GALT1 in SCLC, highlighting its potential as a multi-dimensional biomarker for patient management and as a driving force for individualized cancer treatment.

Limitations

There are still certain limitations in this study. Small cell samples are difficult to obtain and are very precious. Firstly, the sample size of the serum samples included is limited, which may not fully represent the heterogeneity of all patients with SCLC. Moreover, the samples mainly come from a single center, which introduces a certain selection bias. The broad applicability of the results needs to be further verified. Secondly, this study is of a cross-sectional design and lacks long-term follow-up data. Therefore, it cannot assess the dynamic changes of B4GALT1 during the disease course and treatment and the relationship with actual clinical effects. Regarding the mechanism of B4GALT1, this study has not conducted in-depth exploration. Further related functional experiments are needed to clarify this. The article first conducted a preliminary discovery using

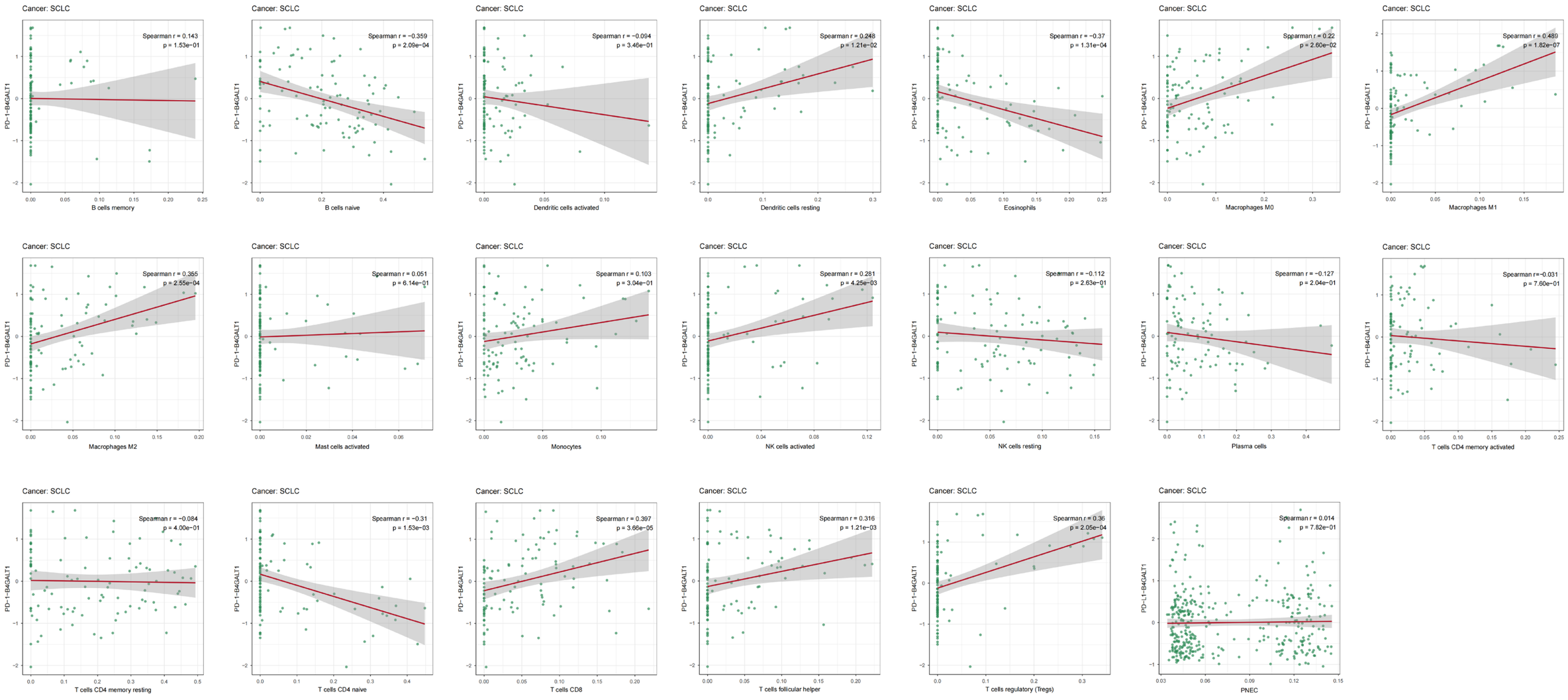

Figure 3 Correlation analysis between B4GALT1 and the SCLC immune microenvironment. NK, natural killer; PNEC, pulmonary neuroendocrine cell; SCLC, small cell lung cancer.

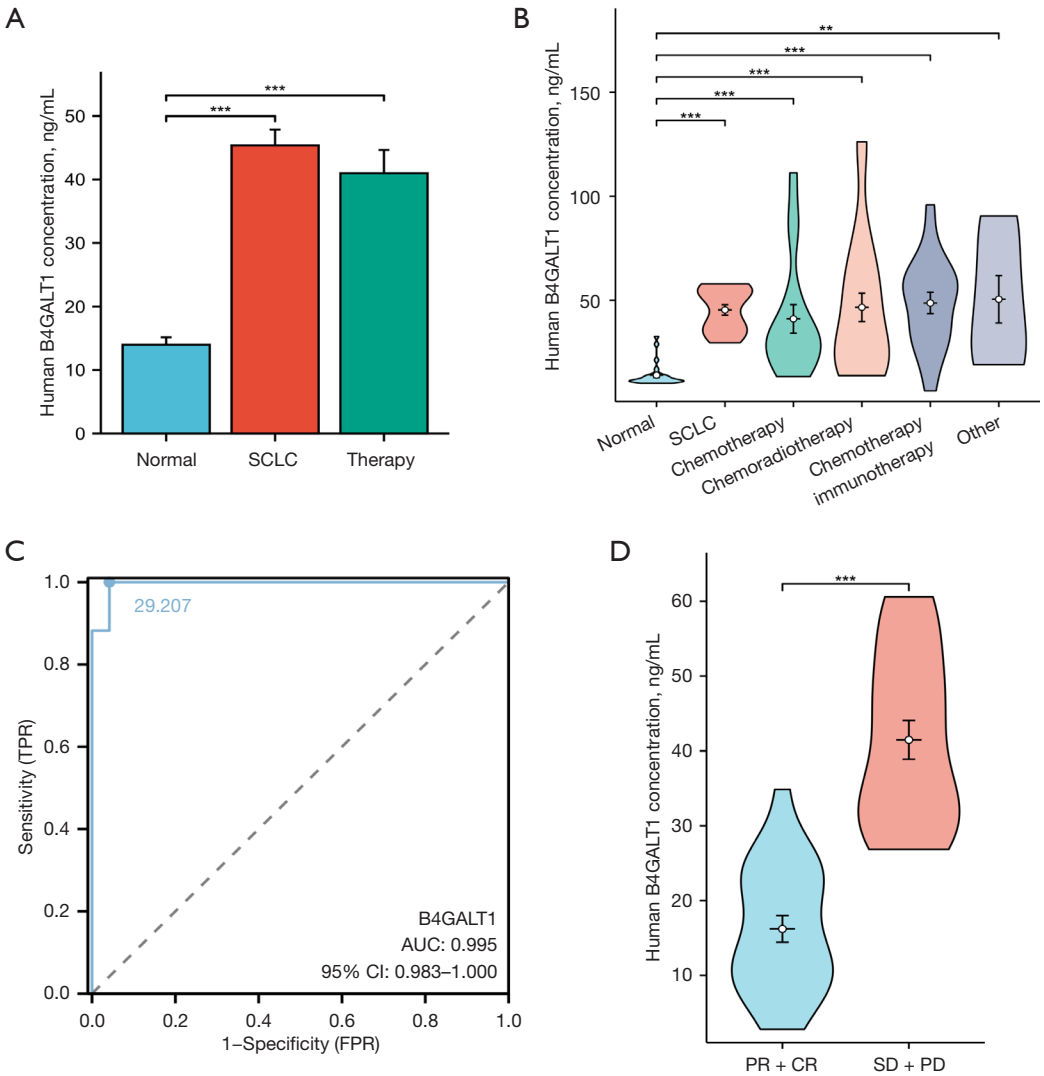

Figure 4 Serum expression of B4GALT1 in humans. (A) B4GALT1 protein expression in patient serum. (B) Compared to healthy controls, B4GALT1 expression in the serum of patients receiving different treatments. (C) ROC analysis. (D) Correlation between serum B4GALT1 protein levels and treatment outcomes in patients with SCLC. **, P<0.01; ***, P<0.001. AUC, area under the curve; CI, confidence interval; CR, complete response; FPR, false positive rate; PD, progressive disease; PR, partial response; ROC, receiver operating characteristic; SCLC, small cell lung cancer; SD, stable disease; TPR, true positive rate.

DIA proteomics, and then verified it through the ELISA method (which is a highly sensitive and highly specific method). This may help to accurately determine the content of B4GALT1, thereby improving the accuracy of diagnosis. Although this remarkable achievement has been achieved, we are well aware that in clinical diagnosis, achieving such nearly perfect results with a single biomarker is extremely rare. We recognize that there may be overfitting in smaller datasets and the need to avoid selection bias. Therefore, we emphasize that these preliminary findings require strict validation in a larger, independent, and geographically diverse cohort. Future research will focus on conducting prospective validation in a broader patient population, including the differentiation from other pulmonary lesions and non-malignant conditions, to confirm the reliability and general applicability of B4GALT1 as a diagnostic biomarker for SCLC. Additionally, exploring combinations of multiple biomarkers with B4GALT1 may further improve diagnostic

accuracy and overcome the limitations of a single biomarker.

Conclusions

Our research results indicate that *B4GALT1* is a multifaceted biomarker in SCLC, demonstrating its significance in prognosis, diagnosis, and prediction of treatment response. We have revealed the close association between B4GALT1 expression and the tumor immune microenvironment, suggesting that it plays a crucial role in regulating anti-tumor immunity. Therefore, therapeutic strategies targeting B4GALT1 or its downstream pathways are expected to become a promising new treatment approach. Additionally, the potential of serum B4GALT1 as a liquid biopsy marker for stratifying SCLC patients, monitoring the disease, and guiding personalized treatment decisions merits further investigation.

Acknowledgments

We would like to express our gratitude to the Laboratory Department of Beijing Chest Hospital for providing clinical samples, and to the Oncology Center of Beijing Chest Hospital for providing the experimental platform.

Footnote

Reporting Checklist: The authors have completed the REMARK reporting checklist. Available at https://jtd.amegroups.com/article/view/10.21037/jtd-2025-1-2610/rc

Data Sharing Statement: Available at https://jtd.amegroups.com/article/view/10.21037/jtd-2025-1-2610/dss

Peer Review File: Available at https://jtd.amegroups.com/article/view/10.21037/jtd-2025-1-2610/prf

Funding: None.

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://jtd.amegroups.com/article/view/10.21037/jtd-2025-1-2610/coif). All authors report that there is one patent application (application No. 202511193954.3) pending under review. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study was conducted in accordance with the Declaration of Helsinki and its subsequent amendments. The study was approved by the Ethics Committee of Beijing Chest Hospital, Capital Medical University (approval number: 2020-18). This project was deemed exempt from the requirement for informed consent by the Ethics Committee of Beijing Chest Hospital.

References

1. Kim SY, Park HS, Chiang AC. Small Cell Lung Cancer: A Review. JAMA 2025;333:1906-17.
2. Lee JH, Saxena A, Giaccone G. Advancements in small cell lung cancer. Semin Cancer Biol 2023;93:123-8.
3. Dingemans AC, Früh M, Ardizzoni A, et al. Small-cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up(☆). Ann Oncol 2021;32:839-53.
4. Kalemkerian GP, Khurshid H, Ismaila N, et al. Systemic Therapy for Small Cell Lung Cancer: ASCO Guideline Rapid Recommendation Update. J Clin Oncol 2025;43:101-5.
5. Yee NS. Liquid Biopsy: A Biomarker-Driven Tool towards Precision Oncology. J Clin Med 2020;9:2556.
6. Fasano R, Serratì S, Rafaschieri T, et al. Small-Cell Lung Cancer: Is Liquid Biopsy a New Tool Able to Predict the Efficacy of Immunotherapy? Biomolecules 2024;14:396.
7. Lou R, Cao Y, Li S, et al. Benchmarking commonly used software suites and analysis workflows for DIA proteomics and phosphoproteomics. Nat Commun 2023;14:94.
8. Hsu TH, Chang YC, Lee YY, et al. B4GALT1-dependent galectin-8 binding with TGF-β receptor suppresses colorectal cancer progression and metastasis. Cell Death Dis 2024;15:654.
9. De Vitis C, D'Ascanio M, Sacconi A, et al. B4GALT1 as a New Biomarker of Idiopathic Pulmonary Fibrosis. Int J

Mol Sci 2022;23:15040.
10. Chang HH, Chen CH, Chou CH, et al. β-1,4-Galactosyltransferase III enhances invasive phenotypes via β1-integrin and predicts poor prognosis in neuroblastoma. Clin Cancer Res 2013;19:1705-16.
11. Nakano M, Saldanha R, Göbel A, et al. Identification of glycan structure alterations on cell membrane proteins in desoxyepothilone B resistant leukemia cells. Mol Cell Proteomics 2011;10:M111.009001.
12. Chen PD, Liao YY, Cheng YC, et al. Decreased B4GALT1 promotes hepatocellular carcinoma cell invasiveness by regulating the laminin-integrin pathway. Oncogenesis 2023;12:49.
13. Zhai F, Li Y, Zheng J, et al. SPOP/NOLC1/B4GALT1 signaling axis enhances paclitaxel resistance in endometrial cancer by inducing O-dysglycosylation. Oncogene 2025;44:1694-708.
14. Ahvaz S, Amini M, Yari A, et al. Downregulation of long noncoding RNA B4GALT1-AS1 is associated with breast cancer development. Sci Rep 2024;14:3114.
15. Lu X, Wang B, Qiu P, et al. Bioinformatics prediction and experimental verification identify B4GALT1 as a diagnostic biomarker of glioblastoma. Int J Biol Macromol 2025;318:144858.
16. Zhang J, Zhang G, Zhang W, et al. Loss of RBMS1 promotes anti-tumor immunity through enabling PD-L1 checkpoint blockade in triple-negative breast cancer. Cell Death Differ 2022;29:2247-61.
17. Cui Y, Li J, Zhang P, et al. B4GALT1 promotes immune escape by regulating the expression of PD-L1 at multiple levels in lung adenocarcinoma. J Exp Clin Cancer Res 2023;42:146.
18. Zhao Y, Zhang D, Meng B, et al. Integrated proteomic and glycoproteomic analysis reveals heterogeneity and molecular signatures of brain metastases from lung adenocarcinomas. Cancer Lett 2024;605:217262.
19. Chen S, Siedhoff HR, Zhang H, et al. Low-intensity blast induces acute glutamatergic hyperexcitability in mouse hippocampus leading to long-term learning deficits and altered expression of proteins involved in synaptic plasticity and serine protease inhibitors. Neurobiol Dis 2022;165:105634.
20. Balázs A, Millar-Büchner P, Mülleder M, et al. Age-Related Differences in Structure and Function of Nasal Epithelial Cultures From Healthy Children and Elderly People. Front Immunol 2022;13:822437.
21. Han C, Shi C, Liu L, et al. Majorbio Cloud 2024: Update single-cell and multiomics workflows. Imeta 2024;3:e217.
22. Liu Q, Zhang J, Guo C, et al. Proteogenomic characterization of small cell lung cancer identifies biological insights and subtype-specific therapeutic strategies. Cell 2024;187:184-203.e28.
23. Stewart CA, Gay CM, Xi Y, et al. Single-cell analyses reveal increased intratumoral heterogeneity after the onset of therapy resistance in small-cell lung cancer. Nat Cancer 2020;1:423-36.

Cite this article as: Wang C, Zhu X, Shi Y, Wang G, Ma T. Multi-omics analysis identified serum *B4GALT1* as a prognostic factor for small cell lung cancer. J Thorac Dis 2026;18(4):353. doi: 10.21037/jtd-2025-1-2610