

A core fucosylation–derived signature reveals extracellular matrix remodeling and predicts poor outcomes in pancreatic ductal adenocarcinoma

Bokai Zhao^{1,2}, Jin Chen^{2,3,4}, Ruijun Tian^{2,3}, and Liwei Cao^{1,2} ✉

¹ Department of Biochemistry, School of Medicine, Southern University of Science and Technology, Shenzhen, 518055, China

² Shenzhen Key Laboratory of Functional Proteomics, Southern University of Science and Technology, Shenzhen, 518055, China

³ Department of Chemistry and Research Center for Chemical Biology and Omics Analysis, School of Science, Southern University of Science and Technology, Shenzhen, 518055, China

⁴ Clinical Center for Molecular Diagnosis and Therapy, the Second Affiliated Hospital of Fujian Medical University, Quanzhou, 362000, China

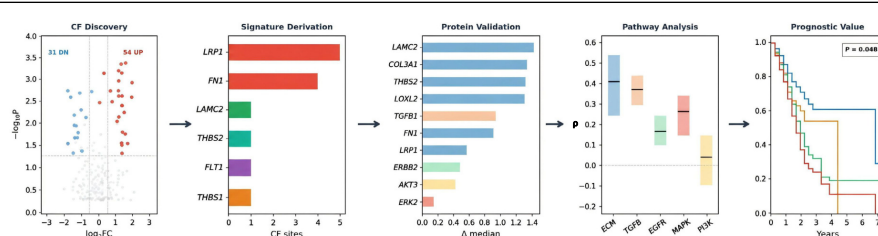


Cite This: *Oral Sci Homeost Med*, 2026, 2, 9610054



Read Online

ABSTRACT: Core fucosylated (CF) glycoproteins are closely associated with the pathogenesis of pancreatic ductal adenocarcinoma (PDAC). However, how core fucosylation alterations connect to critical signaling pathways in PDAC, such as dense extracellular matrix (ECM) remodeling and KRAS–MAPK signaling, remains



Core fucosylation → ECM remodeling (LRP1–FN1–LAMC2) → ECM/integrin–dominant biology with TGFβ, EGFR/ERBB parallel associations → Prognostic value in PDAC

largely unknown. Here, we re-analyzed core fucosylation profiles of six matched PDAC tumor and adjacent normal tissues from a published dataset and identified 54 significantly up-regulated core fucosylation events, which were predominantly associated with ECM remodeling. Leveraging these events, we developed a PDAC prognostic signature comprising three CF glycoproteins, including LRP1, FN1, and LAMC2. This signature was further validated using the tumor-normal comparison in Clinical Proteomic Tumor Analysis Consortium (CPTAC) PDAC cohort. Then, unbiased screening across 1,387 pathways using the Cancer Genome Atlas Program Pancreatic Adenocarcinoma Database (TCGA-PAAD) identified cell–ECM interactions as the dominant association with the signature. Moreover, it was confirmed using tumor-normal comparison across three independent Gene Expression Omnibus (GEO) cohorts (AUC 0.888–0.941) and was associated with adverse prognosis in TCGA-PAAD. Together, these works proposed a core fucosylation-derived prognostic signature linking glycosylation alterations to matrix remodeling and poor outcome in PDAC.

KEYWORDS: core fucosylation, extracellular matrix remodeling, prognosis, pancreatic ductal adenocarcinoma, mass spectrometry

1 Introduction

Protein glycosylation is a widespread post-translational modification in mammalian cells^[1,2]. Among many glycosylation modifications, core fucosylation has drawn increasing attention in recent years because of its involvement in multiple human cancer types^[3,4]. In pancreatic ductal adenocarcinoma (PDAC), altered core fucosylation has been linked to both tumor cell behavior and biomarker discovery^[5]. FUT8, the sole enzyme catalyzing protein core fucosylation, is overexpressed in pancreatic cancer cells, and

core fucosylated (CF) proteins such as serum haptoglobin have been implicated as candidate biomarkers of the disease^[6–8]. However, PDAC-specific CF glycoproteins remain incompletely characterized at the site-specific level.

PDAC is characterized by a dense desmoplastic stroma in which extracellular matrix (ECM) deposition, basement membrane remodeling, and tumor–stromal crosstalk actively shape disease progression^[6–9]. Oncogenic KRAS mutations are present in the vast majority of PDACs, activate RAF–MEK–ERK signaling, and cooperate with the stromal microenvironment to sustain proliferative and invasive programs^[10–14]. Many key proteins at the tumor–stroma interface—including laminins, integrins, and fibronectin—are glycoproteins whose functions may be influenced by core fucosylation^[2]. However, whether core fucosylation alterations are systematically linked to these dominant PDAC

Received: March 30, 2026; Revised: April 21, 2026

Accepted: May 6, 2026

✉ Address correspondence to Liwei Cao, caolw@sustech.edu.cn



清华大学出版社
Tsinghua University Press

SciOpen

programs remains unclear.

A recent site-specific glycoproteomic study catalogued extensive CF glycosites in paired PDAC tumor and adjacent normal tissues^[15]. Many of the altered glycoproteins were ECM-associated, suggesting that core fucosylation changes may play an important role in matrix-centered PDAC biology^[15]. This dataset therefore provides an opportunity to test the hypothesis and develop a core fucosylation-based signature linked to matrix-centered tumor microenvironment remodeling.

Here, we re-analyzed these glycosite-level data to derive a core fucosylation-derived ECM signature centered on three CF glycoproteins, including LRP1, FN1, and LAMC2. This signature was then evaluated through multi-omics validation across The Cancer Genome Atlas Pancreatic Adenocarcinoma Database (TCGA-PAAD), the proteome and phosphoproteome of pancreatic cancer derived from the Clinical Proteomic Tumor Analysis Consortium (CPTAC), and three independent Gene Expression Omnibus (GEO) cohorts to define its relationship to ECM-centered PDAC biology and its clinical utility. Based on these evaluations, we identify a core fucosylation-derived signature that links to ECM remodeling and predicts poor outcomes in PDAC.

2 Methods

2.1 Data sources

The discovery dataset was downloaded from a published site-specific core fucosylation glycoproteomic dataset of PDAC^[15]. This dataset provided paired tumor and adjacent normal glycosite-level data. Transcriptomic expression data (RNA-seq, HiSeq V2), single-sample gene set enrichment analysis (ssGSEA) pathway matrices, and survival annotations for TCGA-PAAD were downloaded from UCSC Xena^[16]. Gene-level protein abundance and phosphosite abundance data were obtained from the CPTAC PDAC discovery cohort^[17] via the LinkedOmics portal^[18]. Sample-level tumor/normal annotations were mapped based on the sample identifiers provided in the cohort. Independent external transcriptomic validation was performed using GEO cohorts GSE15471, GSE28735, and GSE62452.

2.2 Differential glycosite analysis and candidate identification

For each CF event, effect size was defined as the median $\log_2$ (Tumor/normal adjacent tissue (NAT)) across paired samples. Statistical significance was assessed using paired t-tests. At this initial discovery stage, we followed the screening criteria used in the original PDAC core-fucosylation study^[15], and used raw P values together with an effect-size threshold ($|\log_2\text{FC}| \geq 0.585$, corresponding to a 1.5-fold change) to identify altered core fucosylation events. Because the paired discovery cohort was small ($n = 6$), this strategy was intended for candidate screening rather than definitive statistical selection, while downstream prioritization and multi-cohort/multi-omics validation were used to improve robustness. Events meeting both raw $P < 0.05$ and $|\log_2\text{FC}| \geq 0.585$ were considered significantly altered^[15]. Significantly upregulated events were collapsed to the protein level, and proteins harboring multiple upregulated core fucosylation sites were retained for further prioritization. LRP1 and FN1 were selected as the first two components because they harbored the largest numbers of significantly upregulated sites (five and four, respectively).

2.3 Signature development and sensitivity analysis

Further refinement focused on four remaining candidates (LAMC2, THBS2, FLT1, and THBS1), each harboring one significantly upregulated core fucosylation site. Each candidate was combined with LRP1 and FN1 to form a three-component signature. For each one, Spearman correlations were computed between the resulting signature score and representative ssGSEA pathway scores in the TCGA-PAAD cohort. Representative pathways were chosen to summarize five PDAC-related signaling axes: ECM/integrin^[6,8], TGF β ^[14,19], EGFR/ERBB^[20], PI3K/AKT/mTOR^[21], and RAF-MEK-ERK^[11,12]. The final signature was selected based on the most coherent overall correlation pattern across these axes. A supplementary prioritization framework is provided in Supplementary Table S1.

For all subsequent analyses, the LRP1-FN1-LAMC2 signature score was calculated as the sample-wise mean of z-score-normalized expression or abundance values of the three genes/proteins within each cohort. Samples were stratified into signature-high and signature-low groups using the cohort-specific median score as the cutoff.

2.4 Functional enrichment analysis

Gene set enrichment was analyzed using GSEAPy through the Enrichr API^[22,23]. Four gene sets were queried: MSigDB Hallmark, KEGG 2021, Reactome 2022, and GO Biological Process 2021. Terms with adjusted $P < 0.05$ were considered significant.

2.5 Intermediate abundance validation in LinkedOmics/CPTAC

The LRP1-FN1-LAMC2 signature score was evaluated in the CPTAC PDAC cohort using global proteome and N-linked glycoproteome datasets to compare glycosite-derived and abundance-based patterns^[17,18].

2.6 CPTAC proteomic and phosphoproteomic validation

For proteomic analysis, representative proteins spanning ECM structural proteins, integrin subunits, ECM remodeling enzymes, TGF β pathway components, MAPK cascade members, PI3K/AKT/mTOR pathway components, and EGFR/ERBB family members were compared between tumor and adjacent normal tissues. When a gene corresponded to multiple rows, sample-level values were first averaged. The effect size was defined as the difference in median abundance (Δ median). Group differences were assessed using two-sided Mann-Whitney U tests.

For phosphoproteomic analysis, phosphosites were selected to represent five signaling axes: ECM-proximal (ITGB1, ILK, paxillin), TGF β receptor (TGFBR2), MAPK cascade (MAP2K1, MAPK1, MAPK3), EGFR/ERBB (EGFR, ERBB2), and PI3K/AKT (PIK3R1, AKT3). Each phosphosite was evaluated across three dimensions: tumor versus adjacent normal (Δ median; Mann-Whitney U), signature-high versus signature-low in tumor-only samples (Welch's t-test), and continuous Spearman correlations with the proteome-derived signature score.

2.7 Unbiased pathway correlation analysis

In TCGA-PAAD, Spearman correlations were computed between the signature score and ssGSEA enrichment scores for all 1,387 available pathways^[24]. Pathways were grouped into functional categories (ECM, TGF β , EGFR/ERBB, MAPK, PI3K/AKT, and

others) for category-level comparison. Representative terms related to extracellular matrix, integrin, laminin, or focal adhesion were grouped as ECM/adhesion; those related to TGF β /SMAD/activin signaling were grouped as TGF β ; those related to EGFR/ERBB signaling were grouped as EGFR/ERBB; those related to MAPK/MEK/ERK/RAF signaling were grouped as MAPK; and those related to PI3K/AKT/mTOR signaling were grouped as PI3K/AKT/mTOR. Group differences between signature-high and signature-low tumors were assessed using Mann–Whitney U tests.

2.8 External GEO validation

Microarray probes of GSE15471, GSE28735, and GSE62452 were mapped to gene symbols using annotations from the matched platforms. When multiple probes mapped to the same gene, intensity values were collapsed by median. Tumor–normal differences in signature score were assessed using Mann–Whitney U tests. Receiver operating characteristic (ROC) analysis was performed using scikit-learn^[25]. For exploratory assessment of pathway reproducibility, small expression-based proxy gene sets were manually curated from canonical pathway knowledge and representative PDAC/pathway literature. The exact gene lists and proxy-score definitions of these manually curated proxy sets were provided in Supplementary Table S2. For each axis, a proxy score was defined as the mean of z-score-normalized expression values of constituent genes within each cohort. Gene groups represented ECM/adhesion^[6,8], TGF β ^[14,19], EGFR/ERBB^[20], MAPK-related ERK output^[11,12], and PI3K/AKT/mTOR^[21].

2.9 Survival analysis and prognostic evaluation

The prognostic value of the signature was evaluated in 183 TCGA-PAAD patients with complete survival annotation and matched RNA-seq data^[16]. Patients were grouped into signature score-high ($n = 92$) and signature score-low ($n = 91$) groups using the median signature score.

For interaction analysis between the signature and clinical stage, patients were stratified using an American Joint Committee on Cancer (AJCC)-based stage regrouping variable, defined as I+IIA versus IIB+III+IV. This regrouping was chosen to provide a clinically interpretable boundary at the IIA/IIB transition while avoiding the severe imbalance produced by the conventional I+II versus III+IV split in the TCGA-PAAD cohort^[26]. Two patients with unavailable staging were excluded from the four-group analysis.

Time-dependent AUC values were computed from a Cox model incorporating the continuous signature score and the binary AJCC-based stage regrouping variable, using the cumulative/dynamic estimator^[27] implemented in scikit-survival (v0.22)^[28]. Kaplan–Meier estimation and log-rank testing were implemented in lifelines (v0.29). Statistical significance was defined as $P < 0.05$ (two-sided).

2.10 Statistical analysis and software

Statistical analyses were performed in Python 3.13. Core packages included pandas, NumPy, SciPy, matplotlib, scikit-learn^[25], scikit-survival^[28], lifelines, and GSEapy^[22]. Paired glycosite-level comparisons used paired t-tests. Unpaired group comparisons used two-sided Mann–Whitney U tests unless otherwise indicated. Welch’s t-test was used for tumor-only phosphosite comparisons between signature-high and signature-low groups. Multiple-comparison correction used Benjamini–Hochberg FDR where indicated; however, the initial glycosite discovery screen in Section 2.2 followed the criteria used in the original source study and used

raw P values together with an effect-size threshold for candidate selection in the small paired cohort. All tests were two-sided.

3 Results

3.1 Glycosite-level analysis and construction of the LRP1–FN1–LAMC2 signature

Using paired glycosite-level data from a published glycoproteomic dataset^[15], we identified 54 significantly upregulated and 31 significantly downregulated core fucosylation events in PDAC (paired t-test, $P < 0.05$, $|\log_2FC| \geq 0.585$) from a total of 551 quantified glycosites (Fig. 1a).

Functional enrichment of genes harboring upregulated core fucosylation events revealed several PDAC-related biological processes, including extracellular matrix remodeling, epithelial–mesenchymal transition, cell adhesion, glycosylation-related processes, and growth factor signaling (Fig. 1b). Among these, ECM- and adhesion-associated terms were the most prominent across databases, including HALLMARK Epithelial–Mesenchymal Transition, Reactome Extracellular Matrix Organization, GO extracellular matrix structural programs, and KEGG focal adhesion. Glycosylation-related terms were also retained, supporting the glycobiological relevance of the altered gene set^[23]. In detail, three reported ECM-related CF proteins, LRP1, FN1, and LAMC2 were among the significantly up-regulated events. In addition, three well-known PDAC related pathway, PI3K/AKT-, TGF β -, and KRAS/MAPK, were also significantly enriched in upregulated genes, although they were less significant than ECM-associated programs in this analysis. Similar results were also observed in the database-specific enrichment analysis (Fig. S1). Together, these findings suggested intrinsic links between CF proteins and PDAC ECM environment which points to a direction for downstream analysis.

Based on the above observations, we prioritized ECM-associated candidates using a two-stage strategy that combined candidate-level glycosite evidence with signature-level sensitivity analysis^[29]. In Figure 2a, we summarize the number of significantly upregulated CF sites per candidate protein, whereas Figure 2b shows the individual site-level fold-change distribution for each candidate. LRP1 and FN1 were selected as the two anchor components because they combined greater CF site density with consistently positive tumor-to-normal shifts across multiple sites, providing the most substantial glycosite-level support among the prioritized candidates.

Then we investigated the remaining candidates (LAMC2, THBS2, FLT1, and THBS1), although each was supported by only one CF site (Fig. 2a). We therefore integrated the candidate-level evidence with the enrichment pattern observed in Figure 1b. In this context, LAMC2, a basement membrane laminin component implicated in PDAC invasion and prognosis^[30,31], provided biological complementarity to FN1, a matrix glycoprotein^[32], and LRP1, a membrane endocytic receptor^[33].

We then performed a second round of screening by comparing the sensitivity patterns of each combination in a heatmap (Fig. 2c). Specifically, four three-component variants—each containing LRP1 and FN1 plus one alternative third component—were compared across representative ECM/integrin, TGF β , EGFR/ERBB, PI3K/AKT, and MAPK-related axes. The LAMC2-containing variant showed the most coherent ECM/integrin-related correlation

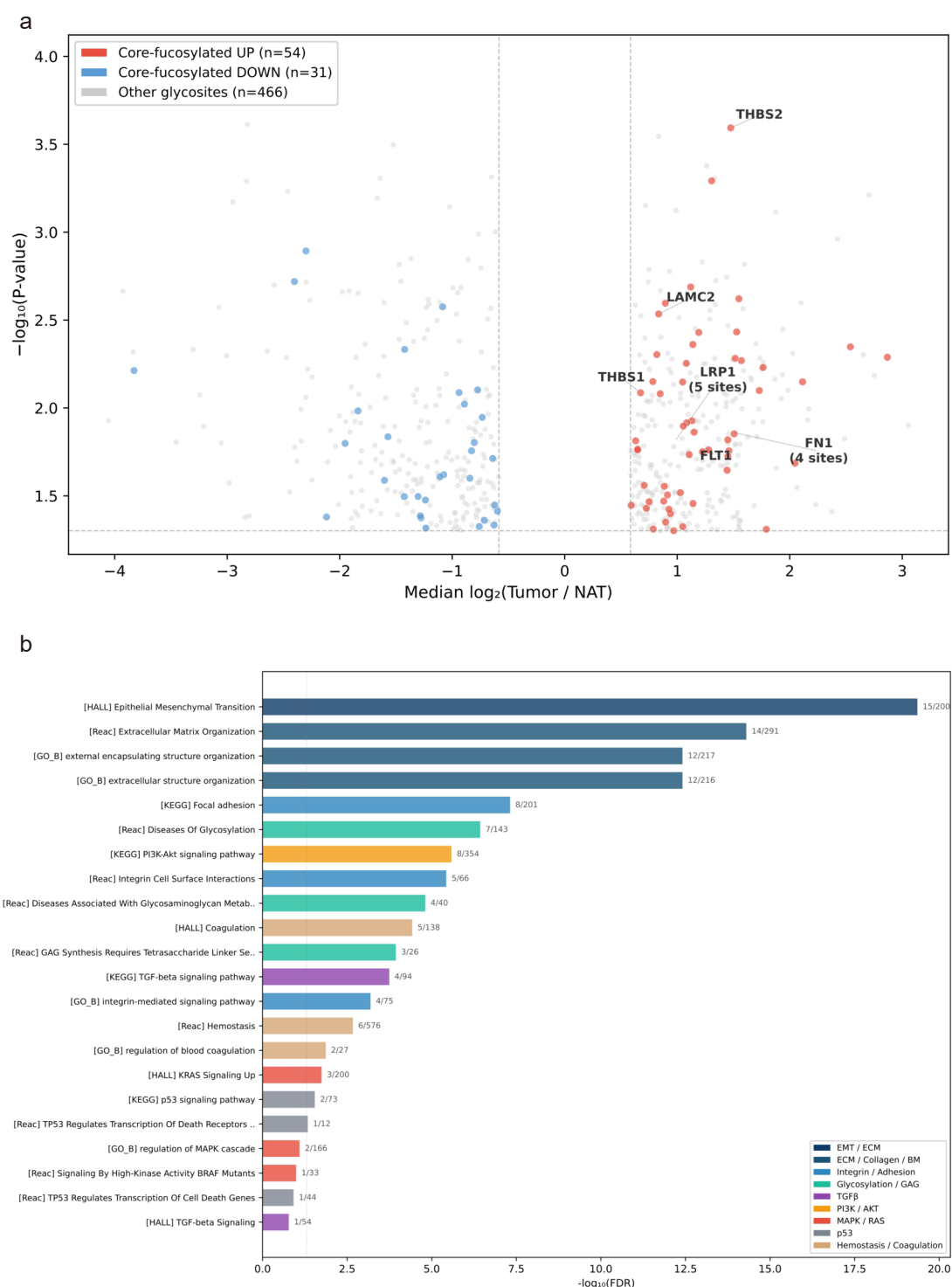


Figure 1 Glycosite-level discovery and functional enrichment of core fucosylation alterations in pancreatic ductal adenocarcinoma (PDAC). (a) Volcano plot of differential core-fucosylated (CF) glycosites in paired tumor and adjacent normal tissues. Red: upregulated (54 events); blue: downregulated (31 events); grey: non-CF events (466). Dashed lines: $|\log_2\text{FC}| \geq 0.585$ and $P < 0.05$. (b) Summary of top enriched terms across HALLMARK, KEGG, Reactome, and Gene Ontology (GO) databases for genes harboring upregulated core fucosylation events. Database-specific enrichment plots are provided in Figure S1.

pattern together with positive TGF β association and the weakest inverse relationship with MEK activation, indicating the most coherent overall alignment with the enrichment-defined biology. Based on these results, the LRP1–FN1–LAMC2 combination was selected as the final core fucosylation-derived signature.

3.2 Validation of the signature using CPTAC proteomic and phosphoproteomic data

Having defined the signature at the glycosite level, we next asked whether it related to tumor-level protein abundance and signaling

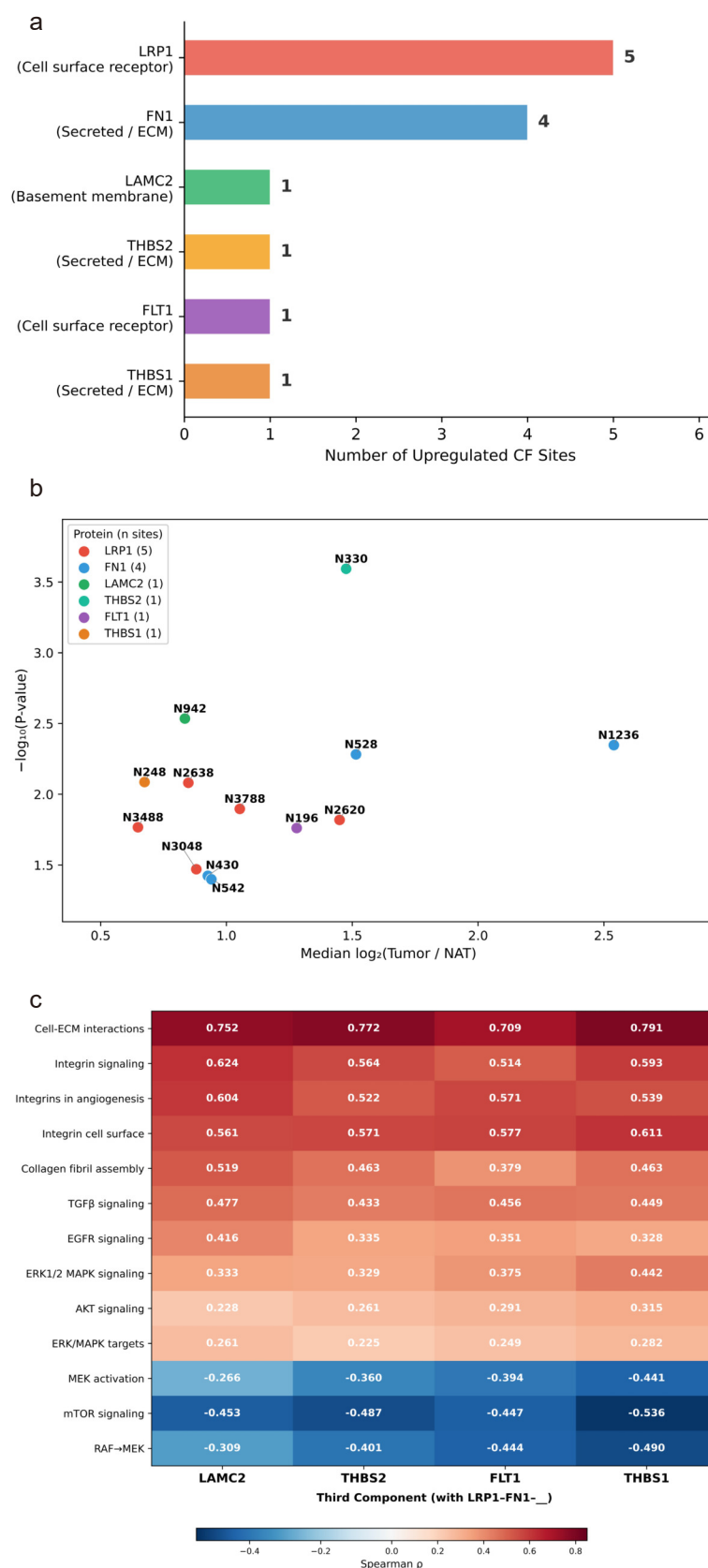


Figure 2 Data-driven derivation of the LRP1–FN1–LAMC2 signature. (a) Number of upregulated core-fucosylated (CF) sites per candidate protein. LRP1 (5 sites) and FN1 (4 sites) showed the highest CF glycosite evidence density. (b) Individual CF site evidence shown as median $\log_2(\text{Tumor/Adjacent normal tissue [NAT]})$ versus $-\log_{10}(\text{P value})$ for prioritized candidates. (c) Sensitivity analysis comparing four three-component variants (LRP1–FN1 + LAMC2 or THBS2 or FLT1 or THBS1) across representative ECM/integrin, TGF β , EGFR/ERBB, PI3K/AKT, and MAPK-related pathways.

activity in independent datasets. For this purpose, we examined the CPTAC PDAC cohort^[17] at two complementary levels: global protein abundance (Fig. 3a) and phosphosite-level support (Fig. 3b).

Figure 3a shows tumor and normal protein abundance differences across representative proteins from ECM/adhesion, TGF β , EGFR/ERBB, PI3K/AKT, and MAPK pathways. Among them, ECM/adhesion proteins showed the most prominent positive shifts in tumors, including LAMC2, COL3A1, THBS1, LOXL2, POSTN, ITGA2, FN1, and LRP1. TGF β 1 was also elevated, whereas EGFR/ERBB- and PI3K/AKT-related proteins showed more modest but still positive support. By contrast, core MAPK proteins displayed comparatively limited abundance differences. This pattern suggests that the most prominent protein-level changes associated with the signature occur in the ECM compartment,

while intracellular kinase components show weaker abundance shifts, consistent with the desmoplastic phenotype of PDAC^[7,9].

Figure 3b summarizes selected phosphosites showing positive support across three complementary criteria: elevation in tumor versus adjacent normal tissue, higher signal in signature-high versus signature-low tumors, and positive correlation with the signature score. Within this framework, TGFBR2 S587 showed positive support consistent with TGF β -related signaling in PDAC, a pathway with context-dependent tumor-suppressive and tumor-promoting roles^[14,19,34]. ERK1 Y204 and ERK2 Y187 were also positively supported, in line with increased ERK1/2 phosphorylation and MAPK pathway activity reported in PDAC^[11,12,35]. EGFR T648 and ERBB2 T671 provided additional support for EGFR/ERBB-associated signaling, consistent with the essential role of EGFR signaling in KRAS-driven PDAC^[20].

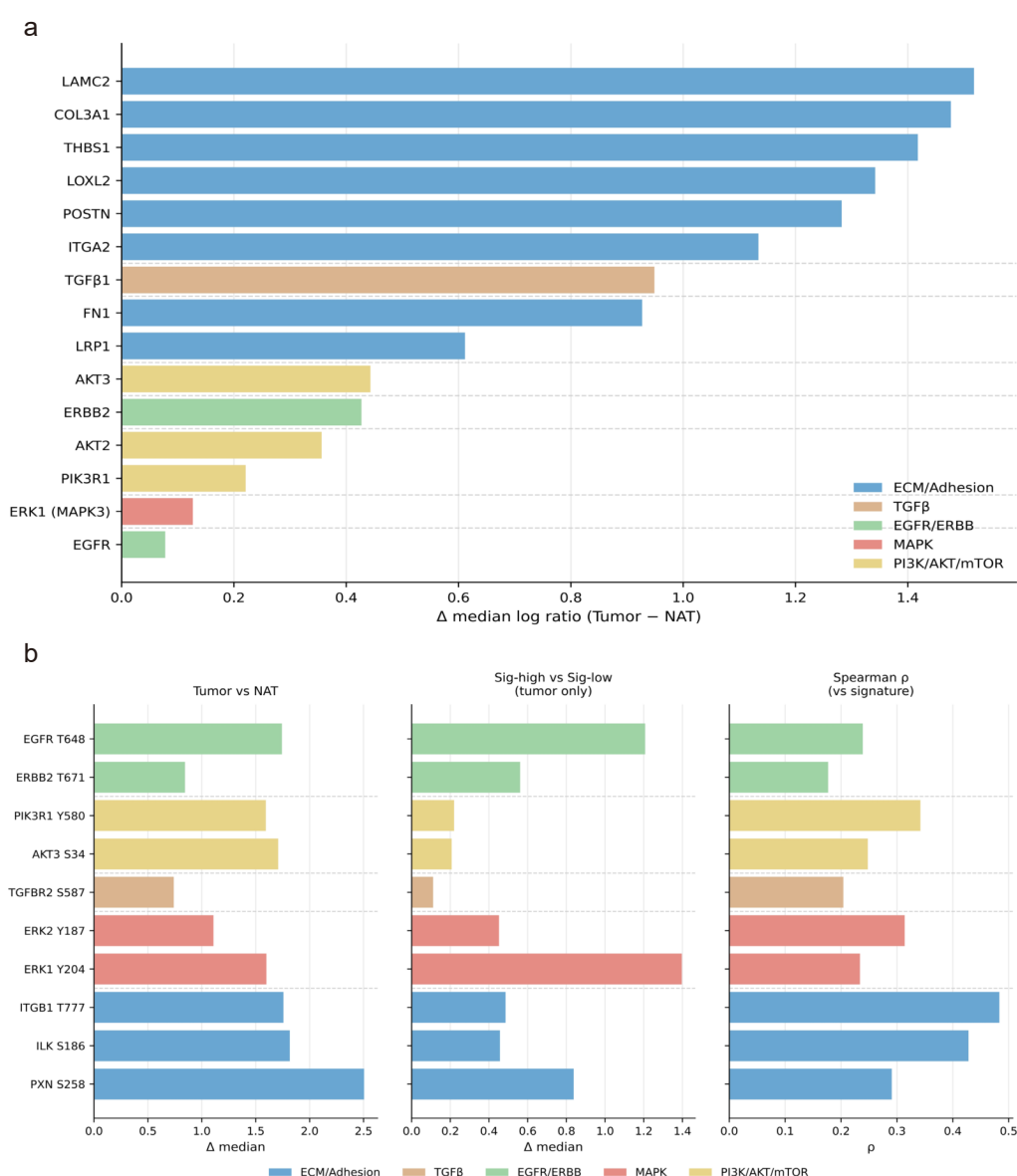


Figure 3 Independent proteomic and phosphoproteomic validation of the LRP1–FN1–LAMC2 signature in the pancreatic ductal adenocarcinoma (PDAC) cohort of Clinical Proteomic Tumor Analysis Consortium (CPTAC). (a) Tumor-versus-normal protein abundance differences across representative ECM/adhesion, TGF β , EGFR/ERBB, PI3K/AKT, and MAPK-related proteins. (b) Selected phosphosites showing positive support across three complementary comparisons: tumor versus adjacent normal tissue (NAT), signature-high versus signature-low tumors, and correlation with the signature score. Complete analysis is provided in Figure S2.

PI3K/AKT-related phosphosites, including PIK3R1 Y580 and AKT3 S34, also showed positive support, consistent with the established importance of PI3K/AKT signaling in PDAC^[21]. ECM-proximal phosphosites showed a more mixed pattern: ITGB1 T777 and PXN S258 were positively associated, whereas ILK S186 showed strong support. Complete phosphosite analysis is provided in Fig. S2.

3.3 Evaluation of the CF signature using transcriptome-based pathway activity

To extend this validation from protein- and phosphosite-level evidence to transcriptome-based pathway activity, we next examined how the signature correlated with broader biological programs in the TCGA-PAAD cohort. Spearman correlations were calculated between the CF signature score and the ssGSEA enrichment scores of 1,387 pathways in the TCGA-PAAD cohort, and representative results are summarized in Figures 4a–c. As shown in Figure 4a, cell–ECM interactions and integrin-related programs showed the strongest positive associations with the CF signature score, whereas TGFβ-, EGFR/ERBB-, and ERK1/2–MAPK-related pathways also retained positive but more moderate correlations. PI3K/AKT-related signaling was represented but showed weaker association in this summary view.

Category-level comparisons further supported this pattern (Fig. 4b). ECM/adhesion pathways showed the highest overall correlation distribution, exceeding MAPK- and PI3K/AKT/mTOR-related categories, while TGFβ-associated pathways also remained positively shifted. These findings are consistent with the ECM-rich biology of PDAC^[6,8] and with evidence that MAPK signaling in this disease is context-dependent^[12,36].

Group-level pathway activity analysis reinforced the same pattern (Fig. 4c). Signature-high tumors showed the largest increase in cell–ECM interaction activity, with additional elevation of integrin-related and TGFβ-related programs. ERK1/2–MAPK activity was also increased but to a more moderate degree, consistent with the limited abundance change of core MEK proteins observed in CPTAC. Together, these results suggest that the signature is most strongly aligned with ECM/integrin-centered biology while retaining coordinated associations with TGFβ and context-dependent MAPK output.

3.4 Evaluation of the CF signature using tumor and normal comparison

We next evaluated the CF signature score and related pathway activities in three independent cohorts. Figure 5a shows PDAC tumor and normal separation across three GEO datasets and Figure 5b summarizes pathway activities associated with the signature across cohorts.

PDAC tumor and normal tissues can be well separated by the CF signature score across all three GEO cohorts (Fig. 5a), with AUCs of 0.941 in GSE15471, 0.906 in GSE28735, and 0.888 in GSE62452. It shows the robustness of the CF signature in separating PDAC tumor and normal tissues. Then we compared activities of PDAC-related pathways in CF signature score-high and signature score-low groups across the three GEO cohorts (Fig. 5b). The exact gene compositions and scoring definitions of these proxy sets were provided in Table S2. ECM/adhesion-related pathway scores showed the most consistent group separation across datasets, supporting the ECM-centered nature of the signature^[6,8]. EGFR/ERBB-related signals were also reproducibly retained,

consistent with prior reports that EGFR signaling remains functionally important in KRAS-driven PDAC^[30]. TGFβ- and MAPK-related associations were present but more variable across cohorts, in keeping with the context-dependent behavior of these pathways in PDAC^[19,36]. By contrast, PI3K/AKT/mTOR-related proxy signals were weaker and less consistent across cohorts, in line with the weaker pathway-level associations observed in TCGA. Together, these analyses supported the robustness of the CF signature in separating tumor and normal tissues and correlation with PDAC related pathways.

3.5 The CF signature is associated with poor outcome in PDAC

We next examined whether this signature was associated with patient outcome in the TCGA-PAAD cohort. Prognostic evaluation was performed at three complementary levels: survival analysis by signature alone (Fig. 6a and 6b), stratification by signature and AJCC-based stage regrouping (Fig. 6c and 6d), and time-dependent discrimination in a Cox model combining signature score with stage-related variables (Fig. 6e).

Survival analysis of the 183 TCGA-PAAD patients showed that progression-free interval (PFI) was significantly shorter in signature-high tumors (log-rank $P = 0.044$; Fig. 6b). Overall survival (OS) displayed the same directional pattern but did not reach conventional significance (log-rank $P = 0.085$; Fig. 6a). Two cases were not included in the four-group stratified analysis because of missing stage information. The consistent directionality supports a potentially prognostic association, although the modest sample size of TCGA-PAAD remains a recognized limitation for survival analysis in this disease^[9,37].

To test whether the signature added prognostic resolution beyond broad clinical grouping, patients were further stratified by signature status and AJCC-based stage regrouping. Stage was regrouped as I+IIA versus IIB+III+IV. This cutoff reflects the transition from node-negative stage IIA to node-positive stage IIB in the AJCC 8th staging framework and provides a clinically interpretable boundary. Prior PDAC validation studies also support the prognostic importance of nodal category^[38,39]. In this four-group model, both OS (log-rank $P = 0.048$; Fig. 6c) and PFI (log-rank $P = 0.023$; Fig. 6d) were significantly separated. Patients with high signature and IIB/III/IV disease had the poorest prognosis, whereas those with low signature and I/IIA disease had the most favorable outcome. Together, these findings suggest that the ECM-centered signature adds prognostic value within stage context.

We next assessed whether the signature also contributed to time-dependent prediction. In a Cox model incorporating continuous signature score ($\beta = 0.114$) and binary AJCC-based stage regrouping ($\beta = 0.493$), time-dependent AUC increased from 0.527 at 1 year to 0.699 at 3 years and 0.837 at 5 years (Fig. 6e). The modest 1-year AUC is compatible with the generally poor short-term survival landscape of PDAC^[9,37], whereas the higher 5-year AUC suggests improved discrimination for longer-term outcome. Together, these findings indicate that the LRP1–FN1–LAMC2 signature is associated with adverse prognosis and may capture durable biology linked to late clinical outcome, consistent with prior glycosylation-related prognostic signatures reported in PDAC^[40].

4 Discussion

In this study, we developed a core fucosylation-derived signature

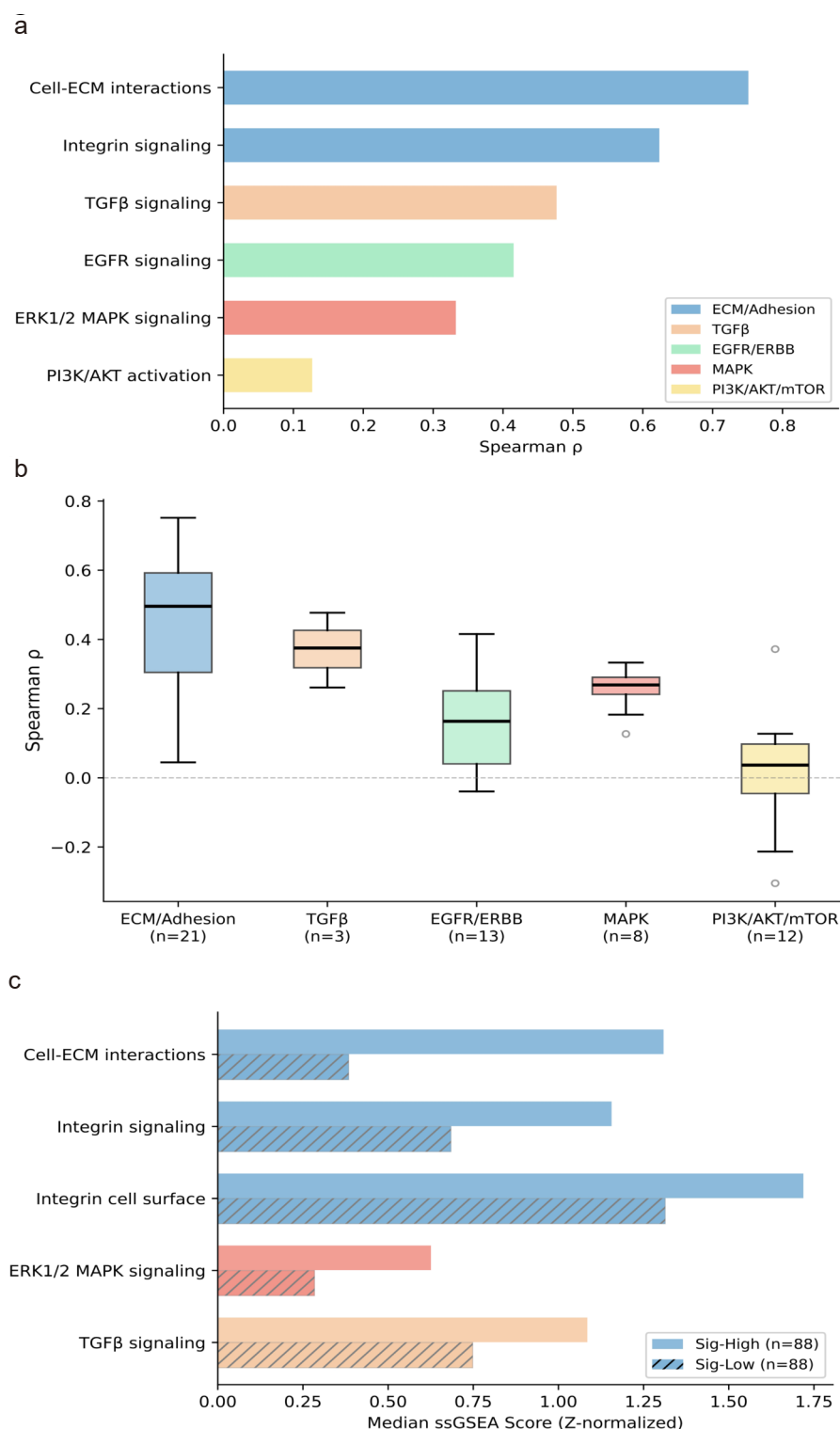


Figure 4 Transcriptome-wide pathway correlation analysis of the signature in The Cancer Genome Atlas Pancreatic Adenocarcinoma Database (TCGA-PAAD). (a) Correlations of the signature with representative pancreatic ductal adenocarcinoma (PDAC)-related pathways. (b) Category-level distributions of signature-pathway correlations. (c) Differences in pathway activity between signature-high and signature-low tumors (median split).

including three CF glycoproteins (LRP1, FN1, and LAMC2) in PDAC and found that it is strongly correlated with ECM/integrin-centered biology. Moreover, it also correlated with several critical pathways in PDAC progression, including TGF β receptor signaling, EGFR/ERBB signaling, and ERK signaling.

CPTAC proteomic data showed that ECM proteins were significantly upregulated in tumors, whereas core MAPK proteins displayed comparatively limited abundance changes^[17]. Transcriptome-wide pathway analysis further showed that cell-ECM interaction programs had the strongest pathway-level

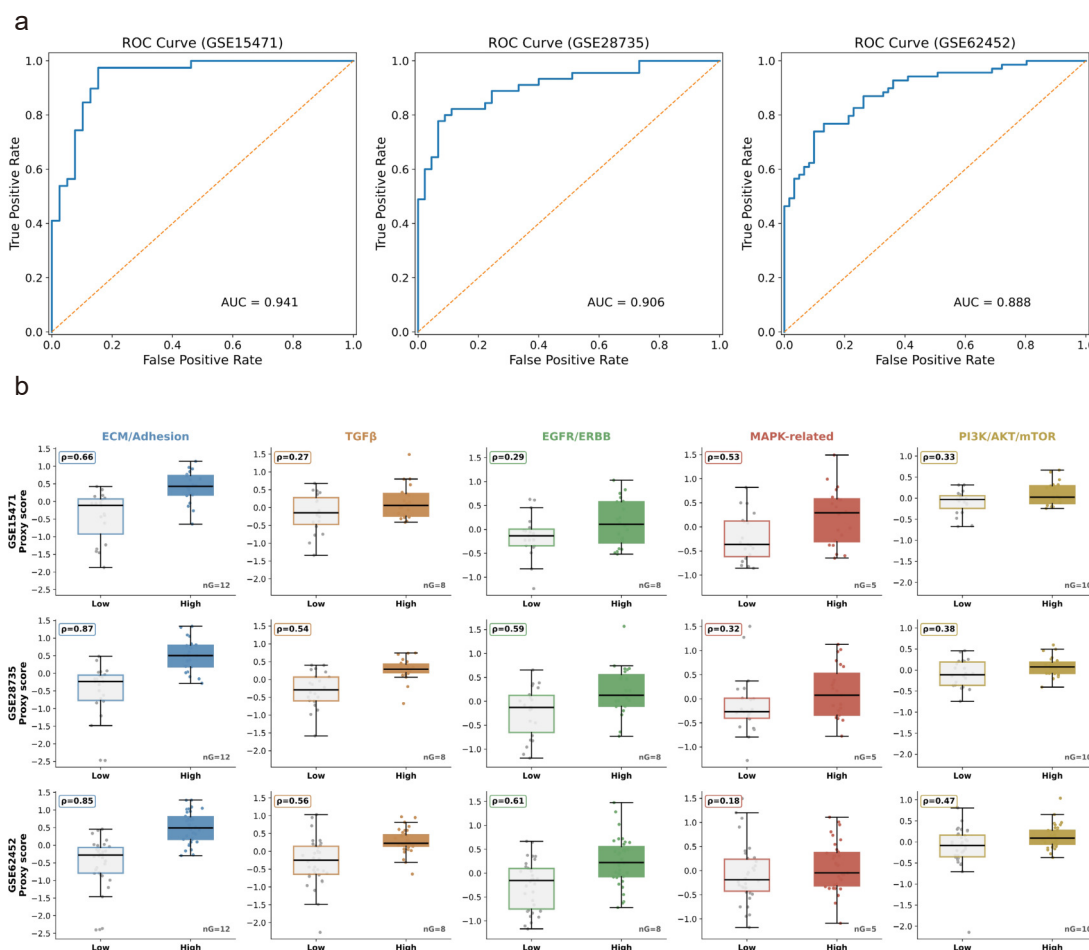


Figure 5 External validation of the signature and pathway proxy states in independent Gene Expression Omnibus (GEO) cohorts. (a) Receiver operating characteristic (ROC) curves showing the performance of the signature in separation of pancreatic ductal adenocarcinoma (PDAC) tumor samples from normal controls derived from the cohorts of GSE15471, GSE28735, and GSE62452. (b) External validation of pathway proxy scores between low- and high-signature tumor groups in the three GEO cohorts, including ECM/adhesion, TGF β , EGFR/ERBB, MAPK-related, and PI3K/AKT/mTOR pathways. The exact gene lists and proxy-score definitions are provided in Table S2.

association with the signature. ECM pathways consistently showed stronger category-level association than MAPK pathways. Validation in three GEO cohorts confirmed the robustness of the signature in tumor and normal separation and showed that ECM/adhesion was the most reproducible pathway. These findings are consistent with the existing knowledge of PDAC, in which ECM remodeling and stromal interactions drive disease progression^[6-8]. In detail, Liot et al. have emphasized that the ECM in PDAC is not merely a structural scaffold but an active driver of tumor cell behavior^[8], and Masugi has highlighted the multilayered heterogeneity of the desmoplastic stroma^[7]. The ECM-dominant pattern observed here is consistent with these characterizations. Garcia et al. further showed that fibroblast heterogeneity within the PDAC stroma contributes to differential ECM deposition patterns^[41].

In addition, phosphoproteomic data provided more insights into these observations. TGFBR2 phosphorylation provided phosphosite-level support for the signature and linked the ECM-associated program to TGF β receptor activity. This connects the ECM program to TGF β receptor activation. TGF β plays dual roles in PDAC depending on SMAD4 status: tumor-suppressive in early disease and pro-fibrotic/pro-invasive in advanced tumors^[14,19].

Zhang et al. have shown that cancer-associated fibroblasts (CAFs) are major mediators of TGF β signaling in the PDAC microenvironment^[19], while Ying et al. placed TGF β within the broader genetic framework of PDAC progression^[14]. The association between TGFBR2 phosphorylation and the ECM-derived signature is consistent with these reports. EGFR T648 and ERBB2 T671 were also positively associated, while EGFR/ERBB-related signals were reproducibly retained across external cohorts. Navas et al. demonstrated that EGFR signaling is essential for KRAS-driven pancreatic tumorigenesis^[20], and our data support this relationship at the phosphoproteomic level. The observation that ERK activation sites were elevated whereas MEK activation scores were negatively correlated suggests context-dependent, non-canonical MAPK signaling. Adamopoulos et al. have reviewed the complex RAF/MEK/ERK landscape in PDAC^[12]. Veghini et al. further showed that differential MAPK activity defines fibroblast subtypes in pancreatic cancer^[36]. Together, these observations are consistent with a context-dependent MAPK output linked to the signature.

From a clinical perspective, the signature was associated with adverse outcome. PFI reached significance and four-group stratification demonstrated an additive effect with AJCC-based stage regrouping. Of note, OS displayed the same trend but did not

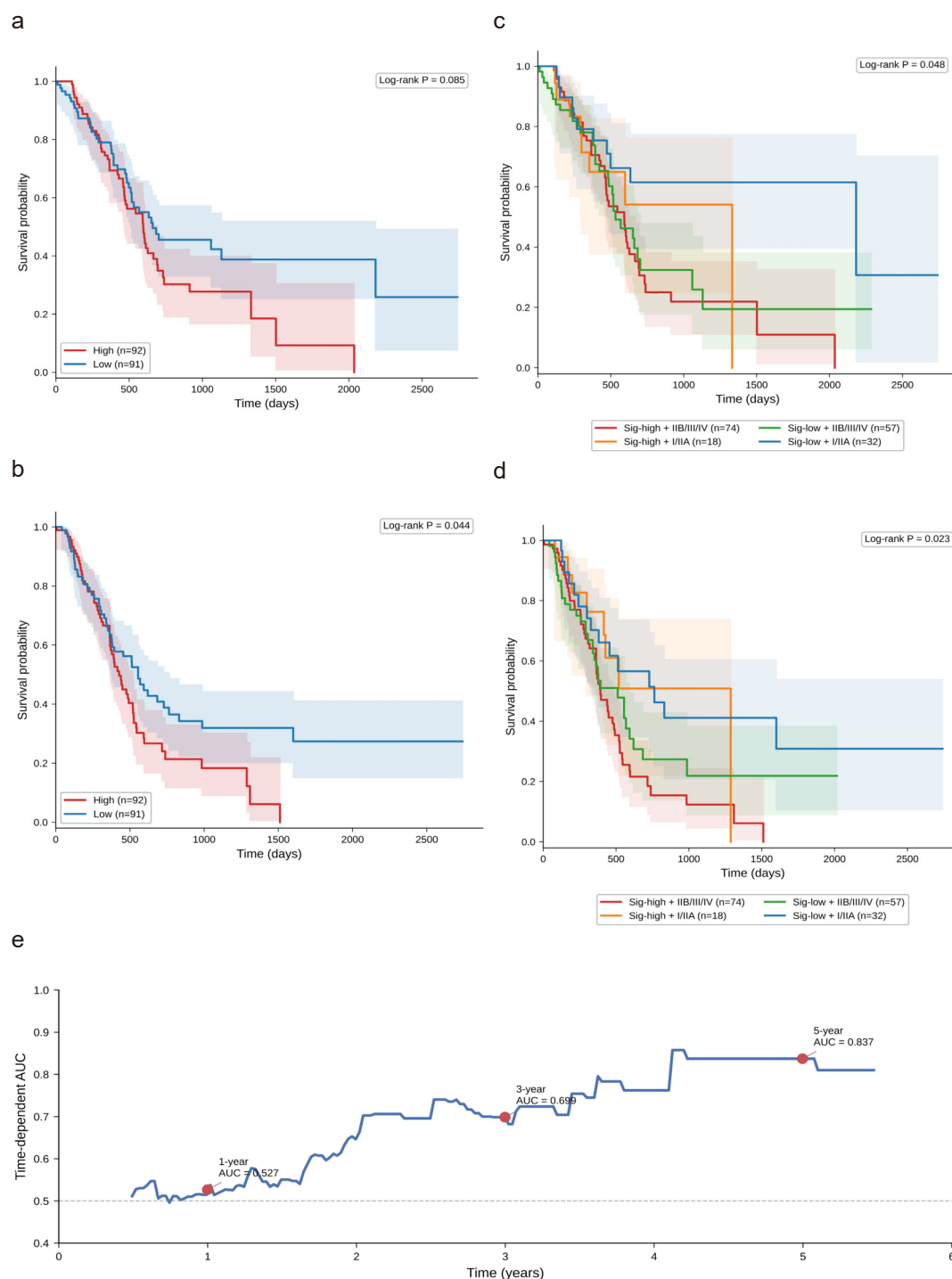


Figure 6 Prognostic evaluation of the signature in The Cancer Genome Atlas Pancreatic Adenocarcinoma Database (TCGA-PAAD). Kaplan–Meier curves for overall survival (OS) (a) and progression-free interval (PFI) (b) by median signature split. Four-group stratification for OS (c) and PFI (d) by signature and American Joint Committee on Cancer (AJCC)-based stage regrouping (I + IIA vs IIB + III + IV). (e) Time-dependent area under the curve (AUC, OS) for a Cox model integrating continuous signature score with stage-related variables.

reach statistical significance. This was likely partially because the assessment of OS was confounded by diverse post-progression palliative therapies, whereas PFI may better reflect the intrinsic aggressiveness of the tumor^[42]. The performance of the signature in separation of PDAC tumor and normal samples across the three GEO cohorts remained robust. The increase in AUC from 1 year to 5 years suggests that the signature may capture durable ECM-

related biology associated with longer-term clinical outcome. These observations are compatible with prior glycosylation-related prognostic signatures in PDAC^[40]. Hu et al. developed a glycosylation-related gene signature for PDAC survival prediction^[40], and our work extends this direction by anchoring the signature to core fucosylation events and ECM biology. The prognostic utility of basement membrane laminins in PDAC has

been reported by van der Zee et al.^[31], further supporting the clinical relevance of the LAMC2 component.

Several translational directions may be considered based on the current findings. Among the candidate translational directions, ECM/integrin-targeted approaches appear most directly supported by the current multi-omics findings. This may involve integrin inhibitors or LOX/LOXL2 inhibitors targeting the ECM–receptor interface^[6,43]. Shihadeh et al. have reviewed stromal-targeting strategies in PDAC, including ECM-directed approaches^[43]. TGFβ receptor targeting is also supported by the phosphosite-level association of TGFBR2 and by the emerging understanding of CAF-mediated TGFβ signaling in PDAC^[19]. ERK/MAPK-directed strategies may be relevant in combination settings, particularly because adaptive rewiring has been implicated in resistance to KRAS inhibition in PDAC^[12,44]. All three directions are hypothesis-generating and require functional validation. In addition, FUT8 itself has also been proposed as a potential therapeutic target^[45], offering a possible complementary direction.

Several limitations should be acknowledged. First, the discovery was based on publicly available glycosite data, and the signature was derived through biologically constrained prioritization rather than prospective screening. Second, the signature measures abundance rather than glycosylation level. Abundance and glycosylation are not equivalent, although intermediate abundance validation confirmed tumor-associated elevation. This distinction between abundance and glycosylation is a known challenge in glycoproteomic studies^[2,15]. Nevertheless, the abundance of LRP1, FN1, and LAMC2 may still serve as a practical surrogate for the underlying CF-associated program because these molecules are themselves ECM-related glycoproteins and central components of the tumor–stroma interface. Notably, prior work has shown that loss of core fucosylation impairs LRP1 function, supporting the biological relevance of core fucosylation for at least one component of the present signature^[46]. Thus, increased abundance in external cohorts may reflect the same matrix-remodeling context in which their CF forms were initially identified, although this cross-cohort validation does not directly measure glycosylation status and future targeted glycoproteomic studies will be needed for direct confirmation. More broadly, the present study should be interpreted as defining a core fucosylation-derived signature rather than directly validating the same site-specific glycosylation events across all external cohorts (Fig. S3). From a translational perspective, the current findings may be more immediately applicable as an expression- or protein-based risk stratification framework, whereas direct clinical deployment of glycosylation-specific measurements will require future targeted glycoproteomic validation in independent PDAC cohorts. Third, functional validation (FUT8 knockdown, integrin blocking) was not performed. Fourth, candidate selection included biologically informed but partly manual elements, although sensitivity analysis provided quantitative support. Fifth, survival analysis was limited to a single TCGA cohort. Sixth, GEO pathway proxy scores used curated small gene sets rather than standard ssGSEA and should be interpreted as supportive evidence.

In summary, we developed a core fucosylation-derived signature related to ECM remodeling, TGFβ receptor signaling, EGFR/ERBB activity, and ERK-related pathways in PDAC. This signature was validated across multi-omics layers in five independent large PDAC cohorts including TCGA PAAD, CPTAC PDAC, and three GEO data sets. These analyses support an important association between

the core fucosylation-derived signature and PDAC progression. This work provides a framework for future investigation of the interface between glycosylation, ECM remodeling, and other signaling pathways in pancreatic cancer.

Electronic Supplementary Material

The online version contains supplementary material available at <https://doi.org/10.26599/OSHM.2026.9610054>.

Acknowledgments

This work was funded by the National Natural Science Foundation of China (22374063 to L.C.) and Shenzhen Science and Technology Program (ZDSYS20230626090803004).

Data available statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Author contribution

L.C. and R.T. conceived the project and designed the experiments. B.Z. and J.C. analyzed the data. B.Z. and L.C. wrote the manuscript.

Ethics approval and consent

N/A

Consent for publication

All authors agree to publish.

Conflicts of interest

The authors declare no conflicts of interest.

References

- [1] He, M. Y.; Zhou, X. X.; Wang, X. Glycosylation: Mechanisms, biological functions and clinical implications. *Signal Transduct. Target. Ther.* **2024**, *9*, 194.
- [2] Liao, C. C.; An, J. X.; Yi, S. Q.; et al. FUT8 and protein core fucosylation in tumours: From diagnosis to treatment. *J. Cancer* **2021**, *12*, 4109–4120.
- [3] Munkley, J. The glycosylation landscape of pancreatic cancer (Review). *Oncol. Lett.* **2019**, *17*, 2569–2575.
- [4] Terao, N. Fucosylation is a common glycosylation type in pancreatic cancer stem cell-like phenotypes. *World J. Gastroenterol.* **2015**, *21*, 3876.
- [5] Tada, K.; Ohta, M.; Hidano, S.; et al. Fucosyltransferase 8 plays a crucial role in the invasion and metastasis of pancreatic ductal adenocarcinoma. *Surg. Today* **2020**, *50*, 767–777.
- [6] Perez, V. M.; Kearney, J. F.; Yeh, J. J. The PDAC extracellular matrix: A review of the ECM protein composition, tumor cell interaction, and therapeutic strategies. *Front. Oncol.* **2021**, *11*, 751311.
- [7] Masugi, Y. The desmoplastic stroma of pancreatic cancer: Multilayered levels of heterogeneity, clinical significance, and therapeutic opportunities. *Cancers* **2022**, *14*, 3293.
- [8] Liot, S.; Balas, J.; Aubert, A.; et al. Stroma involvement in pancreatic ductal adenocarcinoma: An overview focusing on extracellular matrix proteins. *Front. Immunol.* **2021**, *12*, 612271.

- [9] Halbrook, C. J.; Lyssiotis, C. A.; Pasca di Magliano, M.; et al. Pancreatic cancer: Advances and challenges. *Cell* **2023**, *186*, 1729–1754.
- [10] Nusrat, F.; Khanna, A.; Jain, A.; et al. The clinical implications of KRAS mutations and variant allele frequencies in pancreatic ductal adenocarcinoma. *J. Clin. Med.* **2024**, *13*, 2103.
- [11] Eser, S.; Schnieke, A.; Schneider, G.; et al. Oncogenic KRAS signalling in pancreatic cancer. *Br. J. Cancer* **2014**, *111*, 817–822.
- [12] Adamopoulos, C.; Cave, D. D.; Papavassiliou, A. G. Inhibition of the RAF/MEK/ERK signaling cascade in pancreatic cancer: Recent advances and future perspectives. *Int. J. Mol. Sci.* **2024**, *25*, 1631.
- [13] Choucair, K.; Imtiaz, H.; Uddin, M. H.; et al. Targeting KRAS mutations: Orchestrating cancer evolution and therapeutic challenges. *Signal Transduct. Target. Ther.* **2025**, *10*, 385.
- [14] Ying, H. Q.; Kimmelman, A. C.; Bardeesy, N.; et al. Genetics and biology of pancreatic ductal adenocarcinoma. *Genes Dev.* **2025**, *39*, 36–63.
- [15] Cao, L. W.; Lih, T. M.; Hu, Y. W.; et al. Characterization of core fucosylation via sequential enzymatic treatments of intact glycopeptides and mass spectrometry analysis. *Nat. Commun.* **2022**, *13*, 3910.
- [16] Goldman, M. J.; Craft, B.; Hastie, M.; et al. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat. Biotechnol.* **2020**, *38*, 675–678.
- [17] Cao, L. W.; Huang, C.; Cui Zhou, D.; et al. Proteogenomic characterization of pancreatic ductal adenocarcinoma. *Cell* **2021**, *184*, 5031–5052.e26.
- [18] Vasaikar, S. V.; Straub, P.; Wang, J.; et al. LinkedOmics: Analyzing multi-omics data within and across 32 cancer types. *Nucleic Acids Res.* **2018**, *46*, D956–D963.
- [19] Zhang, T. Y.; Ren, Y. X.; Yang, P. F.; et al. Cancer-associated fibroblasts in pancreatic ductal adenocarcinoma. *Cell Death Dis.* **2022**, *13*, 897.
- [20] Navas, C.; Hernández-Porras, I.; Schuhmacher, A. J.; et al. EGF receptor signaling is essential for K-ras oncogene-driven pancreatic ductal adenocarcinoma. *Cancer Cell* **2012**, *22*, 318–330.
- [21] Conway, J. R.; Herrmann, D.; Evans, T. J.; et al. Combating pancreatic cancer with PI3K pathway inhibitors in the era of personalised medicine. *Gut* **2019**, *68*, 742–758.
- [22] Fang, Z. Q.; Liu, X. Y.; Peltz, G. GSEapy: A comprehensive package for performing gene set enrichment analysis in Python. *Bioinformatics* **2023**, *39*, btac757.
- [23] Kuleshov, M. V.; Jones, M. R.; Rouillard, A. D.; et al. Enrichr: A comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **2016**, *44*, W90–W97.
- [24] Barbie, D. A.; Tamayo, P.; Boehm, J. S.; et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* **2009**, *462*, 108–112.
- [25] Pedregosa, F.; Varoquaux, G.; Gramfort, A.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
- [26] Chun, Y. S.; Pawlik, T. M.; Vauthey, J. N. 8th edition of the AJCC cancer staging manual: Pancreas and hepatobiliary cancers. *Ann. Surg. Oncol.* **2018**, *25*, 845–847.
- [27] Uno, H.; Cai, T. X.; Tian, L.; et al. Evaluating prediction rules for t-year survivors with censored regression models. *J. Am. Stat. Assoc.* **2007**, *102*(478), 527–537.
- [28] Pölsterl, S. Scikit-survival: A library for time-to-event analysis built on top of scikit-learn. *J. Mach. Learn. Res.* **2020**, *21*(1), 8747–8752.
- [29] Cima, I.; Schiess, R.; Wild, P.; et al. Cancer genetics-guided discovery of serum biomarker signatures for diagnosis and prognosis of prostate cancer. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 3342–3347.
- [30] Huang, C. Q.; Chen, J. Laminin-332 mediates proliferation, apoptosis, invasion, migration and epithelial-to-mesenchymal transition in pancreatic ductal adenocarcinoma. *Mol. Med. Rep.* **2020**, *23*, 1.
- [31] van der Zee, J. A.; van Eijck, C. H. J.; Hop, W. C. J.; et al. Tumour basement membrane laminin expression predicts outcome following curative resection of pancreatic head cancer. *Br. J. Cancer* **2012**, *107*, 1153–1158.
- [32] Topalovski, M.; Brekken, R. A. Matrix control of pancreatic cancer: New insights into fibronectin signaling. *Cancer Lett.* **2016**, *381*, 252–258.
- [33] Yamamoto, K.; Scilabra, S. D.; Bonelli, S.; et al. Novel insights into the multifaceted and tissue-specific roles of the endocytic receptor LRP1. *J. Biol. Chem.* **2024**, *300*, 107521.
- [34] Zhang, H.; Liu, C. L.; Kong, Y. L.; et al. TGFβ signaling in pancreatic ductal adenocarcinoma. *Tumor Biol.* **2015**, *36*, 1613–1618.
- [35] Sun, H. Z.; Mo, J.; Cheng, R. F.; et al. ENO1 expression and Erk phosphorylation in PDAC and their effects on tumor cell apoptosis in a hypoxic microenvironment. *Cancer Biol. Med.* **2022**, *19*, 1598–1616.
- [36] Veghini, L.; Pasini, D.; Fang, R.; et al. Differential activity of MAPK signalling defines fibroblast subtypes in pancreatic cancer. *Nat. Commun.* **2024**, *15*, 10534.
- [37] Park, W.; Chawla, A.; O'Reilly, E. M. Pancreatic cancer: A review. *JAMA.* **2021**, *326*, 851.
- [38] Amin, M. B.; Greene, F. L.; Edge, S. B.; et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more “personalized” approach to cancer staging. *CA A Cancer J. Clin.* **2017**, *67*, 93–99.
- [39] van Roessel, S.; Kasumova, G. G.; Verheij, J.; et al. International validation of the eighth edition of the American joint committee on cancer (AJCC) TNM staging system in patients with resected pancreatic cancer. *JAMA Surg.* **2018**, *153*, e183617.
- [40] Hu, H. D.; He, B. S.; He, M. G.; et al. A glycosylation-related signature predicts survival in pancreatic cancer. *Aging* **2023**, *15*, 13710–13737.
- [41] Garcia, P. E.; Scales, M. K.; Allen, B. L.; et al. Pancreatic fibroblast heterogeneity: From development to cancer. *Cells* **2020**, *9*, 2464.
- [42] Kasuga, A.; Hamamoto, Y.; Takeuchi, A.; et al. Positive relationship between subsequent chemotherapy and overall survival in pancreatic cancer: Meta-analysis of postprogression survival for first-line chemotherapy. *Cancer Chemother. Pharmacol.* **2017**, *79*, 595–602.
- [43] Shihadeh, H.; Yousef, A.; Al-Leimon, A.; et al. Targeting stromal components in pancreatic ductal adenocarcinoma: A review. *ecancermedicalscience* **2025**, *19*, 1996.
- [44] Dilly, J.; Hoffman, M. T.; Abbassi, L.; et al. Mechanisms of resistance to oncogenic KRAS inhibition in pancreatic cancer. *Cancer Discov.* **2024**, *14*, 2135–2161.
- [45] Mao, C.; Li, J.; Feng, L. L.; et al. Beyond antibody fucosylation: α-(1, 6)-fucosyltransferase (Fut8) as a potential new therapeutic target for cancer immunotherapy. *Antib. Ther.* **2023**, *6*, 87–96.
- [46] Lee, S. H.; Takahashi, M.; Honke, K.; et al. Loss of core fucosylation of low-density lipoprotein receptor-related protein-1 impairs its function, leading to the upregulation of serum levels of insulin-like growth factor-binding protein 3 in Fut8^{-/-} mice. *J. Biochem.* **2006**, *139*, 391–398.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the original author(s) and the source, provide a link to the license, and indicate if changes were made. See <https://creativecommons.org/licenses/by/4.0/>

© The author(s) 2026. Published by Tsinghua University Press.