

RESEARCH

Open Access



Multi-omics integration and machine learning uncover molecular basal-like subtype of pancreatic cancer and implicate A2ML1 in promoting tumor epithelial-mesenchymal transition

Jiachen Ge^{1†}, Jianping Cai^{1†}, Gaolei Zhang¹, Deyu Li^{1*} and Lianyuan Tao^{1,2*} 

Abstract

Background Pancreatic cancer (PC) represents a highly heterogeneous malignancy with poor prognosis, where precise molecular subtyping facilitates comprehensive understanding of disease progression.

Methods Transcriptomic, methylation, and mutational data from 168 PC samples were integrated. The molecular subtypes were established and validated across 13 independent cohorts utilizing ten distinct classification methods. Prognostic genes were identified to construct predictive models through various machine learning approaches. Following the identification of A2ML1 as a key gene, its expression profile was detected using RT-qPCR, western blotting, and immunohistochemistry. We subsequently conducted functional experiments to elucidate the mechanism of A2ML1 in promoting PC progression.

Results Through multi-omics integration, we classified PC into two molecular subtypes with distinct prognostic outcomes. Pathway enrichment analysis revealed subtype-specific characteristics, subsequently validated in external cohorts. Using 23 prognostic genes, we developed and validated a prognostic signature through 101 machine learning algorithms and their combinations, with ridge regression demonstrating optimal performance. This signature demonstrated superior accuracy compared to multiple published signatures. The risk scores showed significant correlations with drug sensitivity, clinical characteristics, and patient outcomes. We further validated that A2ML1 expression was significantly elevated in PC tissues compared to normal counterparts. A2ML1 promoted PC progression through downregulation of LZTR1 expression and subsequent activation of the KRAS/MAPK pathway, ultimately driving epithelial-mesenchymal transition (EMT).

[†]Jiachen Ge and Jianping Cai contributed equally to this work.

*Correspondence:

Deyu Li
lideyu19@hotmail.com
Lianyuan Tao
tly2007tly@hotmail.com

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Conclusion In this study, we uncovered molecular basal-like subtype of PC and developed a prognostic signature using ridge regression. A2ML1 was identified as a crucial regulator of EMT in PC progression.

Keywords Pancreatic cancer, Molecular subtype, Machine learning, Epithelial-mesenchymal transition, KRAS/MAPK pathway

Introduction

Pancreatic cancer (PC) is one of the most commonly diagnosed cancer worldwide, with approximately 495,773 new cases reported in 2020 [1]. It is the 7th leading cause of cancer-related deaths, accounting for 466,003 deaths globally in 2020. Notably, PC has one of the lowest survival rates among all cancers, with a 5-year survival rate of less than 10% [1]. PC is found to have global disparities. In Europe and North America, the incidence rate is 7–9 per 100,000, while in Africa, it is 2–3 per 100,000, but survival rates are significantly lower in the latter [2]. In brief, high-income countries report higher incidence rates, but low- and middle-income countries face higher mortality due to limited access to early diagnosis and treatment. PC imposes a significant economic burden due to its late diagnosis and high treatment costs [3]. In the United States alone, the annual cost of PC care is estimated to exceed \$2.6 billion [3]. For each person, the risk factors of PC have been known, including obesity, diabetes, chronic pancreatitis, et al. [4]. Genetic predisposition also plays a significant role in PC development [4].

At present, clinical assessment of PC prognosis primarily relies on TNM staging and pathological examination [5, 6]. Nevertheless, PC is characterized by highly heterogeneous, both intertumorally and intratumorally [7]. Molecular subtyping of PC is a transformative approach that categorizes tumors based on their genetic, transcriptomic, proteomic, and epigenetic profiles [8]. Precise molecular classification is of paramount importance for the diagnosis and therapeutic management of PC. First, it helps to unravel their unique biological behaviors, such as differential responses to therapy, metastatic potential, and prognosis [9]. Second, molecular subtyping enables the identification of specific therapeutic targets and biomarkers for each subtype [10]. This approach moves away from a “one-size-fits-all” treatment strategy toward precision medicine. Third, molecular subtype helps us predict disease progression and tailor follow-up strategies. Mofitt et al. indicated that the “basal-like” and “squamous” subtype were linked to aggressive behavior and poor survival, while the “classical” subtype tended to have a better prognosis [11]. Finally, researchers can develop targeted therapies and design clinical trials for specific patient populations by identifying subtype-specific molecular pathways [12]. This approach increases the likelihood of therapeutic success and reduces the risk of adverse effects in non-responsive patients.

The Cancer Genome Atlas (TCGA) has released comprehensive genomic, transcriptomic, and epigenomic sequencing data for PC, significantly facilitating molecular subtyping of this malignancy. Previous studies have proposed various molecular subtype approaches for PC, including methodologies based on PDL1 and CD8 expression patterns, as well as those focusing on tumor immune microenvironment characteristics [13, 14]. However, these classification systems are predominantly based on mRNA expression profiles, which have inherent limitations in accuracy. DNA methylation patterns of key genes play critical roles in PC progression and may serve as valuable diagnostic and prognostic biomarkers [15]. The implementation of multi-omics-based classification frameworks could significantly enhance the precision of molecular subtyping and facilitate deeper biological understanding of PC pathogenesis [16]. In this study, we integrated mRNA, lncRNA, and microRNA (miRNA) expression profiles, genomic mutations, and epigenomic DNA methylation data to develop a novel integrated consensus subtype of PC. We combined 10 multi-omics integration strategies to explore a stable subtype. Subsequently, we identified prognosis-related genes through a comprehensive analysis of 13 independent cohorts. Leveraging these markers, we systematically constructed prognostic models using an ensemble approach that incorporated 10 distinct machine learning algorithms and 101 algorithmic combinations. This robust computational framework enabled us to identify critical genes, which were subsequently subjected to functional validation through experimental investigations. Our study significantly contributes to the refinement of molecular subtyping and precise stratification in PC, thereby establishing a crucial foundation for the development of personalized therapeutic strategies.

Methods

Data acquisition

We firstly collected multi-omics data from the TCGA Pancreatic Cancer (PAAD) cohort (<https://xenabrowse.r.net/datapages/>), including patients with complete transcriptome expression (mRNA, miRNA, and lncRNA), DNA methylation, somatic mutations, and clinical information available. A total of 168 PC patients in TCGA-PAAD cohort were enrolled for this consensus analysis. Gene expression and clinical information of GSE21501 [17], GSE28735 [18], GSE57495 [19], GSE62452 [20], GSE71729 [11], GSE78229 [21], GSE79668 [22], and

GSE85916 were downloaded from Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/>) database. Gene expression and clinical information of PACA_AU_arraySeq, PACA_AU_RNA_seq and PACA_CA_seq were downloaded from International Cancer Genome Consortium (ICGC, <https://dcc.icgc.org/>). The information of E-MATB-6134 cohort was downloaded from ArrayExpress (<https://www.ebi.ac.uk/arrayexpress/>) [23]. Single-cell RNA sequencing (scRNA-seq) data of PC were acquired from multiple cohorts available, including CRA001160 [24], GSE111672 [25], GSE154778 [26], GSE155698 [27], GSE158356 [28], GSE212966 [29], GSE214295 [30], and GSE217845 [31].

Identification of the molecular subtypes based on multi-omics data in TCGA-PAAD cohort

“MOVICS” (Multi-Omics Integration and Clustering in Cancer Subtyping) is an innovative R package specifically designed for comprehensive multi-omics analysis in cancer research [32]. This powerful computational tool implements ten state-of-the-art clustering algorithms, enabling us to perform robust molecular subtyping of cancers based on integrated multi-omics data. We employed the “getElites” function to select the top 10% most variable features, including mRNA, miRNA, lncRNA, and methylation profiles, based on standard deviation ranking. The mutation data were systematically processed and transformed into a matrix format based on mutation counts. Finally, we integrated PC multi-omics data, which mRNA, miRNA, lncRNA, DNA methylation, and mutation. The “getClustNum” function was used to explore the optimal cluster number by calculating clustering prediction index (CPI) and Gap-statistics. The “getMOIC” function was utilized to systematically apply and compare ten different clustering methodologies, including SNE, PINSPplus, NEMO, COCA, LRAcluster, ConsensusClustering, IntNMF, CIMLR, MoCluster, and iClusterBayes. Utilizing the “getConsensusMOIC” algorithm, we constructed a consensus matrix to assess the robustness of clustering concordance among different methodologies, accompanied by the generation of a comprehensive consensus heatmap for pattern visualization. The “getSilhouette” function was employed to quantitatively assess sample similarity and evaluate the quality of molecular subtyping through silhouette coefficient analysis. The “getStdiz” and “getMoHeatmap” functions were used to get a standardized data and a comprehensive multi-omics heatmap, respectively. We conducted a comparative analysis of survival outcomes across different molecular subtypes.

Gene set enrichment analysis and gene set variation analysis

The “runDEA” function was used to run differential expression analysis with edgeR method. We subsequently used “runMarker” function to identify biomarkers and determine significantly differential expressed genes (DEGs) for each subtype. DEGs were filtered using an adjusted p-value cut-off of 0.05 and subsequently sorted in descending order according to their absolute log2 fold change values. Gene Set Enrichment Analysis (GSEA) was systematically performed to elucidate the Gene Ontology Biological Process (GOBP, <https://www.gsea-msigdb.org/gsea/msigdb/index.jsp>) alterations across different molecular subtypes of PC. Using the “runGSEA” function, we conducted GSEA to calculate activity of functional pathways for each identified subtype. Gene Set Variation Analysis (GSVA) was performed to characterize pathway activity profiles across distinct molecular subtypes of PC. The enrichment score of each sample was calculated based on the hallmark gene sets.

Differential patterns of immune infiltration and immune-related gene expression

To comprehensively characterize the tumor immune microenvironment, we employed six established computational algorithms to quantify tumor-infiltrating immune cell abundance: Tumor Immune Estimation Resource (TIMER; <http://timer.comp-genomics.org>) [33], Estimating Relative Subsets of RNA Transcripts (CIBERSORT; <https://cibersort.stanford.edu>) [34], xCell (<http://xCell.ucsf.edu>) [35], MCP-counter (<http://github.com/ebrecht/MCPcounter>) [36], quanTIseq (<http://icbi.at/quantiseq>) [37], and EPIC (<http://epic.gfellerlab.org>) [38]. These algorithms were systematically applied to all samples to ensure immune cell profiling, leveraging their distinct computational approaches for optimal immune microenvironment characterization. From a previous study, we obtained the gene sets of immunomodulators (chemokines, receptors, MHC, and immunostimulators) and immune checkpoints [39].

Validation of the molecular subtypes in external cohorts

We identified the biomarkers of different molecular subtypes in TCGA-PAAD cohort. Then, we conducted the nearest template prediction (NTP) based on these biomarkers to predict the subtypes in external cohorts, including GSE21501, GSE28735, GSE57495, GSE62452, GSE71729, GSE78229, GSE79668, GSE85916, E-MATB-6134, PACA_AU_arraySeq, PACA_AU_RNA_seq, and PACA_CA_seq cohorts. To evaluate the effectiveness of the classification, we compared the staging and prognostic differences among patients belonging to different classifications.

Identification of the prognostic-related genes

To identify the prognostic-related genes, we initially performed an intersection analysis of genes with available expression data across all 13 cohorts. We performed univariate Cox regression analysis and identified genes that consistently demonstrated a significance level of $P < 0.1$ with concordant trends across at least five cohorts.

Establishment of the prognostic signature based on machine learning methods

The dataset with the largest patient population, E-MATB-6134, was designated as the training cohort. The other 12 datasets were regarded as the external validation cohorts. Overall survival (OS), defined as the duration from initial diagnosis to death from any cause, served as the primary endpoint of our study. For surviving patients, the follow-up duration was censored at the last documented contact. We systematically integrated the expression profiles of genes with OS data using a comprehensive analytical framework incorporating ten distinct machine learning algorithms, resulting in 101 unique algorithm combinations. The selected algorithms included random survival forest (RSF), survival support vector machine (survival-SVM), ridge regression, elastic net (Enet), least absolute shrinkage and selection operator (LASSO), supervised principal components (SuperPC), generalized boosted regression modeling (GBM), partial least squares regression for Cox (plsRcox), CoxBoost, and stepwise Cox regression. Among these, five algorithms incorporated intrinsic variable selection capabilities, including RSF, LASSO, CoxBoost, backward stepwise Cox, and both versions of stepwise Cox regression. Model development and evaluation were conducted using a leave-one-out cross-validation (LOOCV) framework, with predictive performance assessed through Harrell's concordance index (C-index) calculated independently for these datasets. The optimal model was selected based on the highest average C-index across these cohorts. We then computed risk scores for each patient and established prognostic stratification using the optimal score as the threshold, dividing patients into distinct low-risk and high-risk subgroups.

Evaluation of the predictive performance of the signature

C-index was employed to quantitatively assess the predictive performance of prognostic factors. We firstly evaluated the prognostic performance of the signature against established clinical factors known to impact PC outcomes. For comprehensive benchmarking of this signature's predictive efficacy, we conducted a systematic literature review of established pancreatic cancer prognostic models. The published algorithms of prior prognostic models were retrieved, comprising gene names and regression coefficients. Application of these formulas

yielded individual risk scores across training and validation datasets. Subsequently, we calculated the C-index for each signature across all these cohorts. Ultimately, the predictive performance of each signature was quantitatively assessed by comparing C-index values across all cohorts.

Drug sensitivity prediction

Drug sensitivity of the samples was predicted using the "pRRophetic" package. We compared the differences in the 50% inhibitory concentration (IC50) values of the candidate drugs across different molecular subtypes. A lower IC50 value indicated higher drug sensitivity.

Construction of a prognostic model

The risk score was integrated with clinically relevant prognostic variables to develop a prediction model of PC. The clinical variables comprised age, gender, tumor grade, TNM stage, anatomical site, and adjuvant therapy history. Univariate and multivariate Cox regression analyses were performed on these factors to establish the prognostic model. The prognostic model was dynamically visualized via the "DynNom" package, generating an interactive nomogram and online calculator. Performance of the prognostic model was rigorously evaluated through calibration curves, decision curve analysis (DCA), and receiver operating characteristic (ROC) curves.

Single-cell and spatial transcriptional landscape of PC

Following established computational pipelines, we performed comprehensive single-cell RNA sequencing (scRNA-seq) analyses on eight independent cohorts, including CRA001160, GSE111672, GSE154778, GSE155698, GSE158356, GSE212966, GSE214295, and GSE217845. Specifically, we implemented stringent quality control measures by excluding genes detected in fewer than three cells and cells expressing fewer than 200 genes. Normalization was performed by scaling the total transcript counts per cell to 10,000, followed by log2 transformation. For each dataset, we identified the top 2000 highly variable genes that exhibited the most significant expression variations across cells. After expression level scaling, we conducted principal component analysis (PCA) for linear dimensionality reduction. The optimal number of principal components (PCs) was determined through a dual validation approach using JackStrawPlot and ElbowPlot analyses. Subsequent cell clustering was performed using the first 12 PCs with a graph-based clustering algorithm at a resolution parameter of 0.8. The resulting cellular subpopulations were visualized through uniform manifold approximation and projection (UMAP) dimensional reduction.

Cell cluster annotation was conducted using a dual-validation strategy, combining canonical marker gene identification with cross-validation against original publication annotations. Comparative analysis of the related gene expression patterns was conducted across distinct cell clusters to identify cell-type-specific expression profiles.

The spatial transcriptomics dataset GSE272362 was analyzed to characterize expression levels and tissue localization of target genes in PC specimens [40]. This dataset comprised diverse PC specimens, including primary tumors, normal pancreatic tissues, and matched hepatic/lymph node metastases. Cell annotation data for the samples were obtained from the Zenodo repository (<https://zenodo.org/records/10712047>).

Cell lines and culture

We acquired a panel of PC lines from the American Type Culture Collection (ATCC, Manassas, VA, USA), including one human normal pancreatic duct epithelial cell line (H6C7) and four PC cell lines (T3M-4, PANC-1, BxPC-3, and Capan-2). All cell lines were maintained in Dulbecco's Modified Eagle Medium (DMEM) (Invitrogen, Carlsbad, CA, USA) supplemented with 10% heat-inactivated fetal bovine serum (FBS, Invitrogen) and 1% penicillin-streptomycin antibiotic solution (100 U/mL penicillin and 100 µg/mL streptomycin), and cultured under standard conditions at 37 °C in a humidified atmosphere containing 5% CO₂.

Real-time quantitative polymerase chain reaction (RT-qPCR)

The utilization of human cancer tissues in this study was approved by the Ethics Committee of Henan Provincial People's Hospital, Zhengzhou University People's Hospital (Approval No. 58, 2019). We retrospectively collected paired tumor and adjacent normal tissues from patients who met the following inclusion criteria: (1) provision of written informed consent prior to surgery; (2) no history of neoadjuvant therapy; (3) underwent radical resection followed by adjuvant chemotherapy; (4) age ≥ 18 years; and (5) normal cardiac, hepatic, and renal function.

Real-time quantitative polymerase chain reaction (RT-qPCR) was performed to validate the gene expression profiles in both cultured cell lines and clinical tissue specimens. Total RNA was isolated from cells and tissues using TRIzol reagent (Takara, Japan). Subsequently, 1 µg of total RNA was reverse transcribed into complementary DNA (cDNA) using the PrimeScript RT Reagent Kit with gDNA Eraser (Takara, Japan). Quantitative PCR amplification was conducted using TB Green Premix Ex Taq II (Takara, Japan) on a real-time PCR system. The housekeeping gene β-actin served as an internal control for data normalization. Relative mRNA expression levels

were quantified using the 2^{-ΔΔCt} method. Detailed primer sequences used in this study are provided in Supplementary Table 1.

Immunohistochemistry assay

Immunohistochemistry was performed to evaluate the expression patterns of genes in PC specimens. Formalin-fixed, paraffin-embedded (FFPE) tissue Sect. (4% formaldehyde, 5 µm thickness) were processed through the following protocol. Sections were initially incubated at 60 °C for 1 h, followed by deparaffinization in xylene and rehydration through a graded ethanol series. Antigen retrieval was performed using citrate buffer in a microwave heating system. To quench endogenous peroxidase activity, sections were treated with 3% hydrogen peroxide for 15 min at room temperature, followed by three phosphate-buffered saline (PBS) washes. Primary antibody incubation was conducted at 4 °C overnight in a humidified chamber. After thorough washing with PBS, sections were incubated with horseradish peroxidase (HRP)-conjugated secondary antibody (ZSGB-Bio, PV-6000) for 30 min at room temperature. Chromogenic detection was achieved using 3,3'-diaminobenzidine (DAB) as substrate, and sections were counterstained with hematoxylin. Protein localization and expression intensity were assessed and imaged using light microscopy (Nikon Eclipse 80i, Japan) by two independent pathologists blinded to the clinical information. We listed the antibodies used in this study in Supplementary Table 2.

Immunofluorescence assay

Cells were seeded onto six-well plates equipped with 10 mm coverslips and incubated for 24 h to allow for adherence. After a thorough washing with PBS to remove any media or debris, the cells were fixed with 4% paraformaldehyde to maintain their structural integrity. The cells were then incubated with the primary antibody overnight at 4 °C to facilitate specific binding to the target antigen. Following removal of the primary antibody, the cells were incubated with a fluorochrome-conjugated secondary antibody to visualize the primary antibody binding, along with 4',6-diamidino-2-phenylindole (DAPI) to stain the cell nuclei. Finally, the cells were imaged using a fluorescence microscope to capture the immunofluorescence staining pattern.

Western blotting

Protein extraction was performed by lysing cells, tumor tissues, and paired normal tissues in RIPA buffer (Beyotime Biotechnology, China) supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF) on ice for 30 min. The lysates were centrifuged for 15 min at 4 °C. Then, the supernatant was collected for protein quantification using a BCA protein assay kit (Applygen Technologies,

China) according to the manufacturer's protocol. Equal amounts of protein samples (50 µg) were separated by 8% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and subsequently transferred onto polyvinylidene fluoride (PVDF) membranes (Millipore, USA) using a semi-dry transfer system. Membranes were blocked with 5% non-fat dry milk in Tris-buffered saline containing 0.1% Tween-20 (TBST) for 2 h at room temperature, followed by incubation with primary antibodies at 4 °C overnight. After three washes with TBST, membranes were incubated with HRP-conjugated secondary antibodies for 2 h at room temperature. Protein bands were visualized using enhanced chemiluminescence (ECL) substrate and imaged with a ChemiDoc™ CRS+ Molecular Imager (Bio-Rad, Hercules, CA, USA). β-actin served as the internal control for protein loading normalization.

Lentiviral-mediated transfection

For gene manipulation studies, lentiviral particles containing either short hairpin RNA (shRNA) sequences or gene overexpression constructs were commercially synthesized by RiboBio (Guangzhou, China) and Genomeditech Co. (Shanghai, China). Transfection was performed using DharmaFECT4 transfection reagent (Thermo Scientific, USA) according to the manufacturer's optimized protocol. Following an 8-hour incubation period, the transfection medium was replaced with fresh complete growth medium. Transfection efficiency was quantitatively assessed 48 h post-transfection through parallel analyses using both RT-qPCR and western blotting. Sequences of the shRNAs used in this study are listed in Supplementary Table 3.

Cell counting kit-8 assay

To evaluate the proliferative capacity of human PC cell lines *in vitro*, we performed quantitative cell proliferation assays using the Cell Counting Kit-8 (CCK-8) assay. Cells were harvested during their exponential growth phase and plated in 96-well culture plates at a density of 2,000 cells per well in 100 µL of complete growth medium. The plates were maintained under standard culture conditions (37 °C and 5% CO₂) throughout the experiment period. Cell proliferation was assessed at four time points (24, 48, 72, and 96 h), with the initial measurement (0 h) serving as the baseline control. At each designated time point, 10 µL of CCK-8 reagent (APExBIO, USA) was added to each well, followed by incubation for 2 h under standard culture conditions. The metabolic activity of viable cells, which correlates with cell proliferation, was quantified by measuring the optical density (OD) at 450 nm using a microplate reader. All measurements were performed in triplicate, and background absorbance

from blank wells containing medium alone was subtracted from the experimental values.

Colony formation assay

The proliferative potential of various cell lines was assessed using a standardized colony formation assay. Single-cell suspensions were prepared through enzymatic dissociation and precise pipetting to ensure uniform cell distribution. Cells were seeded in 6-well culture plates at a density of 500 cells per well in 2 mL of complete growth medium. The plates were incubated under optimal conditions (37 °C and 5% CO₂) for 14 days, allowing for undisturbed colony formation. Following the incubation period, colonies consisting of ≥ 50 cells were identified by visual inspection. The culture medium was carefully aspirated, and colonies were fixed with 4% paraformaldehyde (Maokang Biotechnology Co., Ltd, Shanghai, China) for 30 min at room temperature. After fixation, cells were stained with 0.1% crystal violet solution for 30 min to facilitate colony visualization. Excess stain was removed by gentle washing with PBS.

Wound healing assay

Cells were plated in 6-well culture plates at a density of 5×10^5 cells/well and allowed to adhere overnight in complete growth medium. A uniform wound was created by scraping the monolayer with a sterile P200 pipette tip, followed by three washes with PBS to remove cellular debris. Cell migration was monitored at 12-hour intervals, and images were captured using a light microscope.

Cell migration and invasion assays

Cell migratory and invasive capacities were evaluated using Transwell chambers (Corning 3422, USA) following the manufacturer's protocol with modifications. Briefly, cells were harvested during the logarithmic growth phase, centrifuged, and resuspended in serum-free DMEM medium. For migration assays, 5×10^4 cells in 200 µL serum-free medium were seeded into the upper chamber of uncoated Transwell inserts. For invasion assays, an equivalent number of cells were plated in Matrigel-coated upper chambers. The lower chambers were filled with 600 µL of complete medium containing 10% FBS as a chemoattractant. Following incubation, non-migrated/non-invaded cells on the upper membrane surface were removed using cotton swabs. Migrated or invaded cells on the lower surface were fixed with 4% paraformaldehyde for 20 min at room temperature and stained with 0.1% crystal violet solution for 20 min.

Flow cytometric analysis of apoptosis and cell cycle distribution

Cells in the logarithmic growth phase were harvested, trypsinized, and washed twice with ice-cold PBS. The

cell pellet was resuspended at a concentration of 1×10^6 cells/mL. For apoptosis detection, 100 μ L of cell suspension was incubated with 5 μ L annexin V-FITC (KeyGen Biotech, China) and 10 μ L propidium iodide (PI, KeyGen Biotech, China) in the dark at room temperature for 15 min. Following incubation, 400 μ L of binding buffer was added, and samples were immediately analyzed using a FACSCalibur flow cytometer (BD Biosciences, USA).

For cell cycle distribution analysis, harvested cells were fixed in 70% ice-cold ethanol at 4 °C overnight. Fixed cells were washed twice with PBS and treated with RNase A to remove RNA interference. Cellular DNA was stained with 50 μ g/mL PI in the dark at room temperature for 30 min. Cell cycle distribution was analyzed using the FACSCalibur flow cytometer.

Co-immunoprecipitation assay

For protein-protein interaction studies, transfected cells were harvested and lysed in ice-cold immunoprecipitation (IP) buffer supplemented with Protease Inhibitor Cocktail (Roche, Switzerland). Cell lysates were centrifuged at 4 °C, and the supernatant was collected for IP. Equal amounts of protein were incubated overnight at 4 °C with rotation using specific primary antibodies. Normal IgG from the same host species served as a negative control. Protein A-sepharose beads (GE Healthcare, USA) were added to the antigen-antibody complex and incubated at 4 °C with gentle rotation. The immune complexes were washed three times with cold IP buffer and eluted in Laemmli buffer at 95 °C. The eluted proteins were resolved by SDS-PAGE and analyzed by western blotting using appropriate secondary antibodies.

Ras activation assay

The activation status of Ras was determined using the Active Ras Pull-Down and Detection Kit (Thermo Fisher Scientific) according to the manufacturer's protocol. Briefly, cells were lysed in the provided lysis buffer supplemented with protease inhibitor cocktail (Roche, Switzerland). The active GTP-bound form of Ras was isolated by incubating cell lysates with Raf1-RBD agarose beads at 4 °C with gentle agitation. The beads were then washed three times with lysis buffer, and bound proteins were eluted in Laemmli buffer. For quantitative analysis, the pull-down samples were subjected to western blotting using the mouse anti-Ras (1:1000, Thermo Fisher Scientific) and HRP-conjugated secondary antibodies. Ras activation levels were normalized to total Ras expression.

Ubiquitination assay

To investigate the protein ubiquitination level, cells were treated with 10 μ M MG132 (C2211, Sigma-Aldrich, USA), a cell-permeable proteasome inhibitor, for 6 h to prevent protein degradation. Cells were then lysed in IP

buffer supplemented with Protease Inhibitor Cocktail (Roche, Switzerland) and Phosphatase inhibitor (Sigma-Aldrich, USA) for 30 min on ice. Cell lysates were centrifuged at 4 °C, and the supernatant was collected for IP. Equal amounts of protein were incubated overnight at 4 °C with rotation using either the primary antibody or control IgG. Protein A/G magnetic beads (Thermo Fisher Scientific, USA) were added and incubated for an additional 2 h at 4 °C. The immunoprecipitates were washed five times with cold IP buffer and eluted in Laemmli buffer. Ubiquitinated protein was detected by western blot analysis using anti-ubiquitin antibody. The membranes were probed with HRP-conjugated secondary antibodies and visualized using ECL substrate.

Three-dimensional organoid culture and functional characterization

Human PC tissue specimens were mechanically dissociated and enzymatically digested in DMEM containing 1 mg/mL collagenase IV (Sigma-Aldrich, Germany) at 37 °C for 15 min. The digestion was quenched with ice-cold PBS supplemented with 10% FBS. After centrifugation and filtration, the tissue fragments were resuspended in Cultrex Reduced Growth Factor Basement Membrane Extract (Trevigen, USA) and plated in 24-well plates. The embedded tissue was cultured in organoid growth medium (Biozellen, USA) containing advanced DMEM/F12 supplemented with N21-MAX, N2-MAX, glutamine, nicotinamide, HEPES, A83-01, N-acetylcysteine, FGF-10, BMP7, EGF, HGF, Noggin, Wnt-3 α , R-Spondin 1, Gastrin I, Forskolin, and Y-27,632. The cultures were maintained at 37 °C in a humidified 5% CO₂ incubator, with medium changes every 2–3 days.

For organoid propagation, the cultures were harvested using Cultrex Organoid Harvesting Solution (3700-100-01, R&D Systems, USA). The organoids were mechanically dissociated and enzymatically digested into single cells, followed by centrifugation and PBS washing. The cells were then re-embedded in fresh Cultrex matrix and cultured in proliferation medium containing advanced DMEM/F12 supplemented with N21-MAX, N2-MAX, glutamine, nicotinamide, HEPES, A83-01, N-acetylcysteine, FGF-10, BMP7, EGF, HGF, R-Spondin 1, Gastrin I, and Forskolin.

For assessment of epithelial-mesenchymal transition (EMT), organoids were fixed in 4% paraformaldehyde at 4 °C for 60 min, permeabilized with 0.5% Triton X-100, and blocked with 5% bovine serum albumin (BSA) for 60 min at room temperature. The samples were incubated with the primary anti-E-Cadherin and anti-Vimentin antibodies overnight at 4 °C, followed by the secondary antibody for 2 h at room temperature. Nuclei were counterstained with DAPI for 15 min. Fluorescence images were acquired using a confocal microscope (Leica TCS SP8, Germany).

Statistical analysis

Continuous variables were expressed as mean $\pm$ standard deviation (SD), while categorical variables were presented as absolute numbers with corresponding percentages. Univariate Cox regression analysis was initially performed to identify potential prognostic factors using backward stepwise selection to determine independent predictors of survival. The strength of association between covariates and survival outcomes was quantified using hazard ratios (HRs) with corresponding 95% confidence intervals (CIs). Statistical significance was defined as a two-tailed P -value < 0.05 , with the following notation used to indicate varying levels of significance: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, and $****P < 0.0001$. All statistical analyses were conducted using R statistical software (version 4.1.0), and Python (version 3.8).

Results

Identification of the two molecular subtypes in TCGA-PAAD cohort

According to the mentioned standards, we totally enrolled 187 mRNAs, 57 miRNAs, 190 lncRNAs, 144 DNA methylation sites, and 35 somatic mutations. Through comprehensive evaluation utilizing both CAP and Gap-statistics methodology, we identified two distinct clusters as the optimal classification number (Fig. 1A). The heatmap illustrated the expression profiles across multi-omics, highlighting the top 10 most variable features within each omic category (Fig. 1B). The Silhouette plot demonstrated a good clustering effect for the two molecular subtypes (Fig. 1C). We integrated the typing results from 10 different classification methods and ultimately determined the molecular subtype for each sample (Fig. 1D). The consensus heatmap clearly showcased the distinctive characteristics of the two molecular subtypes (Fig. 1E). The Kaplan-Meier curves demonstrated clear separation between the two subtypes, with CS1 showing consistently poorer prognosis across all time points ($P < 0.001$, Fig. 1F).

Distinct characteristics were explored between the two molecular subtype

To evaluate the efficacy of molecular subtyping, we systematically compared the characteristics of samples from both molecular subtypes, with particular emphasis on immune-related features. Multiple computational algorithms were employed to quantitatively assess immune cell infiltration within the samples (Fig. 2A). The infiltration of various immune cells exhibited significant differences between CS1 and CS2 subtypes, including neutrophil, macrophage, myeloid dendritic cell, and T cell regulatory. The two molecular subtypes also exhibited significant differences in the enriched biological processes (Fig. 2B). Marker genes of the CS1 subtype were significantly enriched in multiple immune-related biological processes, including antimicrobial humoral immune response mediated antimicrobial

peptide, neutrophil migration, and interferon gamma production. Furthermore, these genes were also enriched in various biological processes related to EMT, including cornification, regulation of epidermis development, and regulation of epidermal cell differentiation. The CS2 subtype was enriched in metabolic-related biological processes, including mRNA export from nucleus, negative regulation of voltage gated potassium channel activity, and glutamate receptor signaling pathway. We observed the significant differential enrichment in immune process related gene sets between the two molecular subtypes (Fig. 2C). There are notable differences in the enrichment levels of multiple pathways related to PC progression within the hallmark gene sets between CS1 and CS2 subtypes, including apoptosis, EMT, hypoxia, and KRAS signaling (Fig. 2D). The expression of most immunomodulators was significantly lower in the CS1 subtype compared to the CS2 subtype (Fig. 2E). Furthermore, the expression of multiple immune checkpoints, including CTLA4 and CD276, is notably lower in the CS1 subtype compared to the CS2 subtype (Fig. 2F).

Validation of the molecular subtyping efficacy across multiple PC cohorts

We utilized 12 external PC cohorts to validate the efficacy of molecular subtypes. DEGs between the CS1 and CS2 subtypes were analyzed. The NTP method was employed to predict the molecular subtype assignment for each sample in the external cohort. Next, the Kaplan-Meier method was utilized to compare the prognostic differences between the two molecular subtypes. The prognosis of the CS2 subtype was consistently superior to that of the CS1 subtype across multiple cohorts, with the exception of the GSE28735 cohort ($P < 0.05$, Fig. 3A-F, Supplementary Fig. 1A-F). Comparative analysis revealed distinct clinical profiles between the two subtypes. Compared to CS2, the CS1 subtype exhibited more advanced tumor stages and higher histological grades ($P < 0.05$, Fig. 3G-J). We systematically compared our classification system with previously reported PC typing approaches, including immune-based models and whole-tumor classification schemes. The CS1 subtype was predominantly characterized by Immune C2 and C6 signatures, whereas CS2 showed enrichment of the Immune C3 signature ($P < 0.05$, Fig. 3K). CS1 was enriched for Stroma-Activated and Pure Basal-like, while CS2 favored Pure Classical and Immune Classical subtypes ($P < 0.05$, Fig. 3L).

Development and evaluation of the prognostic signature

We developed and validated a robust prognostic signature utilizing multiple machine learning algorithms. To ensure cross-platform consistency, we identified 195 consensus genes through intersection analysis of gene expression profiles across multiple cohorts. Univariate Cox regression analysis was systematically performed for these genes in each individual cohort (Supplementary

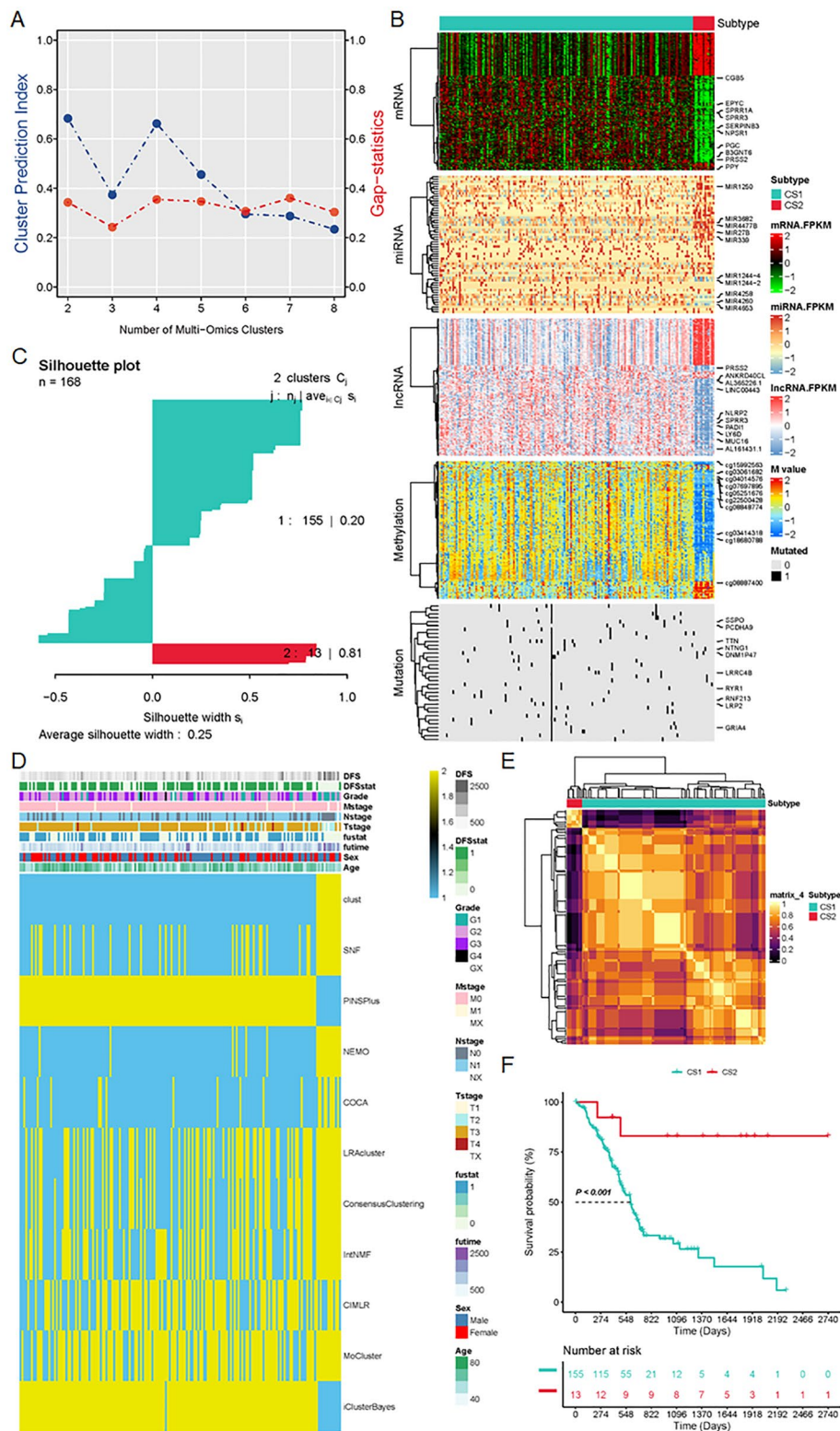
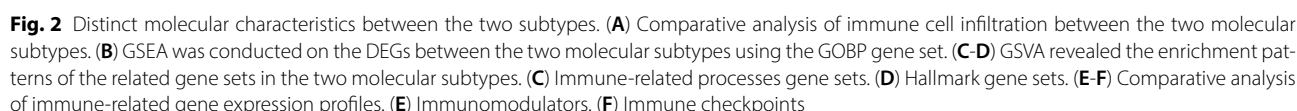


Fig. 1 PC samples within the TCGA-PAAD cohort can be categorized into CS1 and CS2 subtypes. **(A)** The optimal number of clusters was determined through a comprehensive evaluation using both the CAP and Gap-statistics methods. **(B)** Multi-omics features of the two molecular subtypes, including mRNA, miRNA, lncRNA, DNA CpG methylation site, and mutant gene. **(C)** Silhouette plot of the two molecular subtypes. **(D)** A comprehensive heatmap was generated to visualize the distribution of clinical characteristics, integrated with molecular subtyping results. The molecular classification was established through consensus clustering analysis, yielding ten distinct multi-omics clustering methods. **(E)** Consensus clustering matrix for three novel prognostic subtypes based on the 10 algorithms. **(F)** Kaplan-Meier plot of OS for the two subtypes



validation sets. We developed prognostic models using 101 machine learning algorithm combinations in the training set, followed by validation across 12 independent cohorts. The optimal algorithm combination was determined based on the highest average C-index across all validation sets (Fig. 4A). Among all algorithm

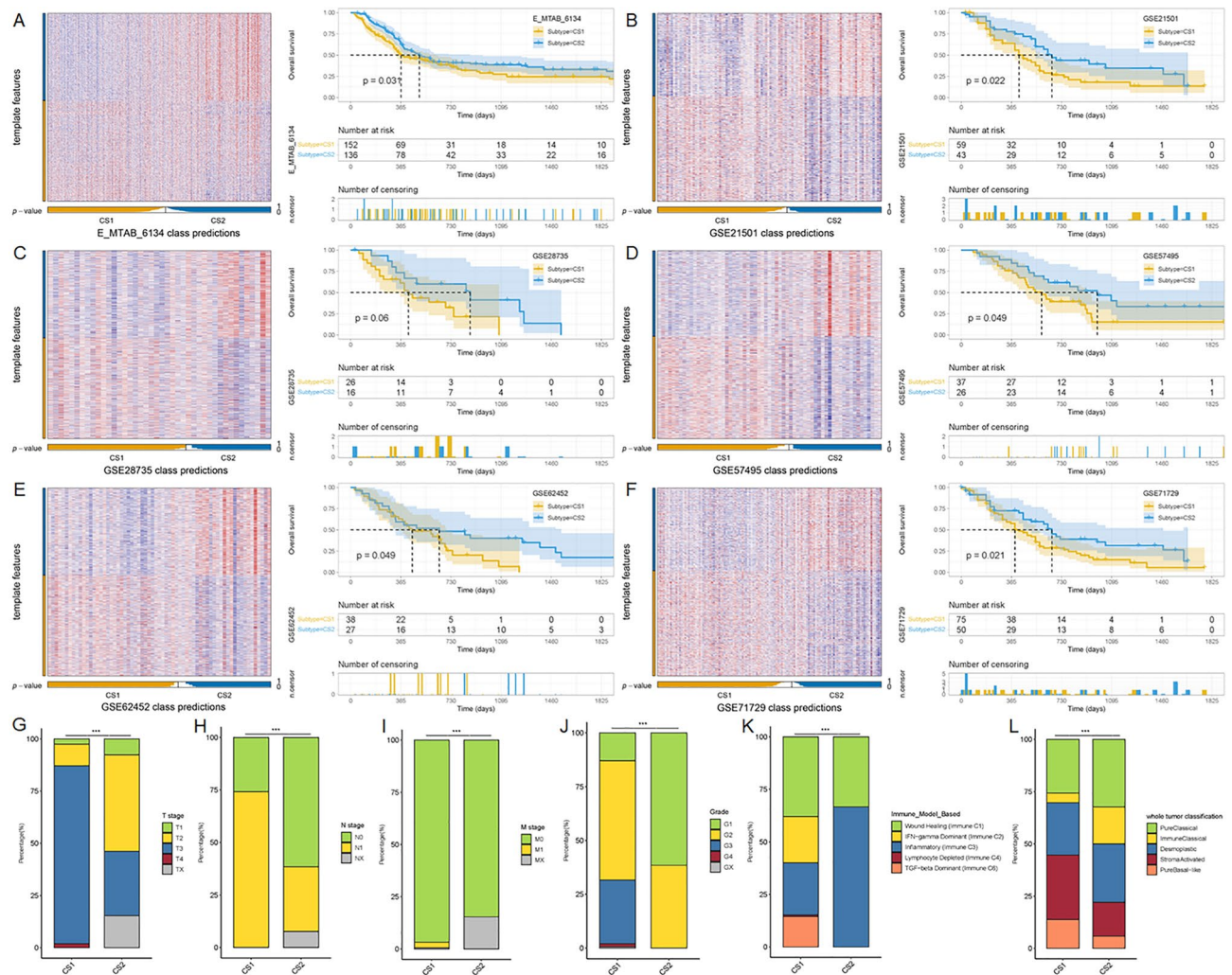


Fig. 3 The robustness and generalizability of the molecular subtyping were validated across 6 independent external cohorts, demonstrating consistent classification performance and clinical relevance. (A) E_MTAB_6134 cohort. (B) GSE21501 cohort. (C) GSE28735 cohort. (D) GSE57495 cohort. (E) GSE62452 cohort. (F) GSE71729 cohort. (G-K) CS1 and CS2 subtypes exhibited distinct TNM staging, histological grading, and immune classification patterns in TCGA-PAAD cohort. (G) T stage. (H) N stage. (I) M stage. (J) Tumor grade. (K) Immune based model. (L) CS1 and CS2 demonstrated divergent whole-tumor classifications

combinations, ridge regression demonstrated the highest mean C-index across the 12 validation cohorts. The optimal λ value was determined at the minimal partial likelihood deviance, and corresponding variable coefficients were subsequently calculated (Fig. 4B). We identified prognostic genes through univariate Cox regression analysis followed by rigorous screening criteria (Fig. 4C and Supplementary Fig. 2A-L). In the ridge regression model, the top five variables with highest coefficients were EREG, FAM83A, C12orf56, ZBED2, and A2ML1 (Supplementary Table 5). Sample-specific risk scores were computed as the sum of each gene's expression multiplied by its ridge coefficient (Risk score = $\sum[\text{gene} \times \text{ridge coefficient}]$). Based on ridge regression, we calculated risk scores for each sample and subsequently stratified them into high- and low-risk groups using the optimal

risk score as the cut-off value (Fig. 4D and Supplementary Fig. 3A-L). With the exception of the GSE21501 cohort, the high-risk group consistently demonstrated significantly worse prognosis compared to the low-risk group across all datasets ($P < 0.05$; Fig. 4E and Supplementary Fig. 4A-L).

Through a systematic literature review, we compiled and analyzed 172 previously reported prognostic signatures for PC (Supplementary Table 6). Inclusion was restricted to formula-based signatures to facilitate direct comparison. Due to the insufficient miRNA annotation coverage across cohorts, we restricted our summary to prognostic signatures derived from mRNA and lncRNA data. These prognostic signatures were constructed based on hallmark features of PC progression, encompassing immune cells, basement membrane, EMT, autophagy,

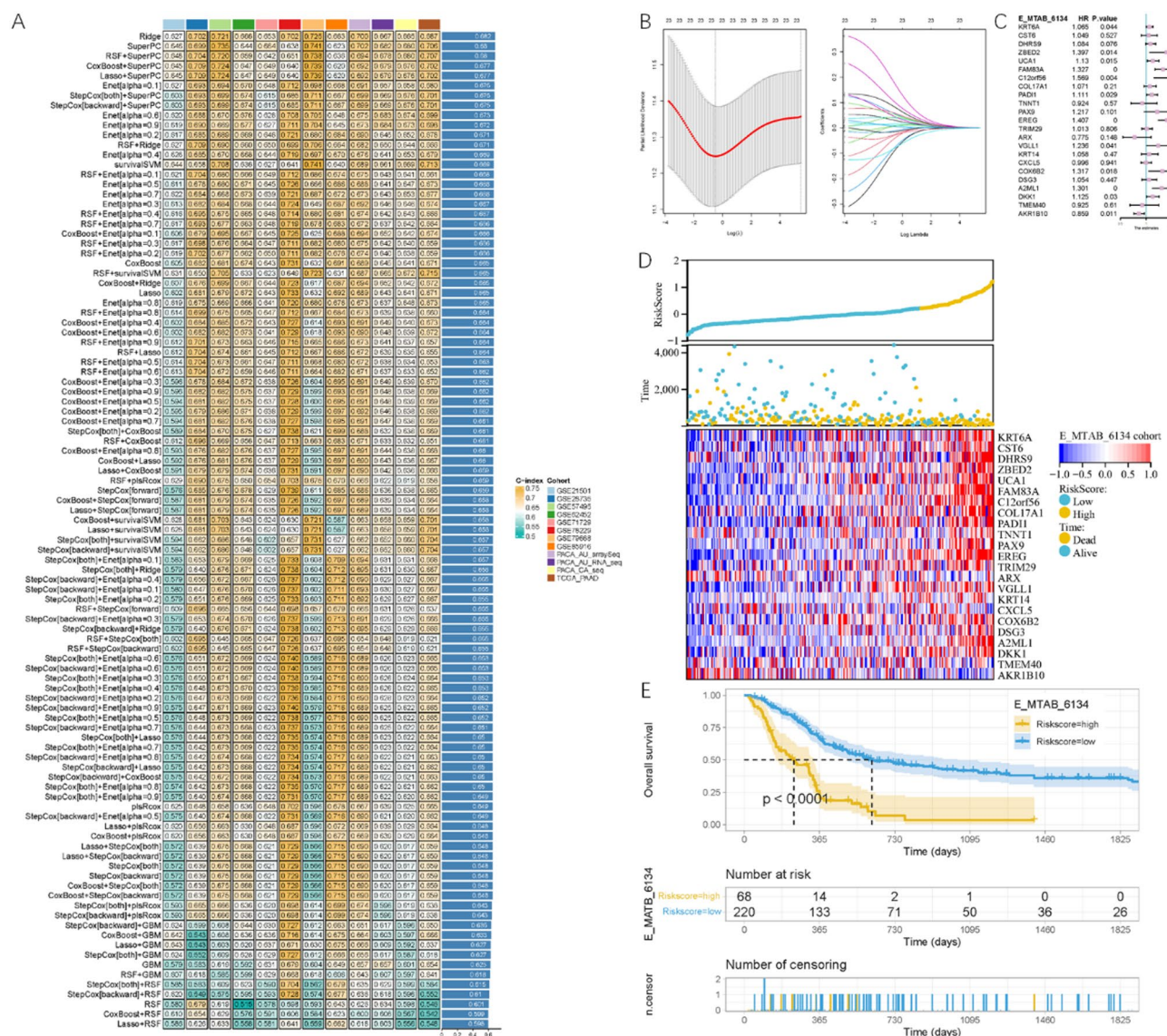


Fig. 4 The prognostic signature was developed using 101 machine learning algorithm combinations. **(A)** The C-index was calculated for each algorithm combination in the validation sets, and the mean C-index values were subsequently determined. **(B)** Left: The optimal λ value was determined at the point of minimal partial likelihood deviance. Right: The coefficients of all variables were calculated across a range of λ values. **(C)** Univariate Cox regression analysis was performed for the 23 candidate genes in the E-MTAB-6134 cohort. **(D)** The E-MTAB-6134 cohort was characterized by risk scores, stratification groups, survival outcomes, and gene expression profiles. **(E)** Survival disparities between high-risk and low-risk groups in the E-MTAB-6134 cohort

ferroptosis, glycolytic pathways et al. Due to inter-cohort probe heterogeneity, the included prognostic signatures varied slightly among cohorts. Consistently top-tier C-index rankings in all cohorts validated the signature's excellent predictive stability (Fig. 5A-G and Supplementary Fig. 5A-F).

This signature revealing the features of PC

Multiple clinical variables significantly impacted PC prognosis, including patient age, tumor stage, histological differentiation grade, and treatment modalities. The signature surpassed clinical variables in prognostic accuracy across all evaluated cohorts (Fig. 6A-E). Using the

well-characterized TCGA-PAAD cohort with comprehensive clinical data, we developed a prognostic model for pancreatic cancer. Based on univariate and multivariate Cox regression analyses, we incorporated age, tumor grade, chemotherapy status, and risk score into the final prognostic model (Fig. 6F, G). For practical application, this model was implemented as both a dynamic nomogram and an online calculator (Fig. 6H, <https://jiacheng.e.shinyapps.io/tcga-paad-prognosis/>). The calibration curve and DCA collectively demonstrated the model's robust predictive performance (Fig. 6I, J). The prognostic model performance showed high accuracy for OS prediction (1-year: AUC = 0.82; 2-3-year: AUC = 0.79; Fig. 6K).



Fig. 5 Comparative C-index analysis in multiple PC cohorts demonstrated superior prognostic performance of this signature versus prior established signatures. **(A)** C-index of 150 signatures in E-MTAB-6134 cohort. **(B)** C-index of 146 signatures in GSE21501 cohort. **(C)** C-index of 151 signatures in GSE28735 cohort. **(D)** C-index of 153 signatures in GSE57495 cohort. **(E)** C-index of 151 signatures in GSE62452 cohort. **(F)** C-index of 146 signatures in GSE71729 cohort. **(G)** C-index of 151 signatures in GSE78229 cohort

GSEA of DEGs between risk groups revealed significant enrichment of metabolic pathways in the low-risk cohort, including GABAergic synapse, protein export, and synaptic vesicle (Fig. 7A). In contrast, the high-risk group exhibited significant enrichment in immune-related disease pathways, including asthma, autoimmune thyroid disease, and graft-versus-host disease (Fig. 7A). Significant differences in drug sensitivity profiles were observed between the two groups (Fig. 7B). The high-risk group demonstrated increased sensitivity to AICAR, Bleomycin, and Epothilone.B ($P < 0.001$). Whereas, the low-risk group showed enhanced response to Axitinib, Elesclomol, and Temsirolimus, with all comparisons reaching statistical significance ($P < 0.001$). We observed

no significant correlation between risk scores and tumor mutation burden (TMB) ($P > 0.05$, Fig. 7C). Risk scores showed significant associations with tumor grade, resection margin status, and histological classification ($P < 0.05$, Fig. 7D-F). A progressive increase in risk scores was observed across tumor grades, with G2 samples showing significantly higher scores than G1, and G3/4 samples exhibiting further elevation compared to G2 ($P < 0.05$, Fig. 7D). It indicated that risk scores were significantly elevated in samples with R1 resection margins compared to R0 margins ($P < 0.05$, Fig. 7E). The basal-like subtype exhibited significantly higher risk scores compared to the classical subtype ($P < 0.001$, Fig. 7F).

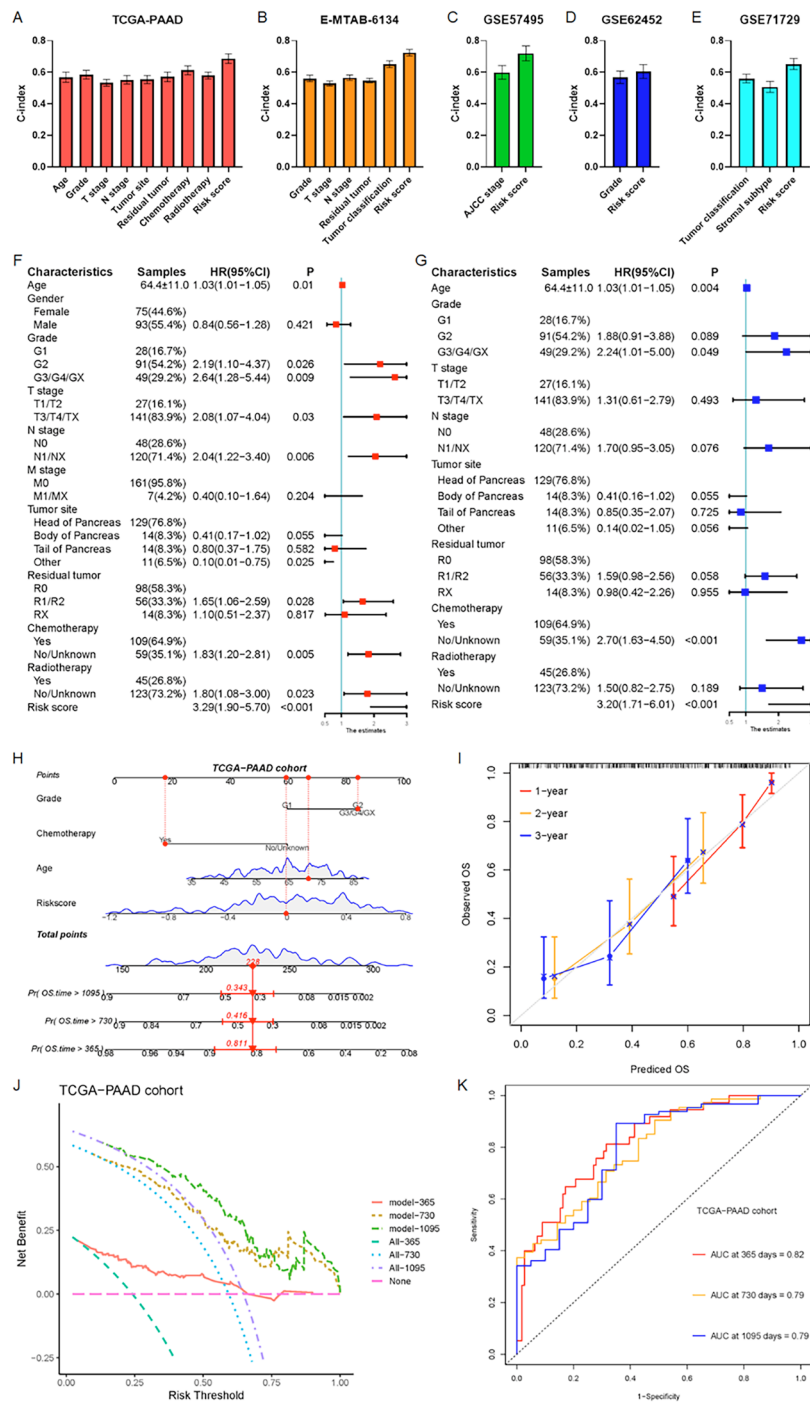


Fig. 6 The integrated prognostic model incorporated both the signature and clinical characteristics for PC. **(A–E)** Comparative prognostic performance analysis between this signature and clinical features was conducted in several independent PC cohorts. **(A)** TCGA-PAAD cohort. **(B)** E-MTAB-6134 cohort. **(C)** GSE57495 cohort. **(D)** GSE62452 cohort. **(E)** GSE71729 cohort. **(F–G)** Univariate and multivariate Cox regression analyses were performed to assess the prognostic value of the risk score and other potential clinical variables in PC. **(F)** Univariate Cox regression. **(G)** Multivariate Cox regression. **(H)** Dynamic nomogram visualization of the prognostic model. **(I)** Calibration curve analysis of the prognostic model. **(J)** DCA. **(K)** ROC curves

A2ML1 was overexpressed in PC and strongly associated with poor clinical outcomes

We systematically compared the expression profiles of these 23 genes between normal and tumor tissues.

In TCGA-PAAD, GSE28735, and GSE62452 cohorts, FAM83A (Family With Sequence Similarity 83 Member A) and A2ML1 (Alpha-2-Macroglobulin Like-1) exhibited significant differential expression between normal

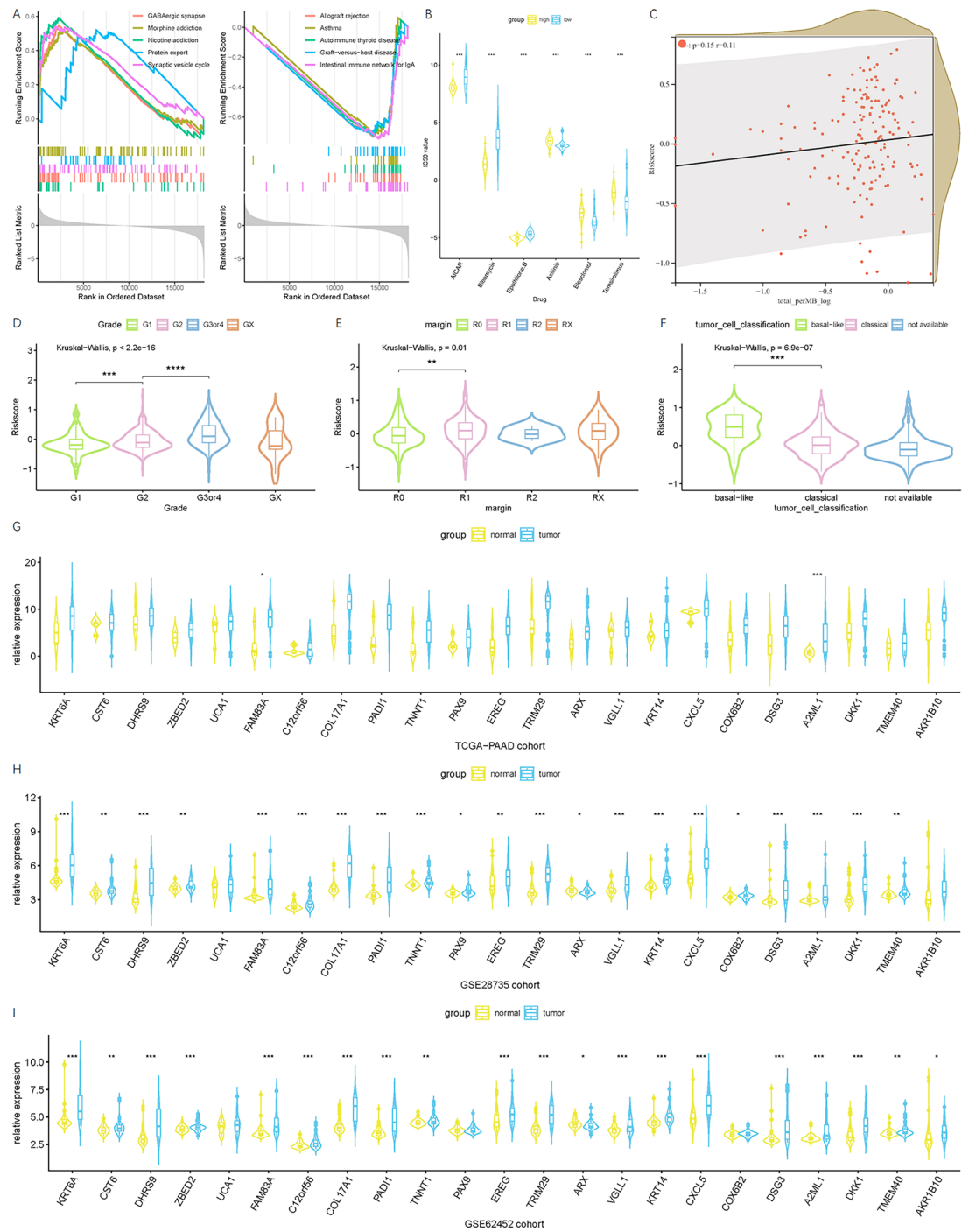


Fig. 7 Distinct clinical characteristics were observed between the low- and high-risk score groups. **(A)** GSEA was performed to identify differentially enriched pathways between low- and high-risk groups. Left: Low-risk group. Right: High-risk group. **(B)** Drug sensitivity analysis revealed distinct therapeutic response patterns between low- and high-risk groups. **(C)** Correlation between risk scores and mutation burdens. **(D–F)** The relationships between risk scores and patients' clinical information. **(D)** Tumor grade. **(E)** Tumor margin. **(F)** Tumor cell classification. **(G–I)** Differential expression of 23 prognosis-related genes between normal and tumor samples in **(G)** TCGA-PAAD cohort, **(H)** GSE28735 cohort, and **(I)** GSE62452 cohort

and tumor tissues, with markedly higher expression levels in tumor samples ($P < 0.05$, Fig. 7G–I). Subsequently, we focused on investigating the role of A2ML1 in PC, given the extensive existing literature on FAM83A [41–45].

Due to the complexity of PC microenvironment, we validated the single-cell expression patterns of FAM83A and A2ML1 across multiple scRNA-seq cohorts. Based on established marker genes and previous research, we

performed sequential dimensionality reduction, clustering, and cell type annotation of single-cell data. Our analysis revealed predominant overexpression of FAM83A and A2ML1 in epithelial cells, ductal cells, and malignant cell populations (Fig. 8A–D and Supplementary Fig. 6A–D). Spatial transcriptomic analysis revealed distinct A2ML1 expression profiles and tissue localization in PC specimens (Fig. 9A–D). The prevalence of A2ML1-positive cells was elevated in tumor tissues versus normal

pancreas, observed consistently across primary lesions, nodal metastases, and liver metastases.

We systematically evaluated A2ML1 expression in five paired PC and adjacent normal tissues using RT-qPCR, western blotting, and immunohistochemistry. A2ML1 was significantly overexpressed in tumor tissues compared to normal counterparts (Fig. 10A–C). We quantified A2ML1 expression in H6C7 and four PC cell lines using RT-qPCR and western blotting analysis. A2ML1

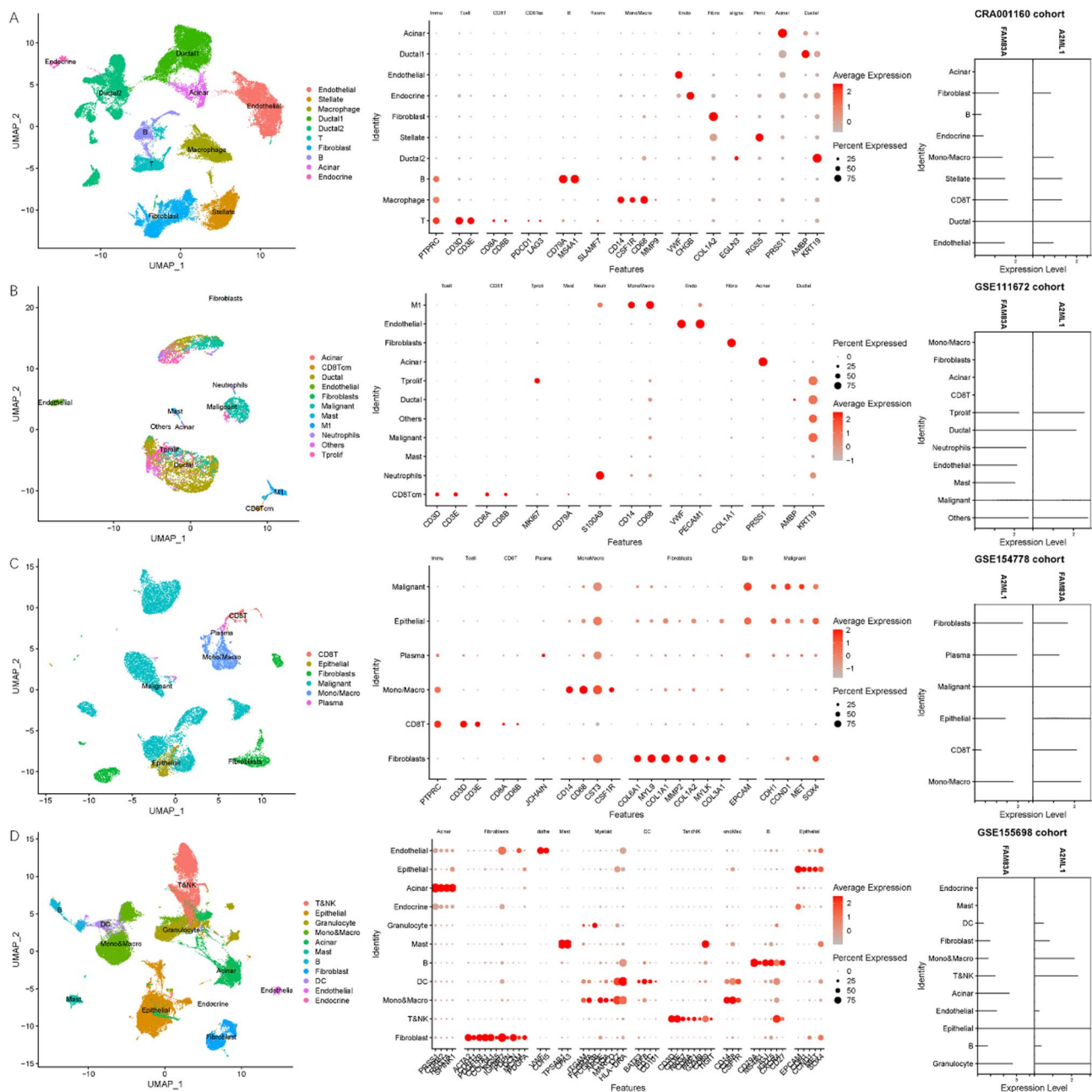


Fig. 8 Four scRNA-seq atlases of PC. We mainly showed the UMAP projections of various cell types, alongside the expression patterns of marker genes and BM-related genes. (A) CRA001160 cohort. (B) GSE111672 cohort. (C) GSE154778 cohort. (D) GSE155698 cohort

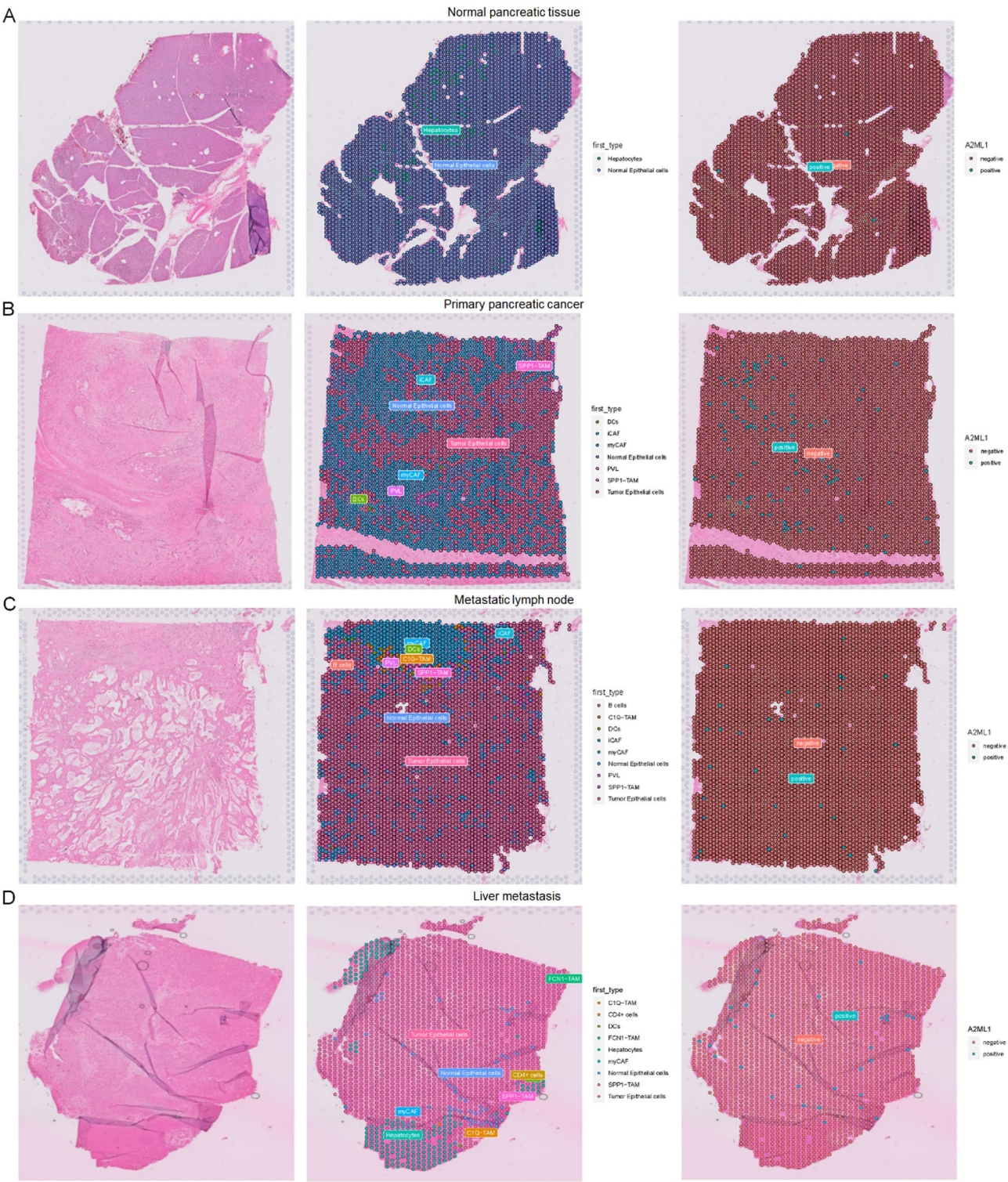


Fig. 9 Spatial transcriptomic profiling was performed to evaluate the distribution patterns of A2ML1-high-expressing cells in pancreatic cancer versus normal pancreatic tissues. **(A)** Normal pancreatic tissues. **(B)** Primary pancreatic cancer. **(C)** Metastatic lymph node. **(D)** Liver metastasis of PC

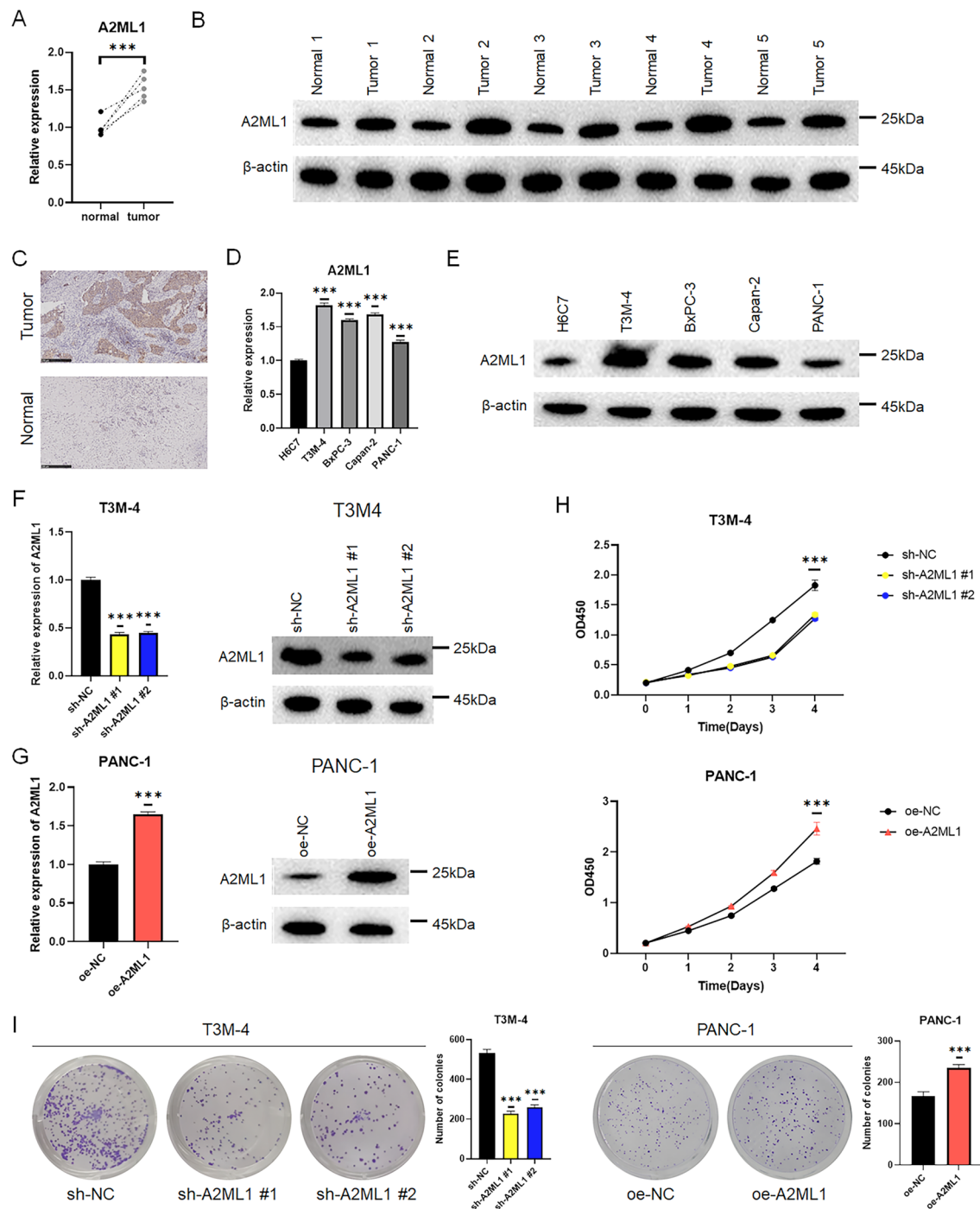


Fig. 10 A2ML1 was significantly overexpressed in PC and functionally promoted tumor cell proliferation. **(A–C)** A2ML1 expression was significantly elevated in PC tissues compared to normal tissues. **(A)** RT-qPCR. **(B)** Western blotting. **(C)** Immunohistochemistry. **(D–E)** A2ML1 expression was markedly upregulated in PC cell lines compared to the normal pancreatic ductal epithelial cell line H6C7. **(D)** RT-qPCR. **(E)** Western blotting. **(F)** A2ML1 was silenced in T3M-4 using RNA interference technology. Left: RT-qPCR. Right: Western blotting. **(G)** A2ML1 was overexpressed in PANC-1 cells through lentiviral transduction. Left: RT-qPCR. Right: Western blotting. **(H)** Cell proliferation was assessed using the CCK-8 assay. Upper: A2ML1 knockdown inhibited PC cell proliferation. Lower: A2ML1 overexpression promoted PC cell proliferation. **(I)** Clonogenic potential of PC cells was evaluated using colony formation assays. Left: A2ML1 knockdown inhibited PC cell proliferation. Right: A2ML1 overexpression promoted PC cell proliferation

levels were significantly elevated in cancer cell lines compared to H6C7 (Fig. 10D, E).

A2ML1 modulated proliferation, migration, invasion, apoptosis, and cell cycle progression in PC cells

A2ML1 knockdown in T3M4 cells and overexpression in PANC-1 cells were performed to investigate its functional roles (Fig. 10F, G). Cell proliferative capacity was assessed using CCK-8 and colony formation assays (Fig. 10H, I). A2ML1 knockdown significantly inhibited PC cell proliferation, whereas its overexpression promoted cell growth. Cell migratory capacity was evaluated using wound healing and Transwell migration assays (Fig. 11A, B). A2ML1 knockdown significantly inhibited PC cell migration, while its overexpression enhanced migratory capacity. Consistent results were observed in the Transwell invasion assays (Fig. 11C). EMT marker alterations were examined at both RNA and protein levels (Fig. 11D). A2ML1 knockdown reduced CDH2 (N-cadherin), VIM (Vimentin), and SNAIL (Snail) expression while increasing CDH1 (E-cadherin) levels. Conversely, A2ML1 overexpression produced opposite effects. Flow cytometry was employed to analyze apoptosis rates and cell cycle distribution (Fig. 11E-F). GSEA revealed significant enrichment of EMT pathways in both CS1 subtype and A2ML1-high PC samples (Supplementary Fig. 7A, B). A2ML1 knockdown significantly promoted apoptosis in PC cells, whereas its overexpression inhibited programmed cell death. A2ML1 knockdown induced G1-phase cell cycle arrest in PC cells.

To further investigate the role of A2ML1 in EMT, we conducted additional experiments using patient-derived organoid models. According to Ronnie et al., PC organoids exhibited two distinct morphological patterns: glandular and solid (Fig. 12A) [46]. TGF- β 1 treatment could induce EMT, as evidenced by characteristic morphological transitions in organoid structures. Our experimental results further validated this observation (Fig. 12B). Notably, A2ML1 overexpression similarly induced mesenchymal transition in PC organoids, characterized by decreased E-cadherin and increased Vimentin expression (Fig. 12C).

A2ML1 facilitated EMT by activating the KRAS/MAPK signaling pathway through downregulation of LZTR1

Given that A2ML1 encoded a protease inhibitor, we explored its potential interacting proteins to further elucidate the mechanisms driving A2ML1-mediated EMT in PC. LZTR1 was identified as a potential A2ML1-interacting protein through bioinformatic prediction (Fig. 12D). Subsequently, we confirmed the protein-level interaction between A2ML1 and LZTR1 using co-immunoprecipitation assay (Fig. 12E). To determine the localization of A2ML1 and LZTR1 in PC cells, we

performed immunofluorescence assays. It indicated that both A2ML1 and LZTR1 were localized in the cytoplasm (Fig. 12F). These findings demonstrated that A2ML1 could interact with LZTR1 in PC cells.

Previous studies indicated that LZTR1 functioned as a crucial regulator of RAS signaling through ubiquitination [47, 48]. As a key member of the RAS family, KRAS played a crucial role in PC progression through multiple oncogenic signaling pathways. A2ML1 could modulate both the expression of LZTR1 and the activation of RAS (Fig. 13A, B). Specifically, silencing A2ML1 enhanced LZTR1 expression, whereas overexpressing A2ML1 suppressed it. Silencing A2ML1 inhibited RAS activity, whereas overexpressing A2ML1 enhanced it. The regulation of RAS activity by A2ML1 was modulated through the expression of LZTR1. We further demonstrated A2ML1-mediated regulation of KRAS ubiquitination, with knockdown enhancing and overexpression suppressing this post-translational modification (Fig. 13C). A2ML1 modulated Ras activity in PC cells, and this regulatory effect could be reversed by LZTR1 (Fig. 13D, Supplementary Fig. 7C-E).

Taiki et al. indicated that Lys-117 of KRAS was a primary target of LZTR1-mediated polyubiquitination [49]. Lys-117 represented a frequently mutated residue in KRAS across various malignancies, including non-small cell lung cancer, colorectal cancer, and multiple myeloma [50, 51]. We hypothesized that A2ML1 could regulate KRAS/MAPK pathway and EMT through LZTR1-mediated modulation of KRAS at the Lys-117 residue. Consequently, we generated a lysine-to-arginine mutation at residue 117 (K117R) in KRAS. In wild-type (WT) PC cells, downregulation of A2ML1 significantly suppressed EMT and MAPK signaling compared to the control group (Fig. 13E-G). However, in KRAS-K117R PC cells, no significant changes in MAPK signaling and EMT were observed upon A2ML1 knockdown relative to controls (Fig. 13E-G). The regulatory mechanism by which EMT and the MAPK pathway were modulated by A2ML1 in PC cells was illustrated in Fig. 14. A schematic overview of the entire study workflow was presented in Fig. 15.

Discussion

Accurate molecular subtyping of PC facilitated comprehensive understanding of disease progression and informs clinical decision-making. In this study, we leveraged multi-omics data from the TCGA-PAAD cohort to uncover a basal-like subtype of PC, which exhibited distinct characteristics of prognosis, immune infiltration patterns, immune-related gene expression profiles, and pathway enrichment. The robustness of this classification system was validated across multiple independent external cohorts. Previous studies have proposed multiple molecular classification systems for PC. We developed

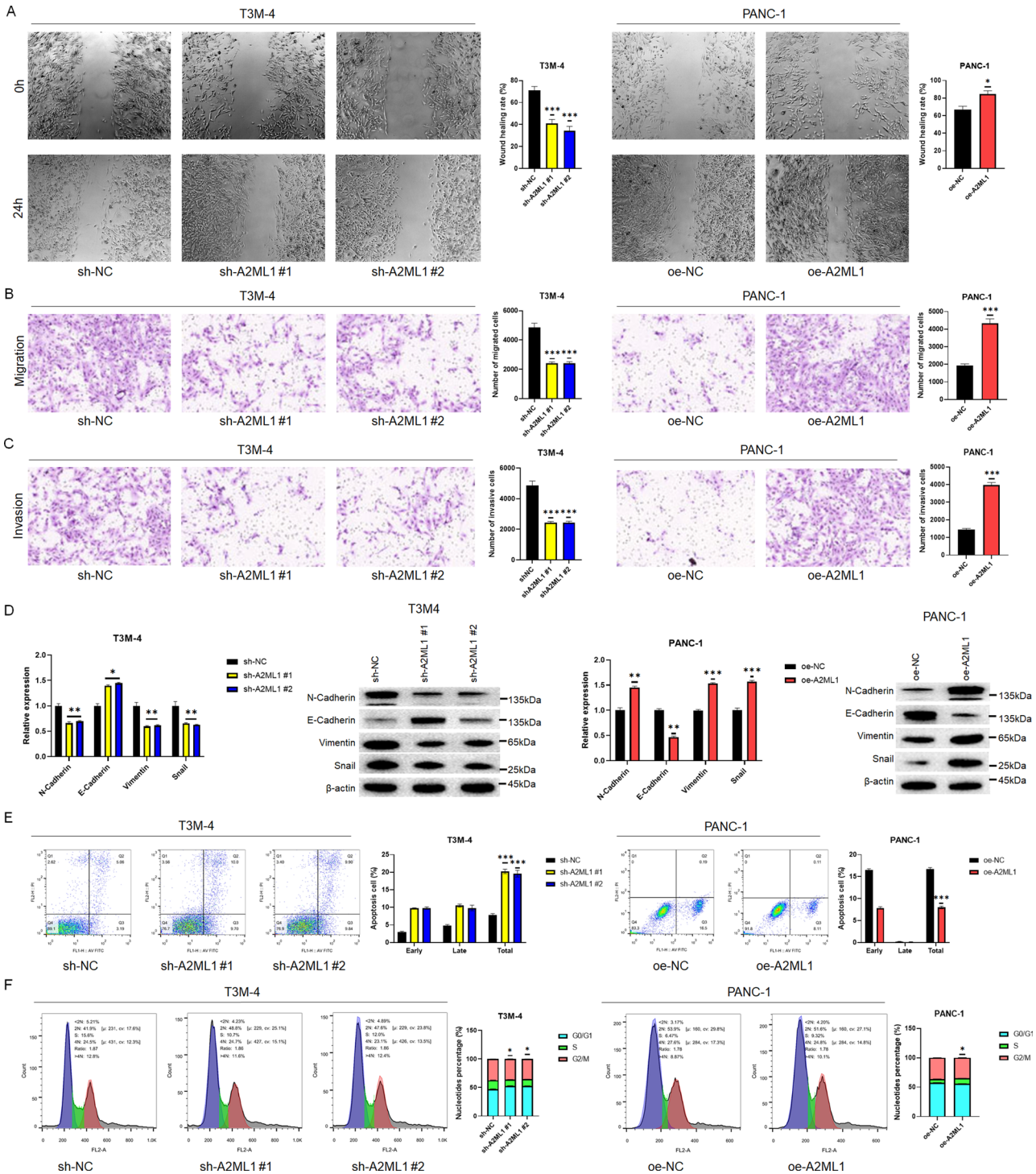


Fig. 11 A2ML1 promoted PC cell migration and invasion, inhibited apoptosis, and induced cell cycle arrest. **(A)** Cell migratory capacity was assessed using the wound healing assay. Left: A2ML1 knockdown inhibited PC cell migration. Right: A2ML1 overexpression promoted PC cell migration. **(B)** Cell migratory capacity was assessed using the Transwell assay. Left: A2ML1 knockdown inhibited PC cell migration. Right: A2ML1 overexpression promoted PC cell migration. **(C)** Cell invasion capacity was assessed using the Transwell assay. Left: A2ML1 knockdown inhibited PC cell migration. Right: A2ML1 overexpression promoted PC cell migration. **(D)** The expression levels of EMT-related proteins were detected using western blotting analysis. Left: A2ML1 knockdown suppressed EMT in PC cells. Right: A2ML1 overexpression enhanced EMT in PC cells. **(E)** Apoptosis rates in PC cells were quantified using flow cytometry. Left: A2ML1 knockdown promoted apoptosis in PC cells. Right: A2ML1 overexpression inhibited apoptosis in PC cells. **(F)** Cell cycle distribution in PC cells was analyzed using flow cytometry. Left: A2ML1 knockdown induced S-phase cell cycle arrest in PC cells. Right: A2ML1 overexpression promoted S-phase cell cycle progression in PC cells

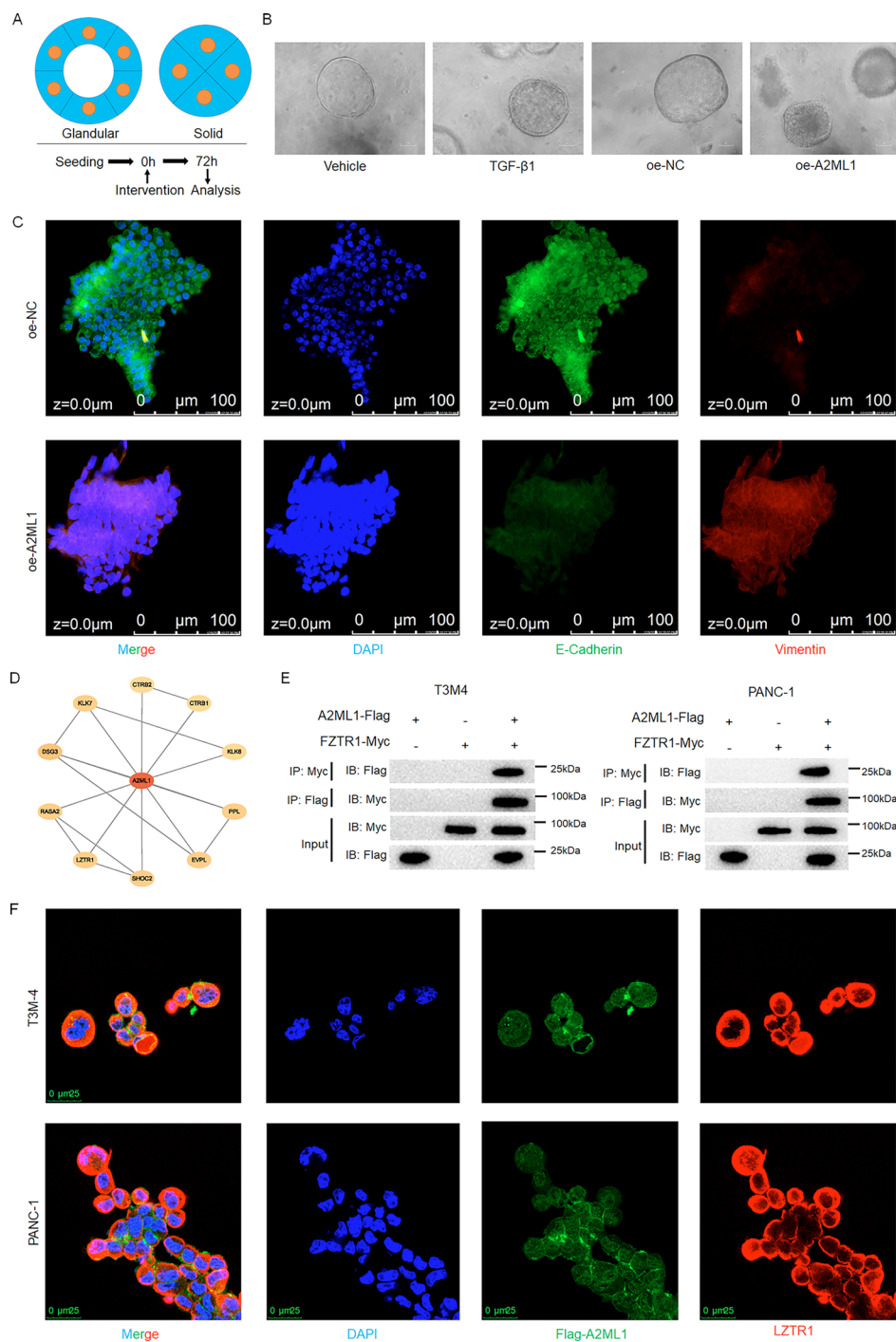


Fig. 12 A2ML1 induced EMT in PC organoids and demonstrated physical interaction with LZTR1. **(A)** Morphological classification and experimental interventions in PC organoids. **(B)** Recombinant TGF-β1 treatment and A2ML1 overexpression promoted morphological alterations in PC organoids. **(C)** Immunofluorescence analysis revealed significant alterations in EMT marker expression patterns in PC organoids. **(D)** Protein-protein interaction network centered on A2ML1. **(E)** Co-immunoprecipitation assays demonstrated a physical interaction between A2ML1 and LZTR1 in PC cell lines. **(F)** Immunofluorescence staining revealed that both A2ML1 and LZTR1 were localized in the cytoplasm of PC cells

and validated a prognostic signature using ridge regression analysis based on 23 prognostic-related genes. The risk scores predicted by this signature showed significant correlations with drug sensitivity, tumor grade, resection margin status, and classical subtypes in PC samples. Bailey et al. identified four molecular subtypes of PC through comprehensive genomic analyses, including squamous, immunogenic, pancreatic progenitor, and

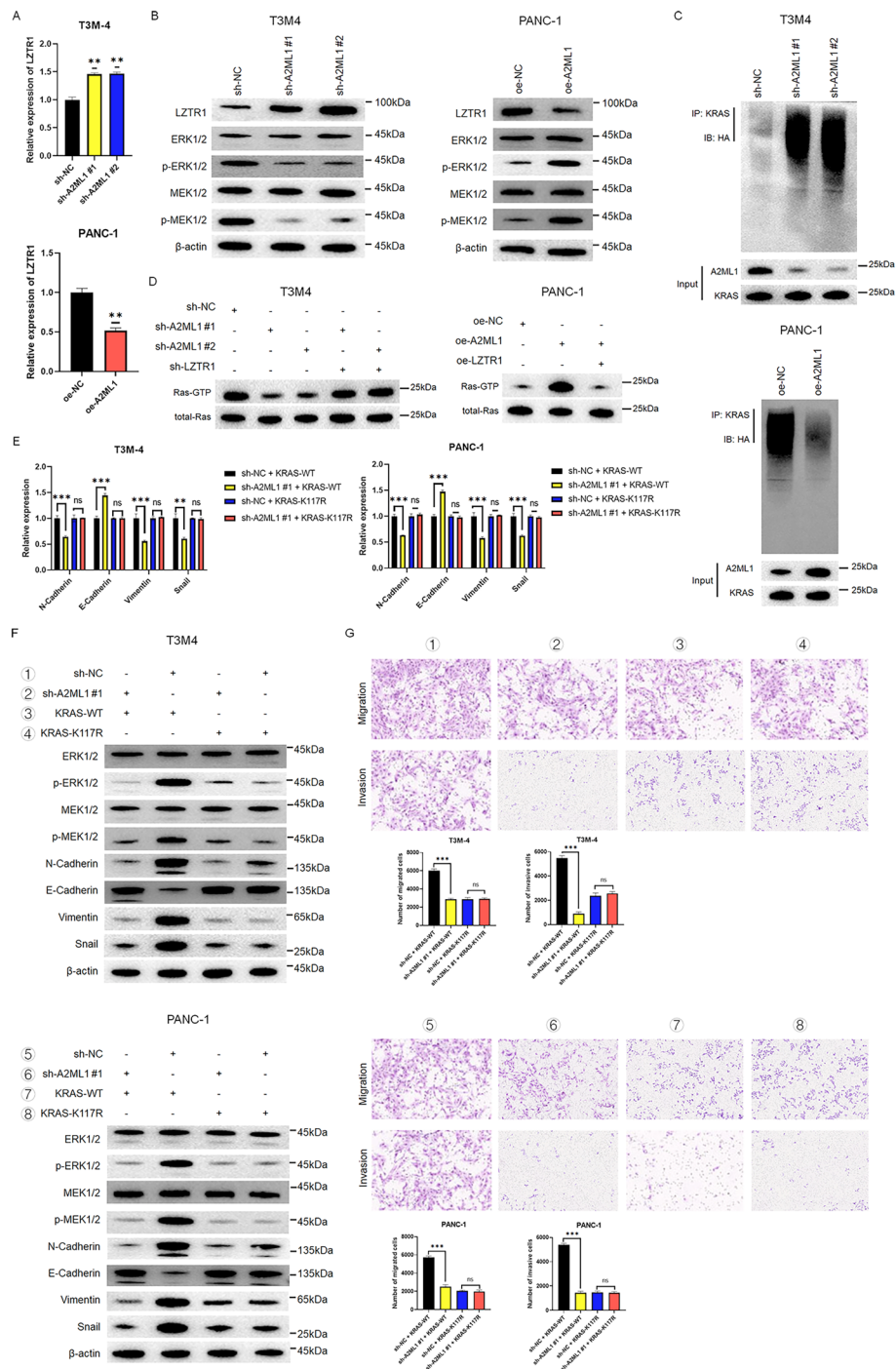


Fig. 13 A2ML1 promotes EMT by downregulating LZTR1 and activating the KRAS/MAPK signaling pathway. **(A)** RT-qPCR was performed to validate the regulatory effect of A2ML1 on LZTR1 expression. Up: Silencing A2ML1 upregulated the expression of LZTR1. Down: Overexpression of A2ML1 suppressed the expression of LZTR1. **(B)** The regulatory roles of A2ML1 in modulating LZTR1 expression and MAPK signaling pathway activity were confirmed by western blotting. Left: Silencing A2ML1 upregulated LZTR1 expression and inhibited the MAPK signaling pathway. Right: Overexpression of A2ML1 down-regulated LZTR1 expression and activated the MAPK signaling pathway. **(C)** A2ML1 modulated the ubiquitination level of KRAS. Up: Silencing A2ML1 enhanced the ubiquitination level of KRAS. Down: Overexpression of A2ML1 reduced the ubiquitination level of KRAS. **(D)** The activation status of Ras was assessed using a Ras activation assay. Left: Silencing A2ML1 suppressed Ras activity, which could be rescued by the knockdown of LZTR1. Right: Overexpression of A2ML1 enhanced Ras activity, which could be rescued by the overexpression of LZTR1. **(E)** RT-qPCR was performed to measure the levels of EMT markers across different experimental groups. Left: T3M-4 cell line. Right: PANC-1 cell line. **(F)** The levels of EMT markers and the activation status of the MAPK pathway were assessed by western blotting. Up: T3M-4 cell line. Down: PANC-1 cell line. **(G)** Transwell assays were performed to evaluate cell migration and invasion capabilities. Up: T3M-4 cell line. Down: PANC-1 cell line

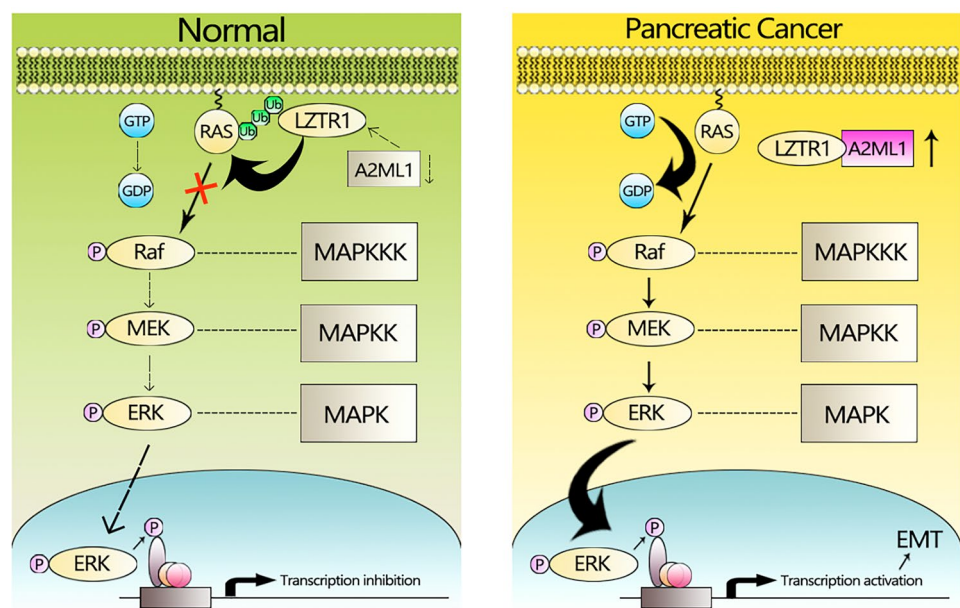


Fig. 14 Schematic diagram illustrating the regulatory mechanism of A2ML1 on the KRAS/MAPK signaling pathway

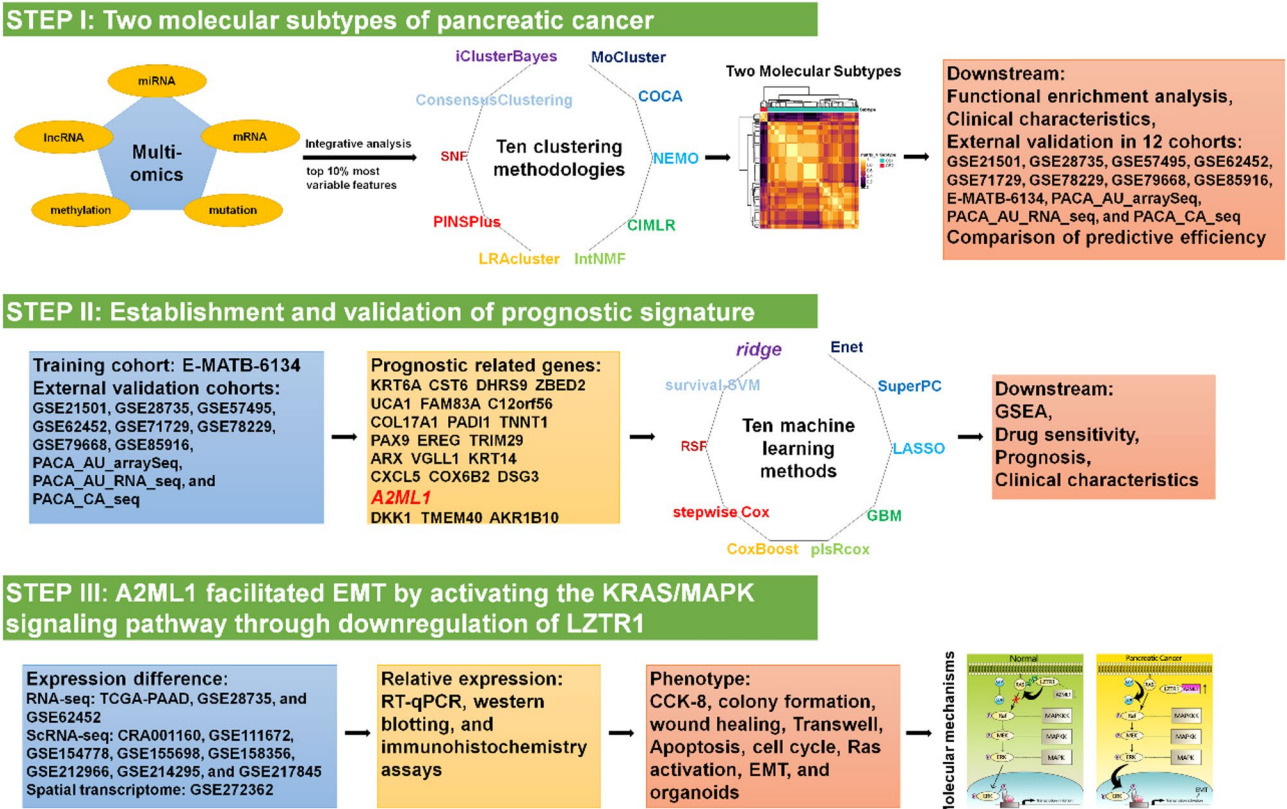


Fig. 15 Flow chart of this study

aberrantly differentiated endocrine exocrine (ADEX) [8]. The squamous subtype was characterized by TP53 and KDM6A mutations with poor prognosis, while the immunogenic subtype showed significant immune cell infiltration. The pancreatic progenitor subtype exhibited transcriptional networks involved in early pancreatic development, and the ADEX subtype demonstrated endocrine and exocrine differentiation features. Collisson et al. proposed a molecular classification system for PC, identifying three distinct subtypes: classical, quasimesenchymal, and exocrine-like [10]. The classical subtype demonstrated epithelial differentiation and relatively better prognosis, while the quasimesenchymal subtype exhibited mesenchymal features and increased aggressiveness. The exocrine-like subtype showed characteristics of normal pancreas tissue with unique metabolic profiles. Moffitt et al. developed a novel dual classification system for PC through virtual microdissection, identifying distinct tumor- and stroma-specific subtypes [11]. The tumor-specific classification comprised classical and basal-like subtypes, characterized by epithelial and mesenchymal features respectively, while the stroma-specific classification includes normal and activated subtypes based on stromal characteristics. Their studies provided valuable insights into the molecular heterogeneity of PC and its potential implications for personalized treatment strategies. Risk scores generated by our predictive model revealed significantly higher values in basal-like compared to classical subtypes, consistent with previous molecular classification findings.

By integrating ten distinct classification algorithms, we identified two molecular subtypes of PC, with CS2 exhibiting significantly better clinical outcomes than CS1. The two subtypes exhibited distinct immune infiltration patterns, including neutrophil (TIMER), macrophage (TIMER), NK cell activated (CIBERSORT), and NK cell (QUANTISEQ). CS1-associated genes were predominantly enriched in immune-related processes, while CS2-related genes showed significant enrichment in RNA transport and glutamate metabolism pathways. PC was characterized by distinct features including angiogenesis, EMT, glycolysis, hypoxia, and inflammatory response, which collectively contribute to its aggressive progression [52, 53]. Multiple signaling pathways played crucial roles in PC initiation and progression, including Hedgehog signaling, p53 pathway, PI3K/AKT/MTOR signaling, and KRAS signaling [54, 55]. Notably, KRAS serve as one of the most important therapeutic targets for PC [56]. The distinct biological behaviors and pathway enrichment patterns between the two molecular subtypes revealed their unique characteristics, potentially explaining their differential clinical outcomes. Furthermore, we identified significant differences in the expression levels of immunomodulators and immune checkpoints between

the two molecular subtypes. Our study primarily focuses on KRAS/MAPK pathway-mediated EMT, a well-documented process in previous studies [57–59]. Targeting the MAPK signaling pathway, a crucial therapeutic target in oncology, could demonstrate efficacy against refractory tumors [60].

In our study, we identified 23 prognostic genes and developed a predictive signature using 101 machine learning approaches, with ridge regression demonstrating optimal performance based on previous research [61, 62]. Sample risk scores were computed by summing the product of each gene's expression level and its corresponding coefficient. This signature consistently ranked among the top-performing prognostic tools when benchmarked against existing signatures in independent PC cohorts. A2ML1 was selected for in-depth functional investigation based on its prognostic significance. Through DEGs analysis between tumor and normal samples, we identified A2ML1 as a key regulator in PC progression. Mechanistically, A2ML1 promoted EMT of PC by downregulating LZTR1 expression and activating the KRAS/MAPK signaling pathway. A2ML1 encoded a protein similar in structure and function to alpha-2-macroglobulin (A2M). As a monomeric protease inhibitor, A2ML1 predominantly expressed in epidermal granular layer, placenta, thymus, and testis, but not in epithelia of kidney, lung, small intestine, or colon [63]. Although its physiological role remained elusive, A2ML1 was hypothesized to modulate epidermal proteases crucial for keratinocyte differentiation and desquamation [64]. Pathologically, it was associated with paraneoplastic pemphigus and otitis media development [65–68]. The unique conformational structure of A2ML1 confers its protease inhibitory function, enabling interactions with both amine and hydroxyl groups [69, 70]. Research on A2ML1 in tumorigenesis remains limited, with only a few studies investigating its roles in non-small cell lung cancer, cervical cancer, glioblastoma, and breast cancer [71–74]. However, these studies primarily focused on correlating A2ML1 expression levels with tumor prognosis, lacking experimental validation of its functional impact and mechanistic roles in cancer progression. To elucidate the molecular mechanisms underlying A2ML1-mediated EMT in PC, we investigated potential protein-protein interactions. We demonstrated the interaction between A2ML1 and LZTR1, with A2ML1 significantly downregulating LZTR1 expression. LZTR1 was a member of the BTB-Kelch superfamily. As a tumor suppressor and Noonan syndrome-associated gene, LZTR1 exhibited germline and somatic mutations in various malignancies, including multiple schwannomas and glioblastoma [75–77]. LZTR1 could serve as an adaptor for the CUL3-based ubiquitin E3 ligase complex. LZTR1 specifically ubiquitinated RAS family proteins, resulting

in either proteasomal degradation or altered cellular localization, ultimately modulating RAS signaling pathways [47–49, 78]. LZTR1-mediated RAS degradation requires ubiquitination at multiple lysine residues. Single-site ubiquitination is insufficient to induce complete degradation. To further elucidate A2ML1-mediated regulation of LZTR1, we generated a K48 site mutation in KRAS, revealing the regulatory role of A2ML1 in KRAS/MAPK signaling.

RAS proteins, which included KRAS, NRAS, and HRAS, played a critical role in the pathogenesis of PC [79]. They were small GTPases that regulated cell proliferation, survival, and differentiation by transmitting signals from cell surface receptors to intracellular pathways, such as the MAPK/ERK and PI3K/AKT pathways. Among these, KRAS was the most frequently mutated oncogene in PC, with mutations occurring in approximately 90% of all cases [80]. The activating mutations in KRAS were most commonly at codon 12 (e.g., G12D, G12V) [81]. These mutations could lead to constitutive activation of downstream signaling pathways, promoting tumor initiation, progression, and metastasis. Nevertheless, targeting mutant KRAS was deemed as a significant challenge in PC therapy [82]. Historically, RAS proteins were considered undruggable due to their high affinity for GTP and the lack of deep hydrophobic pockets for small-molecule inhibitors [83]. Given this challenges, alternative strategies have been explored, including its downstream pathway inhibition, metabolic dependency, and immunotherapy [84, 85]. The MAPK pathway represents a promising therapeutic target for novel drug development [86]. Our study focused on the alterations in MAPK/ERK pathway activity and EMT levels induced by KRAS expression changes in PC. We innovatively uncovered the regulatory role of A2ML1 on KRAS activity and its underlying mechanisms, offering new hope for effective PC treatment.

The findings of this study provided crucial insights into the molecular heterogeneity of PC and its implications for personalized treatment approaches. However, it still had several limitations. First, potential bias might exist due to variability in sample size and platform differences across cohorts. Second, prognostic analysis was limited to consensus genes due to inconsistencies in gene annotation and probe design among datasets. Finally, as the structural biology of A2ML1 became better characterized, future studies should further elucidate its regulatory mechanisms on LZTR1 [69].

In conclusion, we identified molecular basal-like subtype of PC with distinct prognostic outcomes through integrative multi-omics analysis and diverse classification algorithms. Utilizing 23 prognostic genes, we developed an effective predictive signature using ridge regression. A2ML1, significantly overexpressed in PC,

played a crucial role in tumor progression. Mechanistically, A2ML1 promoted EMT by upregulating the KRAS/MAPK pathway through interaction with LZTR1. Our findings provide novel insights for precision medicine in PC treatment.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12967-025-06711-z>.

Supplementary Material 1

Supplementary Material 2: Figure S1: Prediction of molecular subtypes in an external cohort and validate the subtype classification using survival analysis. (A) GSE78229 cohort. (B) GSE79668 cohort. (C) GSE85916 cohort. (D) PACA_AU_arraySeq cohort. (E) PACA_AU_RNA_seq cohort. (F) PACA_CA_seq cohort.

Supplementary Material 3: Figure S2: Univariate Cox regression analysis was systematically conducted for the 23 candidate genes across 12 independent validation cohorts. (A) GSE21501 cohort. (B) GSE28735 cohort. (C) GSE57495 cohort. (D) GSE62452 cohort. (E) GSE71729 cohort. (F) GSE78229 cohort. (G) GSE79668 cohort. (H) GSE85916 cohort. (I) PACA-AU_arraySeq cohort. (J) PACA-AU-RNA_seq cohort. (K) PACA-CA-seq cohort. (L) TCGA-PAAD cohort.

Supplementary Material 4: Figure S3: The heatmap illustrated the sample risk scores, risk groupings, survival statuses, and gene expression profiles across 12 validation cohorts. (A) GSE21501 cohort. (B) GSE28735 cohort. (C) GSE57495 cohort. (D) GSE62452 cohort. (E) GSE71729 cohort. (F) GSE78229 cohort. (G) GSE79668 cohort. (H) GSE85916 cohort. (I) PACA-AU_arraySeq cohort. (J) PACA-AU-RNA_seq cohort. (K) PACA-CA-seq cohort. (L) TCGA-PAAD cohort.

Supplementary Material 5: Figure S4: Survival differences between high-risk and low-risk group samples across 12 validation cohorts. (A) GSE21501 cohort. (B) GSE28735 cohort. (C) GSE57495 cohort. (D) GSE62452 cohort. (E) GSE71729 cohort. (F) GSE78229 cohort. (G) GSE79668 cohort. (H) GSE85916 cohort. (I) PACA-AU_arraySeq cohort. (J) PACA-AU-RNA_seq cohort. (K) PACA-CA-seq cohort. (L) TCGA-PAAD cohort.

Supplementary Material 6: Figure S5: This signature's prognostic value was compared to previous signatures across independent PC cohorts. (A) C-index of 170 signatures in GSE79668 cohort. (B) C-index of 144 signatures in GSE85916 cohort. (C) C-index of 162 signatures in PACA-AU_arraySeq cohort. (D) C-index of 173 signatures in PACA-AU-RNA_seq cohort. (E) C-index of 172 signatures in PACA-CA-seq cohort. (F) C-index of 171 signatures in TCGA-PAAD cohort.

Supplementary Material 7: Figure S6: Four scRNA-seq atlases of PC. We mainly showed the UMAP projections of various cell types, alongside the expression patterns of marker genes and BM-related genes. (A) GSE158356 cohort. (B) GSE212966 cohort. (C) GSE214295 cohort. (D) GSE217845 cohort.

Supplementary Material 8: Figure S7: GSEA identified differential EMT pathway activation among PC groups. RT-qPCR and western blotting analyses demonstrated the expression levels of target genes in PC cell lines. (A-B) Groups were defined by (A) molecular subtypes and (B) A2ML1 expression. (C) LZTR1 was silenced in the T3M-4 cell line. Left: RT-qPCR. Right: western blotting. (D) LZTR1 was overexpressed in the PANC-1 cell line. Left: RT-qPCR. Right: western blotting. (E) A2ML1 was silenced in the PANC-1 cell line. Left: RT-qPCR. Right: western blotting.

Author contributions

Conceptualization, Jiachen Ge, Jianping Cai, Deyu Li and Lianyuan Tao; investigation, Jiachen Ge, Jianping Cai and Deyu Li; visualization, Jiachen Ge, Jianping Cai, Gaolei Zhang and Lianyuan Tao; supervision, Deyu Li and Lianyuan Tao; validation, Jianping Cai and Gaolei Zhang; funding acquisition, Jiachen Ge and Lianyuan Tao; project administration, Lianyuan Tao; software

and methodology, Jiachen Ge, Jianping Cai and Deyu Li; resources, Lianyuan Tao; writing—original draft preparation, Jiachen Ge, Jianping Cai and Deyu Li; writing—review and editing, Gaolei Zhang and Lianyuan Tao. All authors have read and agreed to the published version of the manuscript.

Funding

This study was supported by grants from the Joint Project of Medical Science and Technology Research Program of Henan Province (LHGJ20230007 and LHGJ20230033), Henan Province Health Science and Technology Research Project (242102310141 and 252102311080), the “Three 100” Plan Overseas Study Program for Medical Science and Technology Talents of Henan Province Medical Sciences Academy (H20240044), and Henan Province Health Science and Technology Innovation Excellent Young Talents Training Project (No. YXKC2020044).

Data availability

The public data used in this study are available in TCGA, ArrayExpress, and GEO database. The molecular subtyping workflow, including establishment, validation, and analysis, has been implemented in the R package “MOVICS” (<https://github.com/xlucpu/MOVICS>), which we developed for comprehensive multi-omics integration and visualization. The original data and code of this study is available on the Github website (https://github.com/jcge/pcmovics_machinelearning/). Further inquiries can be directed to the corresponding author.

Declarations

Ethics approval and consent to participate

The studies involving human participants were reviewed and approved by the Ethics Committee of Henan Provincial People's Hospital, People's Hospital of Zhengzhou University (2019 No.58).

Consent for publication

All authors agree to submit the article for publication.

Competing interests

The authors declare no competing interest in this work.

Author details

¹Department of Hepatobiliary Surgery, Henan Provincial People's Hospital, People's Hospital of Zhengzhou University, Zhengzhou 450003, People's Republic of China

²Laboratory of Gene Regulation and Signal Transduction, Departments of Pharmacology and Pathology, School of Medicine, University of California, San Diego, La Jolla, CA, USA

Received: 7 March 2025 / Accepted: 3 June 2025

Published online: 04 July 2025

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71:209–49. <https://doi.org/10.3322/caac.21660>.
2. Li T, Lin C, Wang W. Global, regional, and National burden of pancreatic cancer from 1990 to 2021, its attributable risk factors, and projections to 2050: a systematic analysis of the global burden of disease study 2021. *BMC Cancer*. 2025;25:189. <https://doi.org/10.1186/s12885-025-13597-z>.
3. Huang J, Lok V, Ngai CH, Zhang L, Yuan J, Lao XQ, et al. Worldwide burden of, risk factors for, and trends in pancreatic Cancer. *Gastroenterology*. 2021;160:744–54. <https://doi.org/10.1053/j.gastro.2020.10.007>.
4. Klein AP. Pancreatic cancer epidemiology: Understanding the role of lifestyle and inherited risk factors. *Nat Rev Gastroenterol Hepatol*. 2021;18:493–502. <https://doi.org/10.1038/s41575-021-00457-x>.
5. Hong SB, Lee SS, Kim JH, Kim HJ, Byun JH, Hong SM, et al. Pancreatic Cancer CT: prediction of resectability according to NCCN criteria. *Radiology*. 2018;289:710–18. <https://doi.org/10.1148/radiol.2018180628>.
6. Kim SS, Lee S, Lee HS, Bang S, Han K, Park MS. Retrospective evaluation of treatment response in patients with nonmetastatic pancreatic Cancer using CT and CA 19–9. *Radiology*. 2022;303:548–56. <https://doi.org/10.1148/radiol.212236>.
7. Dreyer SB, Upstill-Goddard R, Legrini A, Biankin AV, Glasgow Precision Oncology L, Jamieson NB, et al. Genomic and molecular analyses identify molecular subtypes of pancreatic Cancer recurrence. *Gastroenterology*. 2022;162:320–24. <https://doi.org/10.1053/j.gastro.2021.09.022>. e4.
8. Bailey P, Chang DK, Nones K, Johns AL, Patch AM, Gingras MC, et al. Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature*. 2016;531:47–52. <https://doi.org/10.1038/nature16965>.
9. Sivapalan L, Kocher HM, Ross-Adams H, Chelala C. The molecular landscape of pancreatic ductal adenocarcinoma. *Pancreatol*. 2022;22:925–36. <https://doi.org/10.1016/j.pan.2022.07.010>.
10. Collisson EA, Sadanandam A, Olson P, Gibb WJ, Truitt M, Gu S, et al. Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nat Med*. 2011;17:500–3. <https://doi.org/10.1038/nm.2344>.
11. Moffitt RA, Marayati R, Flate EL, Volmar KE, Loeza SG, Hoadley KA, et al. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat Genet*. 2015;47:1168–78. <https://doi.org/10.1038/ng.3398>.
12. Waddell N, Pajic M, Patch AM, Chang DK, Kassahn KS, Bailey P, et al. Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature*. 2015;518:495–501. <https://doi.org/10.1038/nature14169>.
13. Danilova L, Ho WJ, Zhu Q, Vithayathil T, De Jesus-Acosta A, Azad NS, et al. Programmed cell death Ligand-1 (PD-L1) and CD8 expression profiling identify an Immunologic subtype of pancreatic ductal adenocarcinomas with favorable survival. *Cancer Immunol Res*. 2019;7:886–95. <https://doi.org/10.1158/2666.CIR-18-0822>.
14. Yang B, Zhou M, Wu Y, Ma Y, Tan Q, Yuan W, et al. The impact of immune microenvironment on the prognosis of pancreatic ductal adenocarcinoma based on Multi-Omics analysis. *Front Immunol*. 2021;12:769047. <https://doi.org/10.3389/fimmu.2021.769047>.
15. Liu F, Gao A, Zhang M, Li Y, Zhang F, Herman JG, et al. Methylation of FAM110C is a synthetic lethal marker for ATR/CHK1 inhibitors in pancreatic cancer. *J Transl Int Med*. 2024;12:274–87. <https://doi.org/10.2478/jtim-2023-0128>.
16. Zhang N, Kandalai S, Zhou X, Hossain F, Zheng Q. Applying multi-omics toward tumor Microbiome research. *Imeta*. 2023;2:e73. <https://doi.org/10.1002/imt2.73>.
17. Stratford JK, Yan F, Hill RA, Major MB, Graves LM, Der CJ, et al. Genetic and Pharmacological Inhibition of TTK impairs pancreatic cancer cell line growth by inducing lethal chromosomal instability. *PLoS ONE*. 2017;12:e0174863. <https://doi.org/10.1371/journal.pone.0174863>.
18. Zhang G, He P, Tan H, Budhu A, Gaedcke J, Ghadimi BM, et al. Integration of metabolomics and transcriptomics revealed a fatty acid network exerting growth inhibitory effects in human pancreatic cancer. *Clin Cancer Res*. 2013;19:4983–93. <https://doi.org/10.1158/1078-0432.CCR-13-0209>.
19. Chen DT, Davis-Yadley AH, Huang PY, Husain K, Centeno BA, Permut-Wey J, et al. Prognostic Fifteen-Gene signature for early stage pancreatic ductal adenocarcinoma. *PLoS ONE*. 2015;10:e0133562. <https://doi.org/10.1371/journal.pone.0133562>.
20. Yang S, He P, Wang J, Schetter A, Tang W, Funamizu N, et al. A novel MIF signaling pathway drives the malignant character of pancreatic Cancer by targeting NR3C2. *Cancer Res*. 2016;76:3838–50. <https://doi.org/10.1158/0008-5472.CAN-15-2841>.
21. Wang J, Yang S, He P, Schetter AJ, Gaedcke J, Ghadimi BM, et al. Endothelial nitric oxide synthase traffic inducer (NOSTRIN) is a negative regulator of disease aggressiveness in pancreatic Cancer. *Clin Cancer Res*. 2016;22:5992–6001. <https://doi.org/10.1158/1078-0432.CCR-16-0511>.
22. Kirby MK, Ramaker RC, Gertz J, Davis NS, Johnston BE, Oliver PG, et al. RNA sequencing of pancreatic adenocarcinoma tumors yields novel expression patterns associated with long-term survival and reveals a role for ANGPTL4. *Mol Oncol*. 2016;10:1169–82. <https://doi.org/10.1016/j.molonc.2016.05.004>.
23. Puleo F, Nicolle R, Blum Y, Cros J, Marisa L, Demetter P, et al. Stratification of pancreatic ductal adenocarcinomas based on tumor and microenvironment features. *Gastroenterology*. 2018;155:1999–2013. <https://doi.org/10.1053/j.gastro.2018.08.033>. e3.
24. Peng J, Sun BF, Chen CY, Zhou JY, Chen YS, Chen H, et al. Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. *Cell Res*. 2019;29:725–38. <https://doi.org/10.1038/s41422-019-0195-y>.
25. Moncada R, Barkley D, Wagner F, Chiodini M, Devlin JC, Baron M, et al. Integrating microarray-based Spatial transcriptomics and single-cell RNA-seq

- reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nat Biotechnol.* 2020;38:333–42. <https://doi.org/10.1038/s41587-019-0392-8>.
26. Lin W, Noel P, Borazanci EH, Lee J, Amini A, Han IW, et al. Single-cell transcriptome analysis of tumor and stromal compartments of pancreatic ductal adenocarcinoma primary tumors and metastatic lesions. *Genome Med.* 2020;12:80. <https://doi.org/10.1186/s13073-020-00776-9>.
27. Steele NG, Carpenter ES, Kemp SB, Siriwhorachai VR, The S, Delrosario L, et al. Multimodal mapping of the tumor and peripheral blood immune landscape in human pancreatic Cancer. *Nat Cancer.* 2020;1:1097–112. <https://doi.org/10.1038/s43018-020-00121-4>.
28. Kemp SB, Steele NG, Carpenter ES, Donahue KL, Bushnell GG, Morris AH, et al. Pancreatic cancer is marked by complement-high blood monocytes and tumor-associated macrophages. *Life Sci Alliance.* 2021;4. <https://doi.org/10.26508/lsa.202000935>.
29. Chen K, Wang Q, Liu X, Tian X, Dong A, Yang Y. Immune profiling and prognostic model of pancreatic cancer using quantitative pathology and single-cell RNA sequencing. *J Transl Med.* 2023;21:210. <https://doi.org/10.1186/s12967-023-04051-4>.
30. Chen K, Ma Y, Zhong X, Lan J, Long D, Tian X, et al. Single-cell transcriptome profiling of primary tumors and paired organoids of pancreatobiliary cancer. *Cancer Lett.* 2024;582:216586. <https://doi.org/10.1016/j.canlet.2023.216586>.
31. Caronni N, La Terza F, Vittoria FM, Barbiera G, Mezzananza L, Cuzzola V, et al. IL-1beta(+) macrophages fuel pathogenic inflammation in pancreatic cancer. *Nature.* 2023;623:415–22. <https://doi.org/10.1038/s41586-023-06685-2>.
32. Lu X, Meng J, Zhou Y, Jiang L, Yan F. MOVICS: an R package for multi-omics integration and visualization in cancer subtyping. *Bioinformatics.* 2021;36:5539–41. <https://doi.org/10.1093/bioinformatics/btaa1018>.
33. Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, et al. TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res.* 2020;48:W509–14. <https://doi.org/10.1093/nar/gkaa407>.
34. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol.* 2018;1711:243–59. https://doi.org/10.1007/978-1-4939-7493-1_12.
35. Aran D, Hu Z, Butte AJ. XCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* 2017;18:220. <https://doi.org/10.1186/s13059-017-1349-1>.
36. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* 2016;17:218. <https://doi.org/10.1186/s13059-016-1070-5>.
37. Finotello F, Mayer C, Plattner C, Laschober G, Rieder D, Hackl H, et al. Molecular and Pharmacological modulators of the tumor immune contexture revealed by Deconvolution of RNA-seq data. *Genome Med.* 2019;11:34. <https://doi.org/10.1186/s13073-019-0638-6>.
38. Racle J, de Jonge K, Baumgaertner P, Speiser DE, Gfeller D. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *Elife.* 2017;6. <https://doi.org/10.7554/eLife.26476>.
39. Hu J, Yu A, Othmane B, Qiu D, Li H, Li C, et al. Siglec15 shapes a non-inflamed tumor microenvironment and predicts the molecular subtype in bladder cancer. *Theranostics.* 2021;11:3089–108. <https://doi.org/10.7150/tno.53649>.
40. Khaliq AM, Rajamohan M, Saeed O, Mansouri K, Adil A, Zhang C, et al. Spatial transcriptomic analysis of primary and metastatic pancreatic cancers highlights tumor microenvironmental heterogeneity. *Nat Genet.* 2024;56:2455–65. <https://doi.org/10.1038/s41588-024-01914-4>.
41. Zhou C, Zhu X, Liu N, Dong X, Zhang X, Huang H, et al. B-lymphoid tyrosine kinase-mediated FAM83A phosphorylation elevates pancreatic tumorigenesis through interacting with beta-catenin. *Signal Transduct Target Ther.* 2023;8:66. <https://doi.org/10.1038/s41392-022-01268-5>.
42. Zhou B, Zhou X, Zhan C, Jin M, Yan S. FAM83A promotes the progression and metastasis of human pancreatic neuroendocrine tumors by inducing the epithelial-mesenchymal transition via the PI3K/AKT and ERK pathways. *J Endocrinol Invest.* 2023;46:1115–30. <https://doi.org/10.1007/s40618-022-01959-4>.
43. Zhou C, Liang Y, Zhou L, Yan Y, Liu N, Zhang R, et al. TSPAN1 promotes autophagy flux and mediates Cooperation between WNT-CTNNB1 signaling and autophagy via the MIR454-FAM83A-TSPAN1 axis in pancreatic cancer. *Autophagy.* 2021;17:3175–95. <https://doi.org/10.1080/15548627.2020.1826689>.
44. Parameswaran N, Bartel CA, Hernandez-Sanchez W, Miskimen KL, Smigiel JM, Khalil AM, et al. A FAM83A positive Feed-back loop drives survival and tumorigenicity of pancreatic ductal adenocarcinomas. *Sci Rep.* 2019;9:13396. <https://doi.org/10.1038/s41598-019-49475-5>.
45. Chen S, Huang J, Liu Z, Liang Q, Zhang N, Jin Y. FAM83A is amplified and promotes cancer stem cell-like traits and chemoresistance in pancreatic cancer. *Oncogenesis.* 2017;6:e300. <https://doi.org/10.1038/oncsis.2017.3>.
46. Low R, Fung KY, Gao H, Preaudet A, Dagley LF, Yousef J, et al. S100 family proteins are linked to organoid morphology and EMT in pancreatic cancer. *Cell Death Differ.* 2023;30:1155–65. <https://doi.org/10.1038/s41418-023-01126-z>.
47. Bigenzahn JW, Collu GM, Kartnig F, Pieraks M, Vladimer GI, Heinz LX, et al. LZTR1 is a regulator of RAS ubiquitination and signaling. *Science.* 2018;362:1171–77. <https://doi.org/10.1126/science.aap8210>.
48. Steklov M, Pandolfi S, Baietti MF, Batiuk A, Carai P, Najm P, et al. Mutations in LZTR1 drive human disease by dysregulating RAS ubiquitination. *Science.* 2018;362:1177–82. <https://doi.org/10.1126/science.aap7607>.
49. Abe T, Umeki I, Kanno SI, Inoue SI, Niihori T, Aoki Y. LZTR1 facilitates polyubiquitination and degradation of RAS-GTPases. *Cell Death Differ.* 2020;27:1023–35. <https://doi.org/10.1038/s41418-019-0395-5>.
50. Johnson C, Burkhardt DL, Haigis KM. Classification of KRAS-Activating mutations and the implications for therapeutic intervention. *Cancer Discov.* 2022;12:913–23. <https://doi.org/10.1158/2159-8290.CD-22-0035>.
51. Cook JH, Melloni GEM, Gulhan DC, Park PJ, Haigis KM. The origins and genetic interactions of KRAS mutations are allele- and tissue-specific. *Nat Commun.* 2021;12:1808. <https://doi.org/10.1038/s41467-021-22125-z>.
52. Chang X, Liu X, Wang H, Yang X, Gu Y. Glycolysis in the progression of pancreatic cancer. *Am J Cancer Res.* 2022;12:861–72.
53. Tao J, Yang G, Zhou W, Qiu J, Chen G, Luo W, et al. Targeting hypoxic tumor microenvironment in pancreatic cancer. *J Hematol Oncol.* 2021;14:14. <https://doi.org/10.1186/s13045-020-01030-w>.
54. Huang L, Yang H, Chen K, Yuan J, Li J, Dai G, et al. The suppressive efficacy of THZ1 depends on KRAS mutation subtype and is associated with super-enhancer activity and the PI3K/AKT/mTOR signalling in pancreatic ductal adenocarcinoma: A hypothesis-generating study. *Clin Transl Med.* 2023;13:e1500. <https://doi.org/10.1002/ctm2.1500>.
55. Wormann SM, Song L, Ai J, Diakopoulos KN, Kurkowski MU, Gorgulu K, et al. Loss of P53 function activates JAK2-STAT3 signaling to promote pancreatic tumor growth, stroma modification, and gemcitabine resistance in mice and is associated with patient survival. *Gastroenterology.* 2016;151:180–e9312. <https://doi.org/10.1053/j.gastro.2016.03.010>.
56. Stickler S, Rath B, Hamilton G. Targeting KRAS in pancreatic cancer. *Oncol Res.* 2024;32:799–805. <https://doi.org/10.32604/or.2024.045356>.
57. Konen JM, Rodriguez BL, Padhye A, Ochieng JK, Gibson L, Diao L, et al. Dual Inhibition of MEK and AXL targets tumor cell heterogeneity and prevents resistant outgrowth mediated by the Epithelial-to-Mesenchymal transition in NSCLC. *Cancer Res.* 2021;81:1398–412. <https://doi.org/10.1158/0008-5472.CA-N-20-1895>.
58. Adachi Y, Ito K, Hayashi Y, Kimura R, Tan TZ, Yamaguchi R, et al. Epithelial-to-Mesenchymal transition is a cause of both intrinsic and acquired resistance to KRAS G12C inhibitor in KRAS G12C-Mutant Non-Small cell lung Cancer. *Clin Cancer Res.* 2020;26:5962–73. <https://doi.org/10.1158/1078-0432.CCR-20-2077>.
59. Blaj C, Schmidt EM, Lamprecht S, Hermeking H, Jung A, Kirchner T, et al. Oncogenic effects of high MAPK activity in colorectal Cancer mark progenitor cells and persist irrespective of RAS mutations. *Cancer Res.* 2017;77:1763–74. <https://doi.org/10.1158/0008-5472.CAN-16-2821>.
60. Wu L, Huang S, Tian W, Liu P, Xie Y, Qiu Y, et al. PIWI-interacting RNA-YBX1 inhibits proliferation and metastasis by the MAPK signaling pathway via YBX1 in triple-negative breast cancer. *Cell Death Discov.* 2024;10:7. <https://doi.org/10.1038/s41420-023-01771-w>.
61. Liu Z, Liu L, Weng S, Guo C, Dang Q, Xu H, et al. Machine learning-based integration develops an immune-derived LncRNA signature for improving outcomes in colorectal cancer. *Nat Commun.* 2022;13:816. <https://doi.org/10.1038/s41467-022-28421-6>.
62. Ge J, Tao M, Zhang G, Cai J, Li D, Tao L. New HCC subtypes based on CD8 Tex-Related LncRNA signature could predict prognosis, immunological and drug sensitivity characteristics of hepatocellular carcinoma. *J Hepatocell Carcinoma.* 2024;11:1331–55. <https://doi.org/10.2147/JHC.S459150>.
63. Galliano MF, Toulza E, Galliano H, Jonca N, Ishida-Yamamoto A, Serre G, et al. A novel protease inhibitor of the alpha2-macroglobulin family expressed in the human epidermis. *J Biol Chem.* 2006;281:5780–9. <https://doi.org/10.1074/jbc.M508017200>.
64. Yamamoto M, Miyai M, Matsumoto Y, Tsuboi R, Hibino T. Kallikrein-related peptidase-7 regulates caspase-14 maturation during keratinocyte

- terminal differentiation by generating an intermediate form. *J Biol Chem*. 2012;287:32825–34. <https://doi.org/10.1074/jbc.M112.357467>.
65. Numata S, Teye K, Tsuruta D, Sogame R, Ishii N, Koga H, et al. Anti-alpha-2-macroglobulin-like-1 autoantibodies are detected frequently and May be pathogenic in paraneoplastic pemphigus. *J Invest Dermatol*. 2013;133:1785–93. <https://doi.org/10.1038/jid.2013.65>.
66. Santos-Cortez RL, Chiong CM, Reyes-Quintos MR, Tantoco ML, Wang X, Acharya A, et al. Rare A2ML1 variants confer susceptibility to otitis media. *Nat Genet*. 2015;47:917–20. <https://doi.org/10.1038/ng.3347>.
67. Larson ED, Magno JPM, Steritz MJ, Llanes E, Cardwell J, Pedro M, et al. A2ML1 and otitis media: novel variants, differential expression, and relevant pathways. *Hum Mutat*. 2019;40:1156–71. <https://doi.org/10.1002/humu.23769>.
68. Elling CL, Gomez HZ, Lee NK, Hirsch SD, Santos-Cortez RLP. The A2ml1-Knock-out mouse as an animal model for non-syndromic otitis media. *Int J Pediatr Otorhinolaryngol*. 2024;181:111980. <https://doi.org/10.1016/j.jiporl.2024.111980>.
69. Nielsen NS, Zarantonello A, Harwood SL, Jensen KT, Kjøge K, Thøgersen IB, et al. Cryo-EM structures of human A2ML1 elucidate the protease-inhibitory mechanism of the A2M family. *Nat Commun*. 2022;13:3033. <https://doi.org/10.1038/s41467-022-30758-x>.
70. Harwood SL, Nielsen NS, Jensen KT, Nielsen PK, Thøgersen IB, Enghild JJ. alpha(2)-Macroglobulin-like protein 1 can conjugate and inhibit proteases through their hydroxyl groups, because of an enhanced reactivity of its thiol ester. *J Biol Chem*. 2020;295:16732–42. <https://doi.org/10.1074/jbc.RA120.015694>.
71. Babu S, Horowitz M, Delgado-Coka LA, Roa-Pena L, Akalin A, Escobar-Hoyos LF, et al. Keratin 17 and A2ML1 are negative prognostic biomarkers in non-small cell lung cancer. *Pathol Res Pract*. 2024;263:155643. <https://doi.org/10.1016/j.prp.2024.155643>.
72. Naveed M, Saad Mughal M, Aziz T, Jabeen K, Ali Khan A, Alhomrani M, et al. The prominence of the Broad-Spectrum protease inhibitor gene A2ML1 as a potential biomarker in cervical cancer diagnostics using immunotherapeutic and Multi-Omics approaches. *Int Immunopharmacol*. 2024;142:113126. <https://doi.org/10.1016/j.intimp.2024.113126>.
73. Hervas-Corpcion I, Navarro-Calvo J, Martin-Climent P, Iriarte-Gahete M, Geribaldi-Doldan N, Castro C, et al. Defining a correlative transcriptional signature associated with bulk histone H3 acetylation levels in adult glioblastomas. *Cells*. 2023;12. <https://doi.org/10.3390/cells12030374>.
74. Li Y, Zhou X, Liu J, Yin Y, Yuan X, Yang R, et al. Differentially expressed genes and key molecules of BRCA1/2-mutant breast cancer: evidence from bioinformatics analyses. *PeerJ*. 2020;8:e8403. <https://doi.org/10.7717/peerj.8403>.
75. Yamamoto GL, Agüena M, Gos M, Hung C, Pilch J, Fahiminiya S, et al. Rare variants in SOS2 and LZTR1 are associated with Noonan syndrome. *J Med Genet*. 2015;52:413–21. <https://doi.org/10.1136/jmedgenet-2015-103018>.
76. Piotrowski A, Xie J, Liu YF, Poplawski AB, Gomes AR, Madanecki P, et al. Germ-line loss-of-function mutations in LZTR1 predispose to an inherited disorder of multiple schwannomas. *Nat Genet*. 2014;46:182–7. <https://doi.org/10.1038/ng.2855>.
77. Frattini V, Trifonov V, Chan JM, Castano A, Lia M, Abate F, et al. The integrated landscape of driver genomic alterations in glioblastoma. *Nat Genet*. 2013;45:1141–9. <https://doi.org/10.1038/ng.2734>.
78. Castel P, Cheng A, Cuevas-Navarro A, Everman DB, Papageorge AG, Simanshu DK, et al. RIT1 oncoproteins escape LZTR1-mediated proteolysis. *Science*. 2019;363:1226–30. <https://doi.org/10.1126/science.aav1444>.
79. Simanshu DK, Nissley DV, McCormick F. RAS proteins and their regulators in human disease. *Cell*. 2017;170:17–33. <https://doi.org/10.1016/j.cell.2017.06.009>.
80. Bryant KL, Mancias JD, Kimmelman AC, Der CJ. KRAS: feeding pancreatic cancer proliferation. *Trends Biochem Sci*. 2014;39:91–100. <https://doi.org/10.1016/j.tibs.2013.12.004>.
81. Ardalan B, Ciner A, Baca Y, Hinton A, Darabi S, Kasi A, et al. Distinct molecular and clinical features of specific variants of KRAS codon 12 in pancreatic adenocarcinoma. *Clin Cancer Res*. 2025. <https://doi.org/10.1158/1078-0432.CCR-24-3149>.
82. Kim D, Xue JY, Lito P. Targeting KRAS(G12C). From inhibitory mechanism to modulation of antitumor effects in patients. *Cell*. 2020;183:850–59. <https://doi.org/10.1016/j.cell.2020.09.044>.
83. Singhal A, Li BT, O'Reilly EM. Targeting KRAS in cancer. *Nat Med*. 2024;30:969–83. <https://doi.org/10.1038/s41591-024-02903-0>.
84. Kerk SA, Papagiannakopoulos T, Shah YM, Lyssiotis CA. Metabolic networks in mutant KRAS-driven tumours: tissue specificities and the microenvironment. *Nat Rev Cancer*. 2021;21:510–25. <https://doi.org/10.1038/s41568-021-00375-9>.
85. Linette GP, Bear AS, Carreno BM. Facts and hopes in immunotherapy strategies targeting antigens derived from KRAS mutations. *Clin Cancer Res*. 2024;30:2017–24. <https://doi.org/10.1158/1078-0432.CCR-23-1212>.
86. Ganesan K, Xu C, Wu S, Sui Y, Du B, Zhang J, et al. Ononin inhibits tumor bone metastasis and osteoclastogenesis by targeting Mitogen-Activated protein kinase pathway in breast Cancer. *Res (Wash D C)*. 2024;7:0553. <https://doi.org/10.34133/research.0553>.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.