



## Transcriptomic landscape of transposable elements reveals *LTR7-PLAAT4* as a potential oncogene and therapeutic target in pancreatic adenocarcinoma

Meilong Shi, Chuanqi Teng, Shan Zhang, et al.

*Genome Res.* 2026 36: 275-290 originally published online January 8, 2026  
Access the most recent version at doi:[10.1101/gr.280528.125](https://doi.org/10.1101/gr.280528.125)

---

**References** This article cites 74 articles, 11 of which can be accessed free at:  
<http://genome.cshlp.org/content/36/2/275.full.html#ref-list-1>

**Open Access** Freely available online through the *Genome Research* Open Access option.

**Creative Commons License** This article, published in *Genome Research*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

**Email Alerting Service** Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

---

To subscribe to *Genome Research* go to:  
<https://genome.cshlp.org/subscriptions>

## Research

# Transcriptomic landscape of transposable elements reveals *LTR7-PLAAT4* as a potential oncogene and therapeutic target in pancreatic adenocarcinoma

Meilong Shi,<sup>1,10</sup> Chuanqi Teng,<sup>1,10</sup> Shan Zhang,<sup>2,10</sup> Xiaobo He,<sup>3,10</sup> Lingyun Xu,<sup>4,5</sup> Fengxian Han,<sup>4,5</sup> Rongqi Wen,<sup>5</sup> Ganjun Yu,<sup>3</sup> Jingwen Liu,<sup>6</sup> Yang Feng,<sup>7</sup> Yanfeng Wu,<sup>3,11</sup> Yan Ren,<sup>6,8,11</sup> Gang Jin,<sup>1,11</sup> and Jing Li<sup>4,5,9,11</sup>

<sup>1</sup>Department of Hepatobiliary Pancreatic Surgery, Changhai Hospital, Second Military Medical University (Naval Medical University), Shanghai 200433, China; <sup>2</sup>Center for Translational Medicine, Second Military Medical University (Naval Medical University), Shanghai 200433, China; <sup>3</sup>National Key Laboratory of Immunity and Inflammation, Institute of Immunology, Naval Medical University, Shanghai 200433, China; <sup>4</sup>School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China; <sup>5</sup>Department of Precision Medicine, Changhai Hospital, Second Military Medical University (Naval Medical University), Shanghai 200433, China; <sup>6</sup>Experiment Center for Science and Technology, Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China; <sup>7</sup>College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China; <sup>8</sup>HIM-BGI Omics Center, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Zhejiang, Hangzhou 310022, China; <sup>9</sup>State Key Laboratory for Macromolecule Drugs and Large-scale Manufacturing, School of Pharmaceutical Sciences, Wenzhou Medical University, Wenzhou 325030, China

Eukaryotic genomes contain numerous transposable elements (TEs), whose dysregulation threatens genome stability and may contribute to cancer. Pancreatic adenocarcinoma (PAAD) is among the deadliest cancers, marked by abundant stroma that obscures tumor-specific molecular signals, complicating bulk-tissue analyses. Here, using 71 patient-derived PAAD organoids, we show that TE activities may potentially promote tumorigenesis and provide a source of novel immunotherapeutic targets. We identify 16 new TE-derived transcripts fused with 15 known oncogenes, exhibiting potential oncogenic function and prognostic value. Notably, *LTR7-PLAAT4*, present in 29% of tumors, encodes a protein variant transcriptionally regulated by *FOXM1* binding to the *LTR7* promoter. *LTR7-PLAAT4* isoform 2 is associated with increased cholesterol ester accumulation and lipid droplet formation mediated through *BSCL2* coexpression, potentially fostering tumor progression. On the immunogenic front, HLA-I immunopeptidomics of AsPC-I cells and DACI3 organoids identify over 11,000 peptides respectively. Although mutation-derived neoantigens are rare, several peptides are originated from TE-chimeric transcripts, including four predicted by TEprof2. The peptide FLIQHLPLV, detected in 27% of organoids, exhibits robust immunogenicity, validated by T2 binding, mass spectrometry and ELISPOT assays with HLA-genotyped PBMCs. Together, these findings suggest that TE activities may contribute to PAAD progression and diversify its immunopeptidome, providing new opportunities for molecular subtyping and potential immunotherapeutic intervention.

[Supplemental material is available for this article.]

Pancreatic adenocarcinoma (PAAD) is one of the most lethal malignant neoplasms across the world and a leading cause of cancer mortality, especially in developed countries (Siegel et al. 2024). By 2030, PAAD is expected to become the second most fatal cancer, following only lung cancer (Rahib et al. 2014). Unlike the majority of other cancers, five- and 10-year survivals for pancreatic cancer have not shown much improvement since the early 1970s (Dbouk et al. 2022). A hallmark of PAAD is its abundant desmoplastic stroma, which significantly complicates the isolation and analysis of molecular signals originating specifically from neoplastic epithelial cells, particularly when using bulk tumor sam-

ples. Studies have estimated the median tumor purity of bulk sequencing samples to be <40%, underscoring the limitations of this approach (Biankin et al. 2012; Bailey et al. 2016; Cao et al. 2021). Several studies have addressed this challenge using sophisticated bioinformatics techniques (Elyada et al. 2019), labor-intensive laser capture microdissection (Maurer et al. 2019), FACS cell sorting (Espinet et al. 2021), and more recently, (spatial) single-cell sequencing approaches (Moncada et al. 2020; Ma and Zhou 2022). Recent advances in organoid models represent a transformative development in cancer research. Organoids, cultured in vitro, retain self-renewal and self-organizing capabilities that mimic the structure and function of their tissue of origin. The high consistency between organoid and bulk tumor samples has been demonstrated across multiple cancer types, including glioblastoma (Jacob et al. 2020) and PAAD (Shi et al. 2022). Patient-derived

<sup>10</sup>These authors contributed equally to this work.

<sup>11</sup>These authors jointly supervised this work.

Corresponding authors: wuyf@immunol.org, reny@genomics.cn, jingang@smmu.edu.cn, ljing@smmu.edu.cn

Article published online before print. Article, supplemental material, and publication date are at <https://www.genome.org/cgi/doi/10.1101/gr.280528.125>. Freely available online through the *Genome Research* Open Access option.

© 2026 Shi et al. This article, published in *Genome Research*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

organoids, which achieve nearly 100% tumor purity while preserving intratumoral heterogeneity (Shi et al. 2022), have been used to generate high-quality sequencing data and minimize confounding by nontumor components. These organoid models also enable high-throughput drug screening, providing valuable insights into therapeutic responses for PAAD, a disease notorious for its lack of effective treatments.

Most studies on PAAD have focused on coding regions, investigating gene mutations (Hayashi et al. 2021) and coding gene expression (Moffitt et al. 2015) to elucidate oncogenic mechanisms, identify biomarkers, and guide treatment choices. In contrast, epigenetic alterations remain underexplored. One study stratified pancreatic ductal adenocarcinoma into two distinct subgroups based on specific DNA elements, such as transposable elements (TEs), although it was limited by the use of labor-intensive FACS sorting with small sample sizes (Espinet et al. 2021). TEs exhibit significant regulatory and functional potential. For instance, repetitive region transcripts can form dsRNAs, activating interferon pathways and remodeling the tumor stroma (Espinet et al. 2021). TE-derived sequences may serve as enhancers (Xie et al. 2013), promoters (Brocks et al. 2017; Jang et al. 2019), or other regulatory elements, and constitute a substantial portion of lincRNAs (Khorkova et al. 2015; Choudhary et al. 2023). Despite their potential importance, the transcriptomic landscape and clinical relevance of TE-derived elements in PAAD remain poorly characterized, partly due to limited sample sizes and the constraints of bulk sequencing.

In this study, we aim to systematically characterize the landscape of TE-chimeric transcripts and their clinical relevance in pancreatic cancer using large-scale patient-derived organoid and bulk sample cohorts. Our objective is to clarify the potential contributions of transposable elements to PAAD pathogenesis and immunogenicity and to explore new opportunities for biomarker discovery and immunotherapeutic development.

## Results

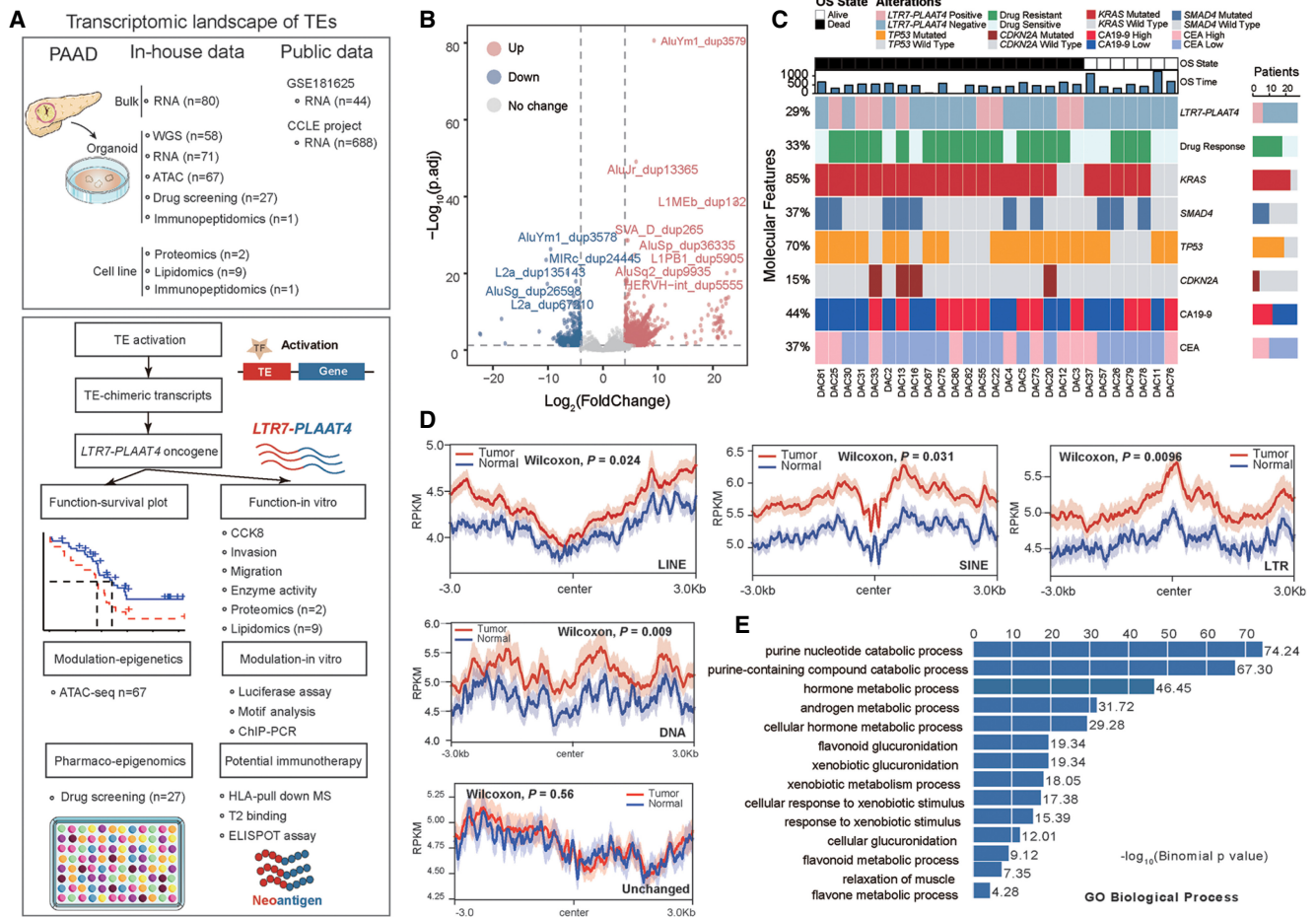
### Integrative analysis of transposable element activation across patient cohorts and cell line data sets

To comprehensively characterize dynamically expressed TEs in PAADs, we analyzed our in-house RNA-seq data sets from patient-derived organoids (Fig. 1A; Supplemental Fig. S1A; Supplemental Table S1; Shi et al. 2022). We conducted analysis of differentially expressed TEs (DETEs) between 66 PAADs and five normal pancreatic ductal organoids using various recommended methods for TE expression quantification (Xiang et al. 2022). We identified 7413 upregulated and 871 downregulated DETE copies in PAADs by featureCounts at a cutoff of fourfold difference and a false discovery rate (FDR) < 0.05, based on multiple previous studies (Fig. 1B; Supplemental Fig. S1A; Supplemental Table S2; Jin et al. 2015; Savytka et al. 2022; Xiang et al. 2022). To validate these findings, we used an independent patient cohort, including 41 bulk tumor and 39 paratumor samples, as well as 41 CCLE (Ghandi et al. 2019) pancreatic cancer cell lines data sets and three hTERT-HPNE pancreatic epithelial normal cell lines (Supplemental Fig. S1A). TE expression dynamics in the cell line data sets were consistent with our bulk tumor cohort (Supplemental Fig. S1B–J). We found that all major classes of TEs exhibited upregulation in PAADs (Supplemental Fig. S2). In the 7413 upregulated DETEs, 36.4% were classified as LINE elements, and 35.8% belonged to SINE, and 0.18% of the 770,551 *LTR* genomic copies

were upregulated. Unlike previous reports linking TE expression to genome instability, we found no association with gene mutations or tumor markers (Fig. 1C; Supplemental Fig. S3). Because epigenetic activation usually leads to TE expression, we interrogated the epigenetic signals based on ATAC-seq data (Shi et al. 2022). We found that upregulated DETEs were associated with more accessible chromatin (Fig. 1D; Supplemental Fig. S4). TEs could act as enhancers and promoters to regulate gene expression. Therefore, we applied the GREAT tool (McLean et al. 2010) to our DETE regions. We found enrichment of genes related to metabolic processes, including the purine nucleotide catabolic process and hormone metabolic process (Fig. 1E).

### TE-chimeric transcripts and onco-exaptation events in the PAAD organoid cohort

To examine the products of TE activation, we utilized TEprof2 (Shah et al. 2023) to analyze TE-derived transcripts, specifically fusion transcripts between TE sequences and known genes, which we referred to as TE-chimeric transcripts. These transcripts displayed different levels in 66 tumor organoids compared to five normal organoids and 39 paratumoral samples. After rigorous filtering, we identified 469 TE-chimeric transcripts from 3513 highly confident TE transcripts as specific to PAADs, which represents a significantly higher count than previously reported (Fig. 2A; Supplemental Fig. S5A; Supplemental Table S3; Shah et al. 2023). Unlike Shah et al. (2023), we could not include GTEx data due to lack of raw RNA-seq files, which are required for splicing-aware alignment. To ensure specificity, we applied a three-step filtering strategy, including tumor enrichment scoring and removal of transcripts overlapping benign promoters from the FANTOM5 database. TE-chimeric transcripts in the organoid data sets were consistent with our bulk tumor cohort and cell line data sets (Supplemental Fig. S5B). A comparative analysis between organoid and bulk samples was conducted to assess the consistency of TE onco-exaptation events, which identified 89 concordant onco-exaptation instances (Supplemental Fig. S5B). Among these, 66% (2305/3513) were novel, presumably due to the low purity of TCGA PAAD samples used in previous studies (Supplemental Fig. S6A). Of the 469 PAAD-specific TE-chimeric transcripts, most originated from intronic (77%) or partially exonic regions (15%) (Fig. 2B). The *LIPA2* subfamily was notably enriched (Fig. 2C), consistent with TCGA pan-cancer analyses. We identified 16 TE-driven fusion events involving 15 known oncogenes, spanning all major TE classes (Fig. 2D; Supplemental Fig. S6B–D; Supplemental Table S4). We assessed associations between TE-chimeric transcripts and overall survival in PAADs (Supplemental Fig. S7; Supplemental Table S4). One of the most abundant fusions, *LTR7-PLAAT4*, was present in 29% of PAAD organoids and correlated with poor prognosis (Figs. 2E, 3A,B; Supplemental Table S5). We found that *LTR7-PLAAT4* chimeric transcripts predicted prognosis and treatment response more accurately than driver mutations or preoperative markers (e.g., CA19-9, CEA) (Fig. 1C). *LTR7-PLAAT4* fusion was also detected in multiple gastrointestinal cancer cell lines, including 24% of pancreatic, 37% of esophageal, and 35% of intestinal lines (Fig. 3C,D). *LTR7* was located ~7 kb upstream of the *PLAAT4* promoter and initiated two transcript isoforms in tumor organoids (Supplemental Fig. S8A–C). To verify the existence of the *LTR7-PLAAT4* isoforms in pancreatic and other cancer cell lines, we conducted nested PCR and Sanger sequencing (Supplemental Fig. S8D–F). Isoform 2 was validated in Capan-1 and SNU-16 cells



**Figure 1.** Activation of transposable elements in PAAD. (A) Overview of the data sets used in this study, including both in-house and publicly available data. Detailed numbers for each data type are provided (top panel). The schematic illustrates the activation of promoters within transposable elements (TEs) in PAAD. Upon activation, these promoters may drive onco-exaptation, resulting in the expression of chimeric transcripts that contribute to oncogenesis, or generate neoantigens that trigger an immune response (bottom panel). (B) Volcano plot showing the differentially expressed TE loci in tumor organoid and normal organoid. The TE-IDs such as “AluYm1\_dup3579” were listed in Supplemental Table S2. Only genes with FDR-adjusted  $P$ -values  $< 0.05$  and absolute  $\log_2$  fold change of  $\geq 4$  are considered significantly differentially expressed. (C) Oncoprint illustrating drug sensitivity, driver gene mutations, tumor biomarker levels, and survival outcomes of patients who underwent organoid-based drug screening. (D) Comparison of ATAC-seq signals located within 3 kb upstream of and downstream from DETEs between organoid tumor and normal samples. Unchanged TE regions (2000 randomly selected) were used as a control. Statistical analysis was performed using the Wilcoxon test, with a  $P$ -value  $< 0.05$  considered statistically significant. PAAD tumor organoid,  $n = 62$ ; Normal organoid,  $n = 5$ . (E) Gene Ontology (GO) enrichment analysis of biological processes associated with DETEs in PAADs.

but absent in hTERT-HPNE normal cells (Fig. 3E; Supplemental Table S6).

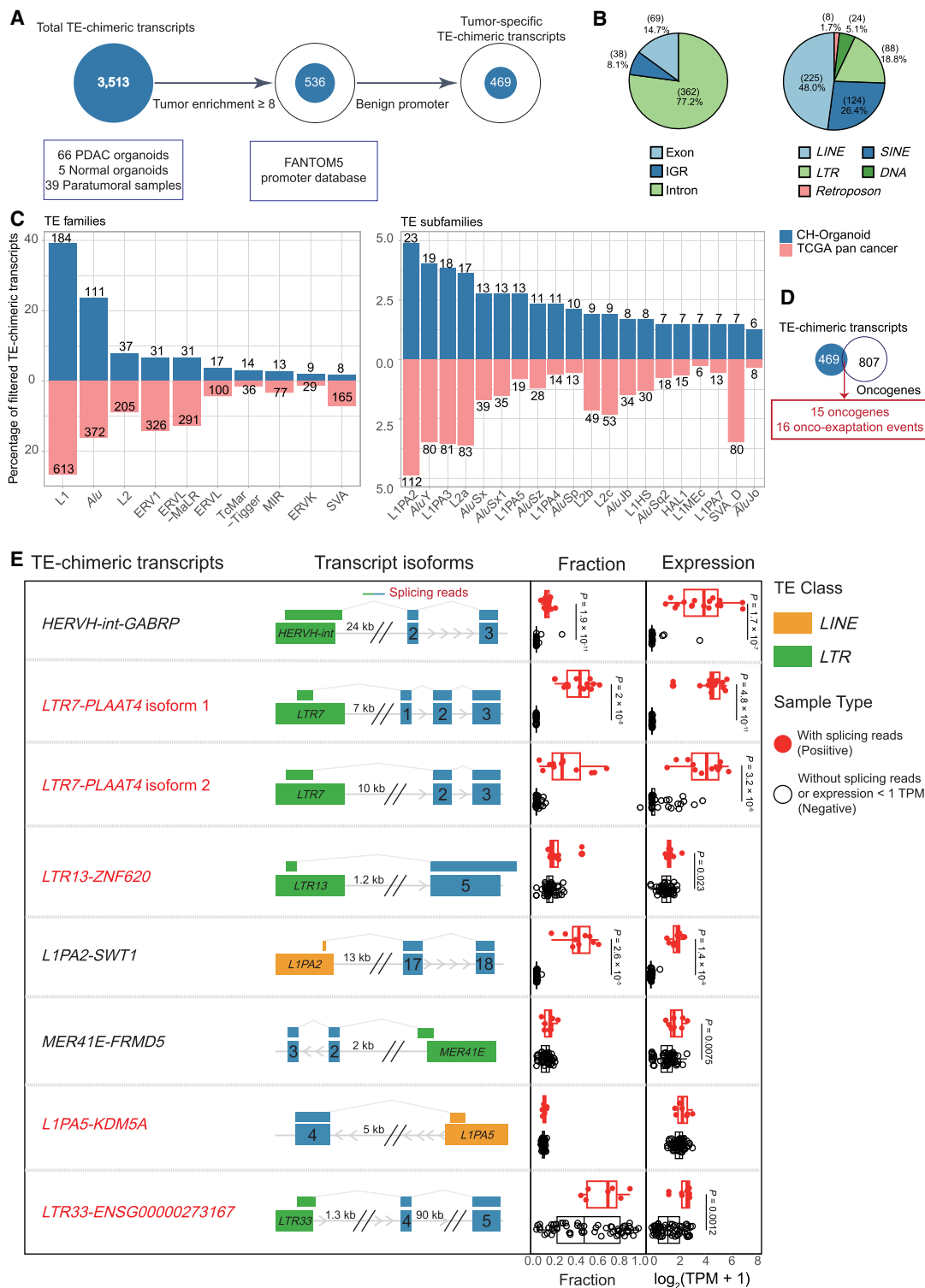
### LTR7 as a potential accessible TF binding site provider driving novel transcription

Next, we profiled chromatin accessibility using ATAC-seq and found that organoids positive for *LTR7-PLAAT4* exhibited higher accessible signals (Fig. 4A). This suggests that an *LTR7* TE is in open chromatin and could serve as an alternative promoter to drive *PLAAT4* expression in pancreatic cancer organoids. Next, we used luciferase assays to test the promoter activities of *LTR7* as well as *MER21B*, a TE fragment immediately adjacent to the 5' end of *LTR7* (Fig. 4B). *MER21B* alone without *LTR7* displayed minimal promoter activity (Fig. 4C), but *MER21B* had a positive *cis*-regulatory effect on *LTR7* transcription. We identified putative transcription factor motifs that were specific to *LTR7* with PROMO (Netanelly et al. 2019). Four TFs associated with these mo-

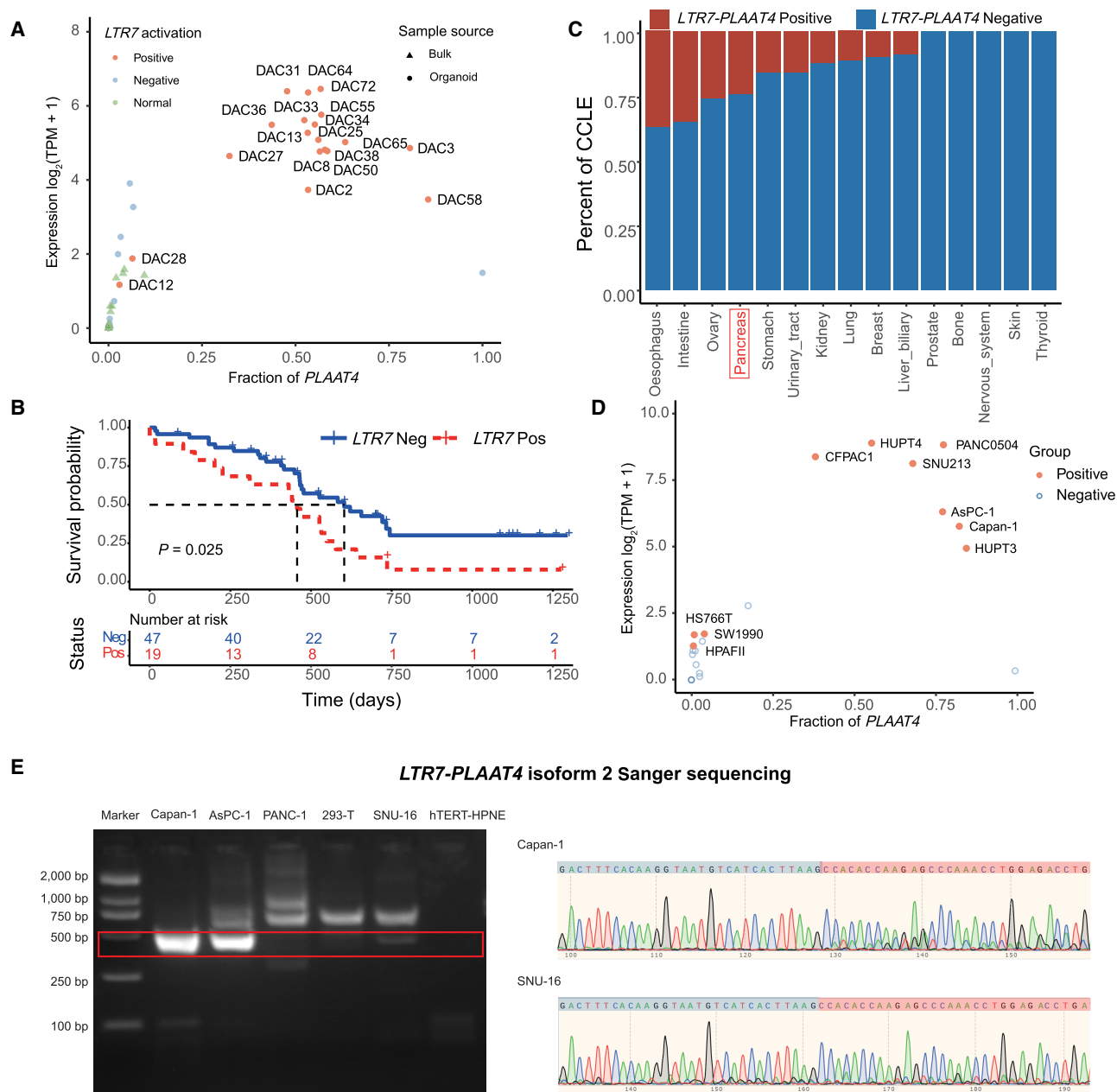
tifs, *FOXM1* (HNF-3alpha), *TFAP2A*, *MAZ*, and *ETV4*, were found to be differentially expressed between tumor and normal samples, with *FOXM1* (HNF-3alpha) being exclusively highly expressed in *LTR7-PLAAT4*-positive samples (Fig. 4D; Supplemental Fig. S9). We validated the binding of *FOXM1* to *LTR7* sequence using ChIP-qPCR (Fig. 4E). We then mutated each binding motif in the *LTR7* sequence and performed luciferase reporter assay and found that, in both 293T and AsPC-1, mutating the *FOXM1* (HNF-3alpha) site significantly reduced luciferase expression (Fig. 4F,G).

### LTR7-PLAAT4 promotes tumorigenesis via lipid remodeling in PAADs

Phospholipase A and acyltransferase 4 (*PLAAT4*) has recently emerged as an oncogene in pancreatic cancer, in contrast to its previously reported tumor suppressor roles in other malignancies (Li et al. 2022). However, the functions of the newly discovered, TE-



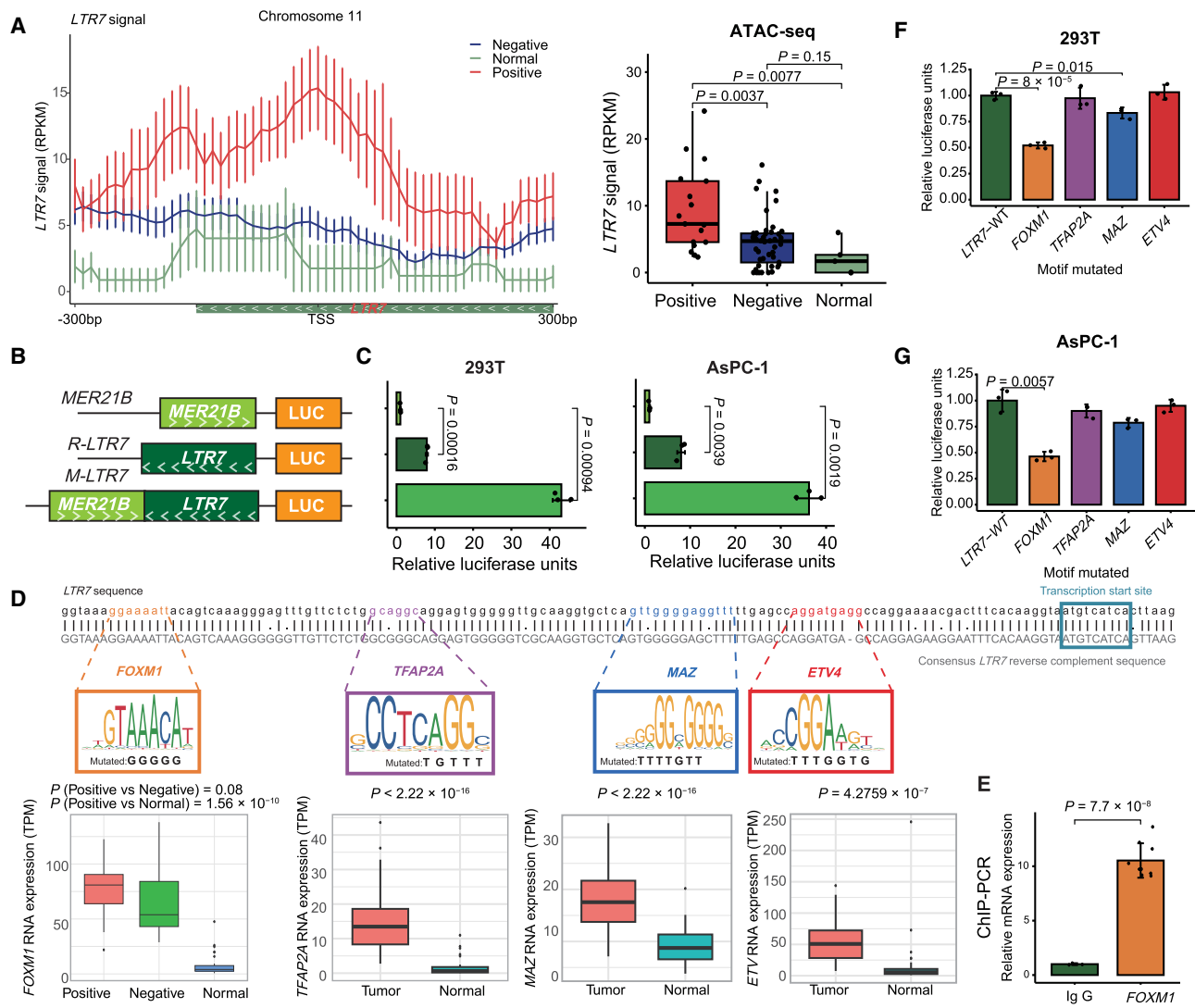
**Figure 2.** Identification of tumor-specific and clinically prognostic TE-chimeric transcripts in PAADs. (A) Diagram illustrating the process of candidate transcript filtering for tumor enrichment using normal organoids and normal bulk samples along with the FANTOM5 promoter atlas. (B) Pie charts depicting the TE location distribution of TE-chimeric transcripts (left) and their distribution across TE classes (right). IGR refers to the intergenic region. (C) Bidirectional bar graphs showing the percentage of families and subfamilies within filtered tumor-associated TE-chimeric transcripts. The upper section represents the in-house organoid cohort, whereas the lower section represents the previously reported TCGA pan-cancer data set. (D) Venn diagram illustrating the intersection between candidate transcripts identified in the organoid cohort and the 807 oncogenes based on the ONGene database and as previously reported. (E) The top eight most prevalent candidates significantly associated with patient survival. The left panel provides labels for the TE-chimeric transcripts and diagrams illustrating the transcript structure of each candidate. In the “Fraction” panel, each bar represents the proportion of total gene expression that is attributable to the TE-chimeric transcript, calculated only across the samples in which the candidate is detected. Validated candidates are denoted by red solid circles, whereas candidates lacking supporting evidence in terms of splicing reads or expression levels below 1 TPM are represented by black hollow circles. The “Expression” panel displays the expression levels of samples featuring the onco-exaptation candidate, with the candidate represented by red solid circles. Samples without supporting splicing reads or expression levels below 1 TPM are denoted by black hollow circles. Each box represents the median and interquartile range, with statistical analysis performed using *t*-tests. TE-chimeric transcripts labeled in red indicate the presence of coding potential.



**Figure 3.** Expression profiling of *LTR7-PLAAT4* across PAAD organoids and various cancer cell lines. (A) Scatterplot illustrating the expression and fraction distribution of *LTR7-PLAAT4* across organoids and bulk samples (n = 110). Orange represents tumor samples positive for *LTR7-PLAAT4*; blue represents tumor samples negative for *LTR7-PLAAT4* (lacking supporting splicing reads or with expression level <1 TPM); green represents normal samples; triangles and circles denote bulk samples and organoid samples, respectively. (B) Comparison of overall survival (OS) between the groups with and without *LTR7-PLAAT4* existence in the organoid cohort (n = 66), assessed using the log-rank test. (C) Frequency distribution bar plot of the *LTR7-PLAAT4* identified in CCLE cancer cell lines data sets, arranged from left to right based on the frequency of detection in cell lines (n = 685). (D) Expression and fraction distribution of *LTR7-PLAAT4* across pancreatic cancer cell lines (n = 41). (E) PCR results following nested PCR were obtained for Capan-1, AsPC-1, and SNU-16 cell lines (which tested positive for *LTR7-PLAAT4* isoform 2), as well as for PANC-1, 293-T, and hTERT-HPNE cell lines (which tested negative for *LTR7-PLAAT4* isoform 2). Representative Sanger sequencing results of Capan-1 and SNU-16 were displayed.

derived *PLAAT4* isoforms are completely unknown. To address this, we individually expressed canonical *PLAAT4* and two LTR7-driven isoforms (isoform 1 and 2) in PANC-1 and AsPC-1 cells. Overexpression of canonical *PLAAT4* and isoform 1 promoted cell viability, whereas isoform 2 notably enhanced migration capacity (Fig. 5A–G; Supplemental Fig. S10). Sequence analysis revealed a robust alternative start codon within the *LTR7*

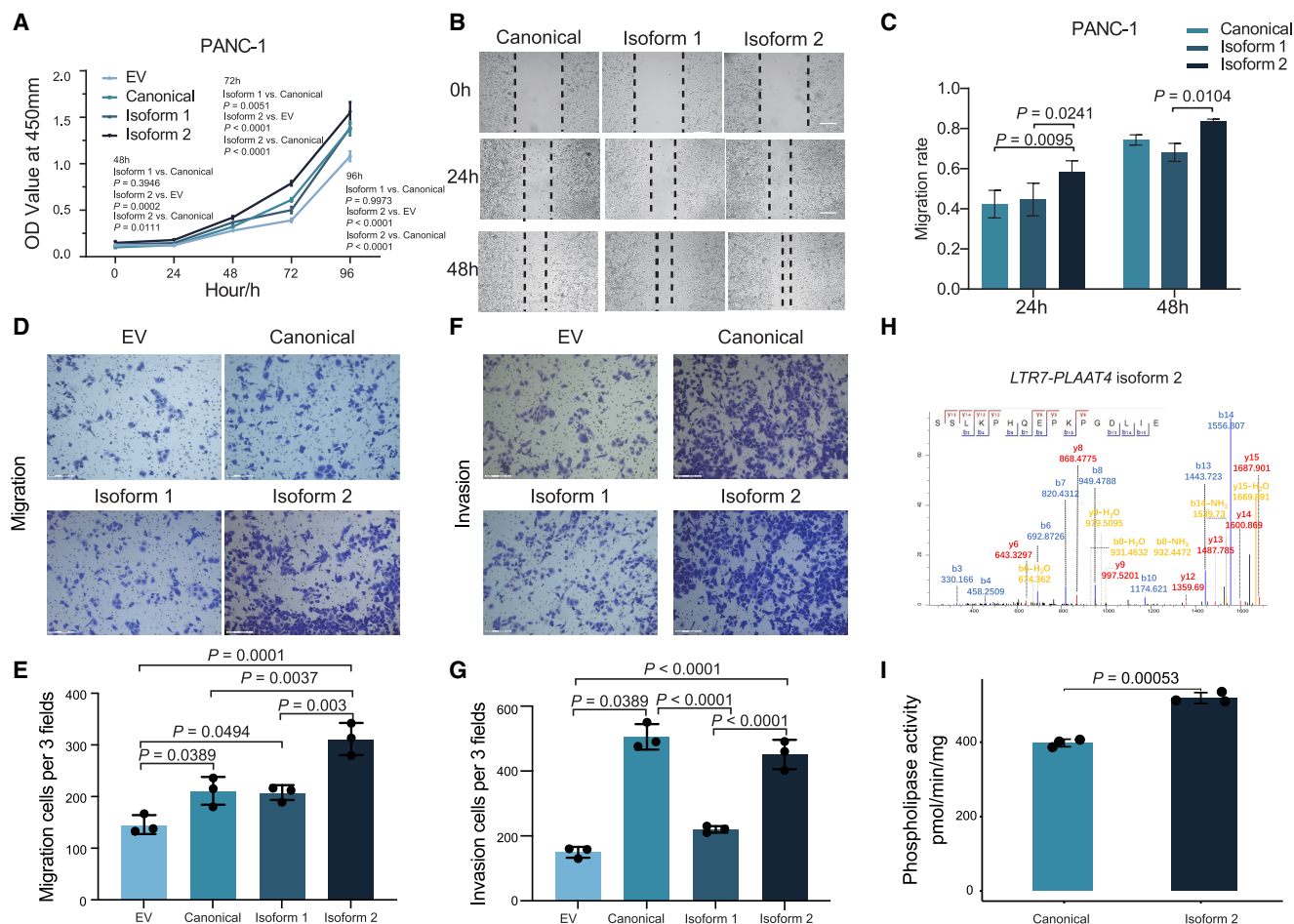
sequence, altering the N-terminus of isoform 2 (Supplemental Fig. S11A), which was validated by mass spectrometry following overexpression (Fig. 5H; Supplemental Fig. S11B; Huang et al. 2024). However, due to the low endogenous abundance of isoform 2 and its high sequence similarity to canonical *PLAAT4*, we were unable to confidently detect isoform 2-specific peptides under physiological conditions. *LTR7-PLAAT4* isoform 2 showed a



**Figure 4.** LTR7 drives *PLAAT4* expression in pancreatic cancer. (A) Line plot displaying the ATAC-seq data signal within  $\pm 300$ -bp range upstream and downstream from the LTR7 transcript start site. The red line represents tumor organoids with LTR7-*PLAAT4* ( $n = 17$ ), the blue line represents tumor organoids without LTR7-*PLAAT4* ( $n = 45$ ), and the green line represents normal organoids ( $n = 5$ ) (left). Statistical analysis was performed using a *t*-test for each group comparison (right). (B, C) Luciferase assays evaluating the transcriptional activity of different TE arrangements in 293T cells (left) and AsPC-1 cells (right) ( $n = 3$  independent experiments), with statistical analysis conducted using a *t*-test. (D) Schematic illustration of luciferase reporter assays measuring LTR7 promoter activity with mutated transcription factor binding motifs. (E) Quantification of FOXM1 enrichment in AsPC-1 cells by ChIP-PCR analysis. (F) Luciferase assays for promoter activity in 293T with mutagenized transcription factor motifs in LTR7 ( $n = 3$  independent experiments). Box plots of RNA expression of transcription factor among different groups, with statistical analysis performed using *t*-tests. (G) Same as F but in AsPC-1 cells. All data are represented as means  $\pm$  standard error (s.e.m.).

significant increase in its phospholipase activity compared to canonical *PLAAT4* (Fig. 5I; Supplemental Fig. S11C). Notably, this N-terminal modification did not affect the membrane localization of the protein (Supplemental Fig. S11D). To elucidate the molecular mechanisms underlying the oncogenic effects of different LTR7-*PLAAT4* transcripts, we performed correlation analysis and identified genes positively correlated with different *PLAAT4* isoforms (Fig. 6A). Our results revealed that genes specifically correlated with canonical *PLAAT4* were enriched in virus defense functions (Supplemental Fig. S12A), whereas those correlated with LTR7-*PLAAT4* isoform 1 were enriched in apoptosis-related functions (Supplemental Fig. S12B). Notably, genes correlated with LTR7-*PLAAT4* isoform 2 were enriched in lipid droplet forma-

tion (Fig. 6B), including lipid droplet-associated enzymes such as *PLA2G4C* and *BSDL2* (Supplemental Fig. S12C). Subsequently, we investigated the impact of constitutive ectopic expression of LTR7-*PLAAT4* isoforms on phospholipid composition using mass spectrometry-based lipidomics (Fig. 6C; Supplemental Table S7). Initially, we observed an increase in phosphatidylcholines (PC) and phosphatidylethanolamine (PE) levels, although the differences were not statistically significant. Consistent with previous reports (Bradley et al. 2022), overexpression of the *PLAAT* family gene showed a slight increase in PC and PE levels. *PLAAT4* facilitates the transfer of fatty acyl groups from PC to LPC and from PE to NAPE (Golczak et al. 2012; Uyama et al. 2012b), and accordingly, we observed increased LPC and LNAPE



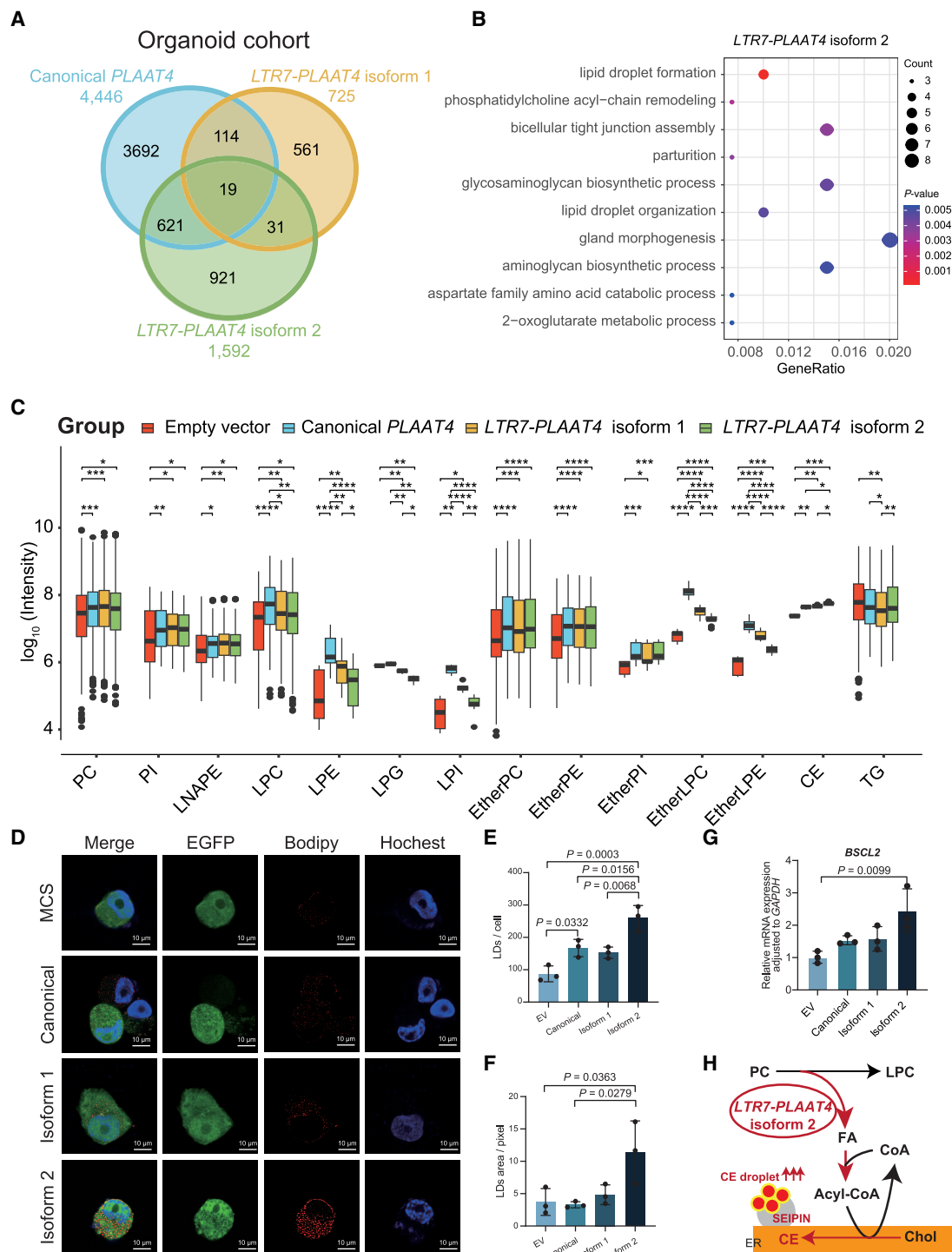
**Figure 5.** Oncogenic properties of LTR7-PLAAT4 isoform 2 in pancreatic cancer cell lines. (A) PANC-1 cell proliferation rate comparison using a CCK-8 assay after LTR7-PLAAT4 isoform 2 overexpression. (B,C) PANC-1 cell migration rate comparison using wound healing assay after different PLAAT4 isoforms overexpression. Representative images of the assay were shown at 50 $\times$  magnification. Scale bar, 200  $\mu$ m. (D,E) PANC-1 cell migration rate comparison using migration assay after different PLAAT4 isoforms overexpression. Scale bar, 150  $\mu$ m. (F,G) Invasion assay was used to detect invasion capability after LTR7-PLAAT4 isoform 2 overexpression in PANC-1 cell lines. Scale bar, 150  $\mu$ m. (H) Secondary mass spectra were obtained from 293T cells overexpressing LTR7-PLAAT4 isoform 2. (I) Phospholipase activity assay of canonical and isoform 2 LTR7-PLAAT4 proteins.

levels in all PLAAT4-overexpressing cell lines, consistent with prior studies (Uyama et al. 2012a) and our phospholipase assay. However, no significant differences in LPC and LNAPE levels were observed among the cell lines overexpressing different PLAAT4 isoforms. The most notable differences among cell lines were in cholesteryl esters (CEs), with the cell line overexpressing PLAAT4 isoform 2 showing the highest CE levels compared to those overexpressing the other two isoforms. Meanwhile, there was a trend toward decreased fatty acid (FA) levels. As a member of the PLA1/PLA2 family, PLAAT4 releases FA from phospholipids (e.g., PC), providing FA for the esterification of cholesterol to form CEs, which are stored in lipid droplets (Tu et al. 2023). There are mainly two types of lipid droplets: CEs and triglycerides (TGs) (Fujimoto and Parton 2011). However, in these cells, PLAAT4-mediated lipid droplet formation primarily involved CE lipid droplets because TG levels decreased in the cell line overexpressing PLAAT4 isoform 2. To further elucidate the relationship between PLAAT4 isoform 2 and cholesterol ester-rich lipid droplets, we conducted in vitro induction of CE-rich lipid droplets. The results demonstrated that both the number and area of lipid droplets in cells expressing PLAAT4 isoform 2 were significantly higher compared to

control cells (Fig. 6D–F). The expression of PLAAT4 isoform 2 was positively correlated with *BSCL2*, which was significantly upregulated in the isoform 2-expressing cells compared to control cells (Fig. 6G). Seipin, derived from the *BSCL2* gene, is a crucial protein involved in forming the lipid droplet assembly complex, including CE droplets (Dumesnil et al. 2023). Lipid droplet formation is crucial for tumors, as it alleviates lipid toxicity, mitigates ER stress, provides energy, and supports cell survival (Zadoorian et al. 2023). These findings suggest that TEs contribute to the diversity of the PLAAT4 gene products and function, thereby contributing to cancer metabolism remodeling (Fig. 6H).

#### Tumors positive for the LTR7-PLAAT4 oncogene exhibit multidrug resistance

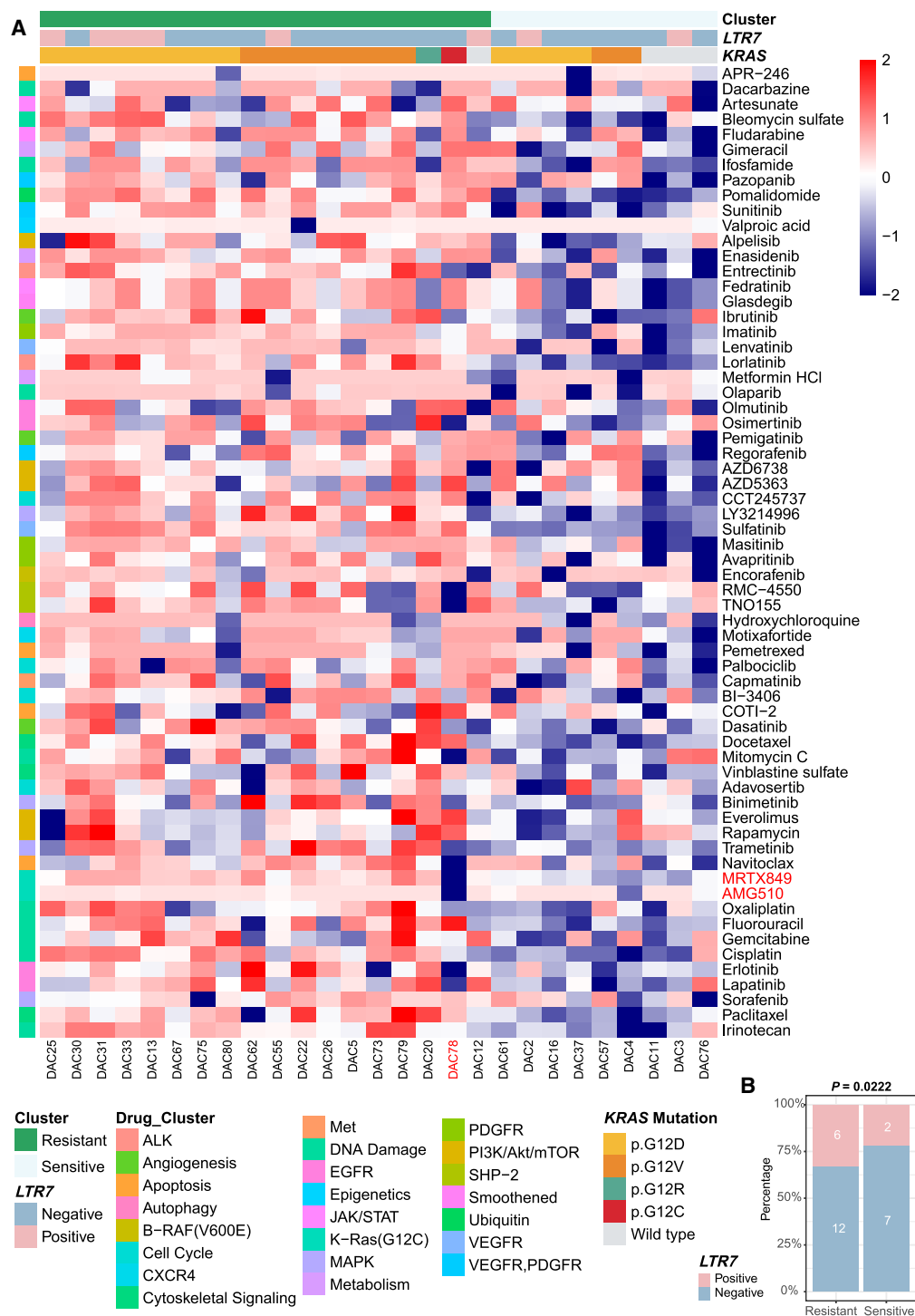
To identify potential therapeutic strategies in light of LTR7-PLAAT4, we conducted a high-throughput drug screen using our organoid platform and evaluated the correlation between drug sensitivity and LTR7-PLAAT4 status in the organoids. A bespoke 64 compound library was assembled for screening, including chemotherapy (n = 14) and targeted drugs (n = 23) (Supplemental



**Figure 6.** Potential oncogenic mechanism of *LTR7-PLAAT4* isoform 2 in pancreatic cancer. (A) Venn diagram showing the number of genes positively correlated with different *PLAAT4* isoforms in organoid cohort. The canonical *PLAAT4*, *LTR7-PLAAT4* isoform 1, and *LTR7-PLAAT4* isoform 2 are marked in light blue, yellow, and green, respectively. (B) Gene Ontology analysis of specific genes correlated with *LTR7-PLAAT4* isoform 2. The size of each bubble represents the number of enriched genes, and the color indicates the *P*-value, with darker colors representing smaller *P*-values. (C) Lipidomics analysis of PANC-1 cells overexpressing *LTR7-PLAAT4* isoform 1, *LTR7-PLAAT4* isoform 2, and canonical *PLAAT4*. Box plots display the distribution of different *PLAAT4* isoforms across various lipid classes. Each box plot includes the median, interquartile range, and outliers for the lipid classes measured. Statistical significance is indicated with (\*)  $P \leq 0.05$ , (\*\*)  $P \leq 0.01$ , (\*\*\*)  $P \leq 0.001$ , (\*\*\*\*)  $P \leq 0.0001$ . The exact *P*-values of panel C are listed in Supplemental Table S6. (D) PANC-1 cells imaged after 24 h of methyl- $\beta$ -cyclodextrin complexed cholesterol (200  $\mu$ M) feeding. BODIPY was added upon imaging for LD labeling and analysis. (E) Analysis of D. Number of LDs per cell. (F) Analysis of D. Area of LDs per pixel. The experiment was independently repeated three times with similar results. (G) Expression of *BCL2* in different *PLAAT4* isoforms overexpressed pancreatic cancer cell lines (PANC-1) measured by qPCR. (H) Schematic illustrating the special oncogenic role of *LTR7-PLAAT4* isoform 2 in PAAD.

Table S8). As positive controls, one organoid (DAC78) bearing the *KRAS* p.G12C mutation exhibited sensitivity to AMG510 and MRTX849 (Fig. 7A). Clustering of compounds based on their AUC revealed a diverse range of sensitivities across the organoids, identifying two distinct clusters. Tumors positive for the *LTR7*-

*PLAAT4* oncogene were predominantly enriched in the resistant subgroup, exhibiting a broad spectrum of drug resistance (Fig. 7B). This was further supported by Genomics of Drug Sensitivity in Cancer (GDSC) data, where *LTR7-PLAAT4*-positive CCLE cell lines also showed pharmacologic resistance (Supplemental Fig.



**Figure 7.** Drug sensitivity profiles of PAAD organoids. (A) Drug response of 35 PAAD organoids to 64 different drugs. Colors ranging from red to blue represent Z-scores based on the Area Under the Curve (AUC). A low Z-score (blue) indicates high drug sensitivity, whereas a high Z-score (red) indicates high resistance. The positive control is highlighted in red font for clarity. (B) Stacked bar chart displaying the distribution of *LTR7* in different states across two groups with  $\chi^2$  test (Bonferroni correction applied).

S13; The Cancer Cell Line Encyclopedia Consortium, The Genomics of Drug Sensitivity in Cancer Consortium 2015).

### TE-chimeric transcript–derived peptides are validated as putative neoantigens

To overcome the resistance associated with *LTR7-PLAAT4* and target lethal cancers driven by aberrant epigenetic activation of TEs, we turned to recent investigations that revealed the capacity of TEs in harboring small open reading frames (smORFs) that may serve as tumor-specific antigens (TSAs) (Yadav et al. 2014; Newey et al. 2019; Shah et al. 2023). In AsPC-1 pancreatic cancer cells, HLA-bound immunopeptidomics identified 12,234 peptides by mass spectrometry (MS), of which 7481 were predicted to bind the endogenous HLA alleles using NetMHCpan-4.1. Annotation against multiple reference databases revealed these peptides originated from benign reference immunopeptidome databases (n = 3701) (Hoenisch Gravel et al. 2023), UniProt-only entries (n = 774), a single mutation-derived ORF from *CAPN2* (n = 1), locally translated TE ORFs (n = 4), TEprof2-predicted ORFs (n = 4), and unannotated genomic regions (n = 2998) (Fig. 8A; Supplemental Table S9). A similar analysis in DAC13 patient-derived pancreatic tumor organoids identified 11,380 HLA-bound peptides. Among them, the TEprof2-derived peptide FLIQHLPLV was predicted to bind HLA-A\*02:01 (Fig. 8B). This peptide demonstrated strong binding in T2 binding assays (Fig. 8C) and was validated through spectral matching with its synthetic counterpart (Fig. 8D). This peptide was also identified in two additional organoid lines, DAC30 and DAC50, yielding a detection frequency of ~27% (3/11). Two other TE-derived peptides, ETVTGVSWSY and FLPDHWAVL, were similarly validated in AsPC-1 cells (Fig. 8D). To investigate their genomic and translational origins, we visualized the corresponding chimeric transcripts using the WashU Epigenome Browser. RNA-seq coverage and getorf-predicted open reading frames provided evidence of local translation from TE loci (Fig. 8E). Only one mutation-derived HLA-presented peptide (*CAPN2*) was identified in AsPC-1, consistent with its rare mutation status as cataloged in the TRON Cell Line Portal (TCLP). No mutation-derived neoantigens were found in DAC13, underscoring the low mutational burden of PAAD and the limited repertoire of conventional mutation-derived TSAs. These observations prompted us to investigate whether TEs might serve as an alternative antigen source. To assess immunogenicity, two TEprof2-derived peptides—FLPDHWAVL and FLIQHLPLV—were tested via ELISPOT assays using PBMCs from HLA-genotyped healthy donors. Notably, FLIQHLPLV elicited a strong IFNG response, indicating potent T cell activation (Fig. 8F). This same peptide was independently identified as a naturally presented HLA ligand in a recent study on pancreatic cancer (Ely et al. 2025), further substantiating its potential as a bona fide neoantigen candidate.

### Discussion

Our study presents the first comprehensive analysis of TE-derived transcripts using a large panel of tumor organoids, providing insights into their potential oncogenic roles and immunotherapeutic relevance. Organoid models offer high tumor cell purity and retain many features of the original tumor, making them suitable platforms for studying gene function and drug response. However, limitations include loss of some tumor heterogeneity and absence of components such as immune cells and vasculature.

*LTR7-PLAAT4* emerged as a top TE-oncogene candidate in PAAD and other cancers. *LTR7*, regulated by *HNF4A*, may interact with m<sup>6</sup>A-modified *HERV-H* transcripts and *TET1* to evade epigenetic silencing (Sun et al. 2023). Functionally, *PLAAT4* participates in lipid metabolism, and our data suggest isoform 2 overexpression promotes cholesteryl ester accumulation—an established metabolic signature in pancreatic cancer (Li et al. 2016). These findings point to a potential role of TE-driven isoform expression in lipid droplet remodeling.

Whereas our overexpression studies in *PLAAT4*-null cells highlight isoform-specific effects, further validation is needed. Isoform-specific loss-of-function studies are technically challenging due to repetitive sequences and shared promoters. Tools like long-read isoform-specific RNA-seq, CRISPRa screening, or bespoke antibodies will be required to test these hypotheses.

We also identified TE-derived peptides that function as neoantigens. In PAAD models with low mutational burden, these TE antigens could represent alternative immunotherapeutic targets. One peptide, FLIQHLPLV, showed strong MHC binding and T cell activation (Chiappinelli et al. 2015; Roulois et al. 2015; Attermann et al. 2018) and was independently validated in a recent report (Shah et al. 2023).

Additional limitations of our study include lack of in vivo validation, incomplete antigen universality due to MHC diversity, and absence of TCR sequencing. Our RNA-seq approach also excluded non-polyadenylated TE transcripts, underscoring the need for deeper profiling.

In summary, our study provides evidence that TEs can function as both drivers and immunological targets in pancreatic cancer. These findings lay the groundwork for future TE-targeted strategies in oncology.

## Methods

### Study design

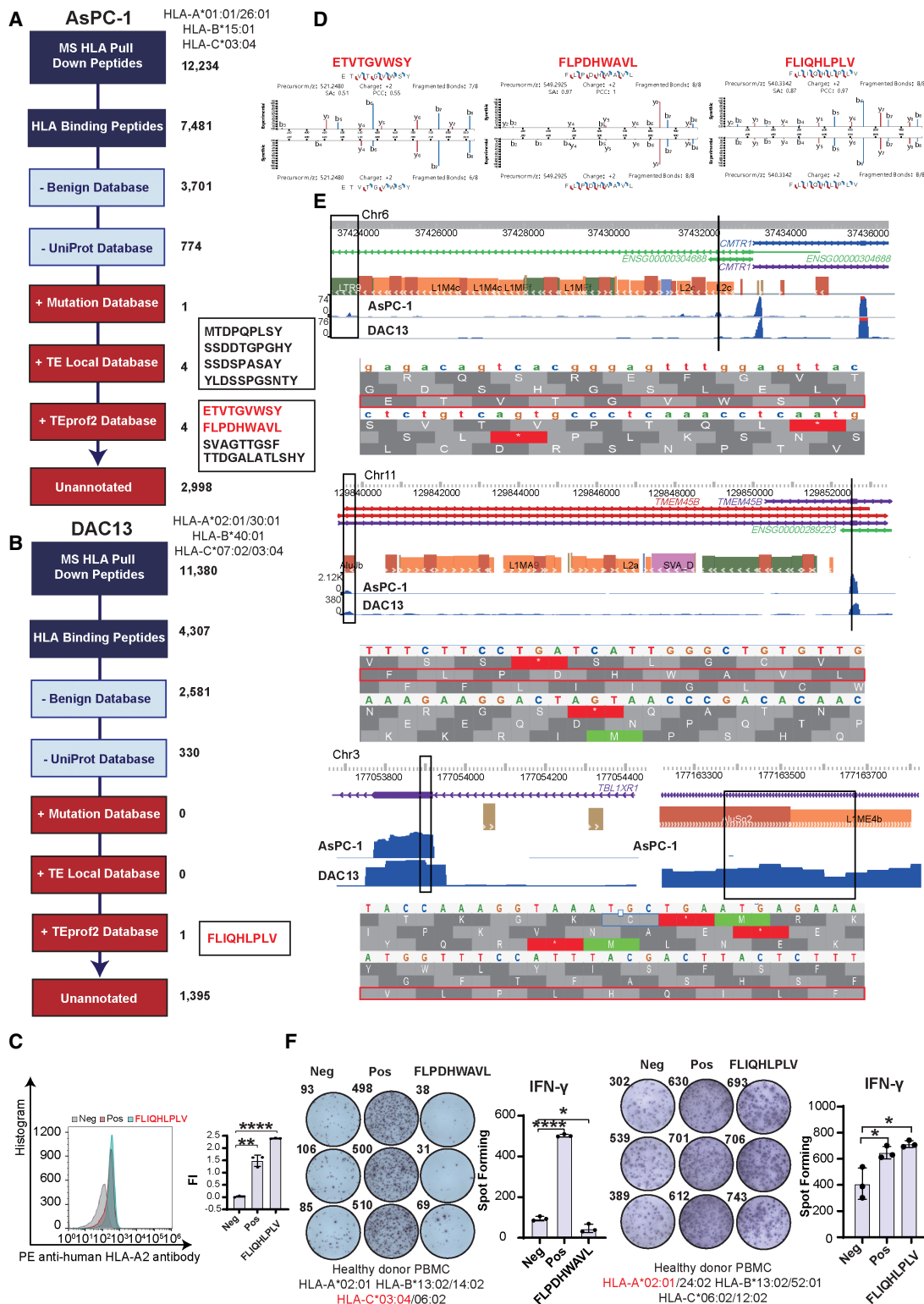
The in-house data set comprised two cohorts: the organoid discovery cohort (Shi et al. 2022) and a new bulk validation cohort. The organoid discovery cohort included 71 patient-derived organoids (PDOs) from pancreatic ductal adenocarcinoma, with 66 samples obtained from tumor tissues and five from paratumoral samples. Whole-genome sequencing (WGS) was performed on tumor and paired blood/normal samples (n = 29), and RNA sequencing was conducted on all 71 organoids. Additionally, 62 tumor organoids and five normal organoids underwent ATAC-seq experiment. Public RNA-seq data from 685 cancer cell lines (CCLE), including 41 PDAC cell lines, and hTERT-HPNE cell lines were retrieved from public repositories (see Supplemental Material).

### RNA-seq and DNA-seq library preparation and sequencing

RNA-seq libraries were generated using the NEBNext Ultra RNA Library Prep kit for Illumina, followed by sequencing on the Illumina HiSeq X TEN platform (150-bp paired-end reads). DNA libraries were prepared from tumor and matched normal samples and sequenced using WGS protocols. Standard protocols for library preparation, sequencing, and QC are provided in the Supplemental Material.

### ATAC-seq library preparation and processing

ATAC-seq was performed on 67 PDOs (62 tumor, five normal) following a modified protocol to remove adapters and transposase sequences before alignment to GRCh38 using BWA-MEM (Jung et al.



**Figure 8.** TE-chimeric transcript-derived peptides as neoantigens. (A) Annotation of 12,234 HLA-bound peptides identified by mass spectrometry (MS) from AsPC-1 cells across multiple reference databases, including benign, UniProt, mutation, locally translated TE ORFs, TEprof2, and unannotated regions. (B) Annotation of 11,380 HLA-bound peptides identified from DAC13 organoids, with one TEprof2-derived peptide (FLIQHLPLV) retained for further analysis. (C) Validation of the binding affinity of the TEprof2-derived peptide FLIQHLPLV to HLA-A\*02:01 using T2 binding assay. (D) Spectral comparison between MS-identified endogenous peptides and synthetic reference peptides: ETVTGVWSY and FLPDHWAVL (AsPC-1), and FLIQHLPLV (DAC13), confirming accurate peptide identification. (E) Visualization of TE-chimeric transcripts and translation evidence using the WashU Epigenome Browser. TE loci are shown with RNA-seq read coverage and getorf-based predicted ORFs; corresponding peptide-coding regions are highlighted. (F) Immunogenicity validation of TE-derived peptides (FLPDHWAVL and FLIQHLPLV) in ELISPOT assays using HLA-genotyped healthy donor PBMCs. Significant IFN $\gamma$  responses indicate these peptides elicit T cell activation.

2022). Peak calling was performed using MACS2 (Zhang et al. 2008). Full details are available in the [Supplemental Material](#).

### RNA-seq and DNA-seq data processing

Raw FASTQ files were processed using fastp v0.23.2 (Chen et al. 2018), aligned to GRCh38 using STAR v2.7.7a (Dobin et al. 2013) for RNA-seq and BWA v.0.7.17 (Li and Durbin 2009) for WGS data. Somatic variants were called using GATK v.4.1.9 (McKenna et al. 2010) MuTect2 with panel-of-normal filtering and annotated using ANNOVAR (Wang et al. 2010). Standard processing pipelines and parameters are outlined in the [Supplemental Material](#).

### TE quantification and differential analysis

TE expression was quantified using featureCounts v2.0.6 (Liao et al. 2014) and TElocal v1.1.1 (Jin et al. 2015) with RepeatMasker annotations. Differentially expressed TEs were identified using DESeq2 (RRID:SCR\_000154) v1.40.2 (Love et al. 2014), with a cutoff of fold-change  $\geq 4$  and FDR  $< 0.05$ .

### Identification and quantification of TE-chimeric transcripts

TE-chimeric transcripts were identified using TEProf2 (Shah et al. 2023), which integrates StringTie v1.3.3 (Pertea et al. 2015) for de novo assembly and transcript quantification. Transcripts overlapping TEs and splicing into downstream genes were filtered using expression (TPM  $> 1$ ) and splice-junction read support. Tumor-specific fusions were prioritized by  $\geq$  eightfold enrichment in tumor versus normal samples. Expression specificity was further refined by filtering against FANTOM5 (Lizio et al. 2015) CAGE peaks, excluding overlapping peaks on the same strand. Fusions were annotated with the ONGene database (Liu et al. 2017) and known onco-exaptation events.

### Possible open reading frame predictions

To evaluate coding potential, we applied CPC2 (Kang et al. 2017) and searched for the longest ORFs starting with Kozak-compatible ATG codons (Kozak 2005). Additional predictions were performed using EMBOSS getorf (Rice et al. 2000) (available at [ftp://emboss.open-bio.org/pub/EMBOSS/](http://ftp://emboss.open-bio.org/pub/EMBOSS/)), retaining sequences with ATG and  $\geq 8$  amino acids. Resulting ORFs were categorized into original, truncated, chimeric-original, or chimeric-truncated. These were forwarded for immunogenicity prediction (see [Supplemental Material](#)).

### Functional association of genes with DETEs

GREAT (McLean et al. 2010) (<http://great.stanford.edu/>) was used to map DETEs to nearby genes (5 kb upstream of to 1 kb downstream from the TSS, with distal extension up to 1000 kb). Enriched GO terms were extracted from GREAT output.

### Visualization of TE-chimeric transcripts

bigWig files were generated from BAM files using deepTools v3.5.6 (Ramírez et al. 2014) and visualized on the WashU Epigenome Browser (Zhuo et al. 2023). Public ChIP-seq data were sourced from CistromeDB (Zheng et al. 2019) (<http://cistrome.org/db/>).

### Sanger sequencing

Nested PCR gene-specific primers (GSPs) were designed to target the fusion sequence of the predicted *LTR7-PLAAT4* isoform 2 and were provided in [Supplemental Table S10](#). Amplified products from two positive (Capan-1, SNU-16) and one negative (hTERT-

HPNE) cell line were sequenced using ABI 3730XL. Sequences were aligned using SnapGene v6.0.2 to confirm fusion junctions.

### Survival analysis

Overall survival was analyzed using the Kaplan-Meier method and log-rank tests. Follow-up data were obtained from institutional records and LinkDoc Technology Co., Ltd. Detailed definitions and statistical tools are listed in the [Supplemental Material](#).

### Cell line and cell culture

AsPC-1, Capan-1, PANC-1, hTERT-HPNE, 293T, and SNU-16 cell lines (RRIDs listed in [Supplemental Material](#)) were obtained from authenticated sources and tested negative for mycoplasma. Cell lines were cultured in medium conditions optimized per cell type, with fetal bovine serum and penicillin/streptomycin supplementation.

### Construction of *LTR7-PLAAT4* overexpressing cell lines

The predicted sequences of *LTR7-PLAAT4* isoform 1 and *LTR7-PLAAT4* isoform 2 were amplified and cloned into pSenti-EF1-EGFP-P2A-Puro-CMV-3xFLAG-WPRE plasmids. Lentiviral vectors (OBio Technology) were then utilized to generate various recombinant vectors for overexpressing *LTR7-PLAAT4* isoform 1, *LTR7-PLAAT4* isoform 2, Canonical *PLAAT4*, and an empty plasmid serving as a negative control (EV). Stable clones were selected with puromycin.

### RT-qPCR

Gene expression was quantified using SYBR Green-based RT-qPCR on an Applied Biosystems QuantStudio 7 Flex system. Amplification specificity was confirmed by melting curve analysis. Data were normalized using the  $\Delta\Delta C_t$  method with three biological and three technical replicates per group. Primer sequences are listed in [Supplemental Table S10](#).

### Western blot

Cells were lysed by sonication in buffer containing protease inhibitors. Proteins (20  $\mu$ g) were resolved by SDS-polyacrylamide gel, specifically a Bis-Tris gel (GenScript, #M00930), transferred to a nitrocellulose membrane (Millipore), and probed with anti-FLAG (Sigma-Aldrich, #F1804, RRID:AB\_262044, 1:1000) and anti- $\beta$ -actin antibodies (66009-1-Ig). Detection was performed with HRP-conjugated secondary antibodies (goat anti-mouse IgG [H+L]) (Thermo Fisher Scientific, #31430, RRID: AB\_228307, 1:10,000) and quantified using ImageJ.

### ChIP-PCR

ChIP-PCR was performed using the CUT&RUN approach with the Hyperactive pG-MNase CUT&RUN kit (Vazyme, #HD101-01) in AsPC-1 cells. Briefly,  $1 \times 10^5$  cells were bound to ConA beads, incubated overnight with FOXM1 antibody (Proteintech, #13147-1-AP, RRID:AB\_2106213) or IgG control (Proteintech, #30000-0-AP, RRID: AB\_2819035), and digested with MNase. DNA was extracted and analyzed by RT-qPCR following the manufacturer's protocol.

### Immunofluorescence staining

PANC-1 cells were seeded on glass coverslips (Genescom, #WHB-12-CS-LC) and fixed with 4% paraformaldehyde. Cells were incubated with 5  $\mu$ g/mL WGA conjugates (Thermo Fisher Scientific, #W32464; diluted in HBSS, #14175095), permeabilized using

0.2% Triton X-100 (Sangon Biotech, #9002-93-1), and blocked with 2.5% Normal Goat Serum (Thermo Fisher Scientific, #R37624). After overnight incubation at 4°C with anti-FLAG antibody (Sigma-Aldrich, #F1804, 1:500), samples were stained with Cy5-conjugated secondary antibody (Abcam, #ab6563, 1:1000) and Hoechst 33342 (Biosharp, #BL803A). Images were acquired using a Zeiss LSM 900 confocal microscope under consistent settings.

### Enzyme activity

A liposome substrate mixture was prepared by combining DOPC (30  $\mu$ L, 10 mM), DOPG (30  $\mu$ L, 10 mM), and Red/Green BODIPY PC-A2 (30  $\mu$ L, 1.5 mM, Thermo Fisher Scientific, #E10217) in DMSO. The mixture (10  $\mu$ L) was added dropwise into 1 mL reaction buffer under vortexing. Protein solutions (0.3 mg/mL, 200  $\mu$ L) were diluted 1:1 with reaction buffer. PLA2 activity was assessed using a standard curve (0–0.5 U/mL) and measured in 96-well microplates (Thermo Fisher Scientific, #M33089). Each well received 50  $\mu$ L enzyme or standard solution and 50  $\mu$ L liposome substrate. Fluorescence was recorded every 2 min for 30 min using a SYNERGY H1 plate reader (Ex: 460 nm; Em: 515 nm).

### LC-MS/MS analysis

For proteomic analysis, 293T cells overexpressing LTR7-PLAAT4 isoform 2 or canonical PLAAT4 ( $2\text{--}5 \times 10^7$  cells) were lysed in ice-cold Cell Lysis Reagent (Sigma-Aldrich, #C2978) supplemented with protease inhibitor cocktail (Bimake, #B14002). Lysates were sonicated every 10 min on ice and centrifuged at 21,000g for 1 h at 4°C. FLAG-tagged proteins were immunoprecipitated using magnetic beads coupled with anti-FLAG antibody (Sigma-Aldrich, #F1804, RRID:AB\_262044).

Immunoprecipitated proteins (100  $\mu$ g) were digested with either Endoproteinase Glu-C or  $\alpha$ -Chymotrypsin (Sigma) at a 1:30 enzyme-to-substrate ratio (w/w) for 12 h at 37°C or 25°C, respectively, in 100 mM sodium bicarbonate (pH 8.0). Peptides were desalted using MonoSpin C18 columns (GL Sciences) and dried using a Speed Vacuum Concentrator (Labogene Aps).

Desalted peptides were loaded onto a PepMap Neo C18 trap column (300  $\mu$ m i.d.  $\times$  5 mm, 5  $\mu$ m, 100 Å; Thermo Fisher Scientific) and separated on a house-packed C18 analytical column (75  $\mu$ m i.d.  $\times$  25 cm, 1.9  $\mu$ m, 100 Å) using the NanoElute system (Bruker Daltonics) at 50°C. Peptides were eluted over a 60-min gradient with buffer B (0.1% formic acid in acetonitrile) as follows: 2%–22% (0–45 min), 22%–35% (45–50 min), 35%–80% (50–55 min), and held at 80% (55–60 min). Buffer A consisted of 0.1% formic acid in water.

### Lipid manipulations and image analysis

PANC-1 cells were delipidated in serum-free medium containing 5% lipoprotein-deficient serum (LPDS) for 2.5 days. For enhanced lipid depletion, DGAT1 and DGAT2 inhibitors (5  $\mu$ M each) were added during the final 18 h. To induce cholesterol ester accumulation, cells were loaded with methyl- $\beta$ -cyclodextrin-complexed cholesterol (20:1 molar ratio) in 5% LPDS for 24 h.

Cells were fixed with 4% paraformaldehyde in 250 mM Hepes buffer (pH 7.4) containing 100  $\mu$ M CaCl<sub>2</sub> and 100  $\mu$ M MgCl<sub>2</sub> and stained with BODIPY 576/589 (10  $\mu$ M; MCE, #HY-D1118) and Hoechst 33342 (5  $\mu$ g/mL; MCE, #HY-15559) for lipid droplet and nuclear visualization. Confocal images were acquired using a Zeiss LSM 900 microscope under consistent settings.

Image analysis was performed in ImageJ (v1.54f). TIFF images were converted to 8-bit grayscale, denoised using a Gaussian blur (sigma=2.0), and thresholded (Otsu or manual). Particles (10–5000 pixels, circularity 0.5–1.0) were detected using the Analyze

Particles function. Edge particles were excluded, and extracted parameters (area, perimeter, circularity) were analyzed in R.

### Immunopeptidomics

Peptides were analyzed on a timsTOF Pro2 system (Bruker Daltonik) in DDA mode. Each acquisition cycle included one MS1 TIMS-MS and 10 PASEF MS/MS scans. Ion accumulation and ramp time were set to 100 msec; ion mobility range was 1/ $K_0$ =0.75–1.30 Vs·cm<sup>-2</sup>. MS/MS precursor ions were measured in the m/z range 100–1700 using CID fragmentation. Dynamic exclusion was set to 30 sec.

### Raw data analysis

MaxQuant (v2.0.1.3; RRID:SCR\_014485) (Tyanova et al. 2016) was used for data analysis with default settings unless otherwise noted. Chymotrypsin or Glu-C was set as the digestion enzyme, with specificity enabled. The minimum peptide length was seven amino acids, allowing up to two missed cleavages. The false discovery rate was controlled using a target-decoy approach at <1% for both peptide spectrum matches and protein groups.

### Promoter and mutagenesis luciferase assay

Based on the results from RNA-seq, ATAC-seq, and ChIP-seq, various promoter sequences derived from TEs were amplified. The minimal promoter sequence of the pGL3-basic luciferase plasmid was removed with KpnI and XhoI restriction enzyme digestion, and the TE-derived promoters were cloned into the pGL3-basic plasmid via Gibson Assembly following the manufacturer's protocol (New England Biolabs, #E2661S). For the mutagenesis assay, we first used PROMO v 3.0.2 (Netanel et al. 2019) to predict the transcription factor (TF) binding sites within the sequences, and then compared the expression levels of these TFs between groups. We selected four TFs for mutation: *FOXM1* (*HNF-3 $\alpha$* ), *TFAP2A*, *MAZ*, and *ETV4*. Specific motifs within the LTR7 promoter-luciferase vector were mutated using the QuikChange Lightning Site-Directed Mutagenesis kit (Agilent, #210518). We used the Neon transfection system (MPK5000) to deliver 400 ng of the promoter-luciferase vector and 200 ng of the pRL-TK *Renilla* vector (Addgene, #E2241) into  $3 \times 10^4$  AsPC-1 cells and  $3 \times 10^4$  293T cells. Luciferase levels were measured after 24 h of incubation using the Dual-Glo Luciferase Assay System (Promega, #E2940).

### Lipid metabolomics mass spectrometry analysis

PANC-1 cells ( $2\text{--}5 \times 10^7$ ) overexpressing different PLAAT4 isoforms were extracted with prechilled HPLC-grade methanol (–80°C), dried, and reconstituted in 100  $\mu$ L isopropanol:methanol (1:1). Lipids were analyzed using a Waters ACQUITY UPLC BEH C18 column (2.1  $\times$  100 mm, 1.7  $\mu$ m) on a UPLC system coupled to an Orbitrap-LTQ-Elite MS (Thermo Fisher Scientific). Gradient and MS parameters are detailed in Supplemental Methods. Data were converted with ProteoWizard v3.0.18178 (Kessner et al. 2008), features extracted using XCMS v3.9.3 (Domingo-Almenara and Siuzdak 2020), and identified by matching to LipidBlast v68 (Tsugawa et al. 2020), with MS1 and MS2 mass tolerances of 0.01 and 0.05 Da, respectively.

### Functional assays

Cell viability was assessed at 0, 24, 48, 72, and 96 h using the CCK-8 kit (Beyotime Biotechnology) and read at 450 nm (Epoch microplate reader, Biotek US). Migration and invasion assays were performed using Transwell inserts (Corning) with or without Matrigel (BD Biosciences) following 24-h serum starvation. Cells

were stained with 0.2% crystal violet and quantified in triplicate experiments.

### High-throughput drug screening of patient-derived organoids

Patient-derived organoids were dissociated and seeded (3000 cells/well) into collagen-precoated 384-well plates. After 24 h, 64 compounds (Selleck) were added in a seven-point dilution series (see Supplemental Table S8) using a Bravo robotic system. Viability was measured after 5 days using CellTiter-Glo (Promega) and read with an Envision plate reader. AUCs were computed using GraphPad Prism 9 and normalized. Clustering was performed with ConsensusClusterPlus v1.64.0 (Wilkerson and Hayes 2010) using parameters: reps=1000, pItem=0.8, pFeature=1, algorithm="km".

### HLA class I antigen isolation from AsPC-1 and organoids

HLA-I peptides were enriched from AsPC-1 cells and PDOs (DAC13) using the NEO Discovery HLA-I Peptide Enrichment kit (Baizhen Biotech), following a modified protocol (Chong et al. 2022). Eluted peptides were analyzed using a timsTOF Pro2 mass spectrometer (Bruker Daltonik) with a 120-min DDA method (MS1 range: m/z 100–1700; PASEF with 10 MS/MS scans, cycle time: 2.27 sec, ion threshold: 1000).

### HLA mass spectrometry analysis

Raw data (.d files) were analyzed using PEAKS DeepNovo Peptidome (Bioinformatics Solutions, Inc.) against the UniProt human-reviewed + NGS libraries. De novo peptides (DeepNovo) were filtered with average local confidence (ALC) > 50%, no enzyme digestion, 20 ppm precursor and 0.05 Da fragment tolerance, and FDR < 1%. Peptides of 7–16 aa were retained and compared to databases: (1) Human Benign Proteome (PXD03878); (2) UniProt (<https://www.uniprot.org/>); and (3) De novo ORF peptide libraries generated using EMBOSS getorf and TEProf2 references. Selected peptides were synthesized, and MS/MS spectra were validated by comparison with eluted HLA peptides.

### HLA typing, binding prediction, and validation

HLA allele types were predicted from RNA-seq using seq2HLA v2.2 (Boegel et al. 2013). Binding affinities of TE-derived peptides were predicted using NetMHCpan-4.1 (Nielsen et al. 2020). Peptides with strong predicted binding (EL-Rank < 0.5) to HLA-A\*02:01 were synthesized for validation.

Binding was tested in T2 cell assays (Tang et al. 2017): cells were incubated with peptide (50  $\mu$ M) and  $\beta$ 2-microglobulin (1  $\mu$ M; Sigma) for 18 h, then stained with PE-labeled anti-HLA-A2.1 (BD Biosciences, #551230, RRID:AB\_394106) and analyzed by flow cytometry (Sony ID7000). Fluorescence intensity (FI) was calculated; FI > 1 indicated strong binding.

### ELISPOT assay

PBMCs from HLA-typed donors were plated ( $2.5 \times 10^5$ /well) on precoated ELISPOT plates (Mabtech, #3420-4APW-2) and stimulated with 20  $\mu$ g/mL peptide for 24 h. After washing, anti-human IFNG antibody (Mabtech) and SA-ALP were applied sequentially (1:1000 dilution). Spots were developed using BCIP substrate and counted with AID ELISPOT Reader and software v7.0.

### Statistical analysis

Two group statistical analysis was performed using the independent samples Student's *t*-test (two-sided) or the Wilcoxon/

Fisher's exact test. Multiple groups statistical analysis was performed using the two-way ANOVA. Data were presented as mean  $\pm$  standard error of mean (SEM); (\*)  $P < 0.05$ , (\*\*)  $P < 0.01$ , (\*\*\*)  $P < 0.001$ , (\*\*\*\*)  $P < 0.0001$ . Sample sizes for each experiment were indicated in the relevant results section or figure legends. All statistical analyses were conducted using R (version 4.3.2) (R Core Team 2024) and GraphPad Prism v8 software.

### Code availability

All source code and custom scripts used in this study are available at GitHub (<https://github.com/MaxlogiLee/PLAAT4>) and as Supplemental Code.

### Data access

All raw and processed sequencing data generated in this study have been submitted to the Genome Sequence Archive (<https://ngdc.cnca.ac.cn/gsa-human>) with the accession number HRA007136.

### Competing interest statement

The authors declare no competing interests.

### Acknowledgments

We thank our patients and their families. This study was supported by the "National Key R&D Program of China" fund (2022YFA1305700, J.Li) granted by the Chinese Ministry of Science and Technology and granted by National Natural Science Foundation of China grant 82372719 (J.Li). We thank Jihang Yuan for experimental assistance.

*Author contributions:* J.Li and G.J. conceived and led the study. Y.W. and Y.R. provided supervision for the wet lab and mass spectrometry, respectively. M.S. conducted data analyses. C.T. and S.Z. carried out the majority of wet lab experiments. X.H. contributed to the T2 binding and immunogeneity experiments. G.Y. checked the analysis of immune peptide experiment results. F.H., R.W., and L.X. provided support for basic statistics, proofreading, data storage, and experiment preparation. Y.F. stored samples in liquid nitrogen. J.Liu conducted the MS experiment. J.Li prepared the manuscript with input from all authors. All authors reviewed the manuscript.

### References

- Attermann AS, Bjerregaard AM, Saini SK, Grønbaek K, Hadrup SR. 2018. Human endogenous retroviruses and their implication for immunotherapeutics of cancer. *Ann Oncol* **29**: 2183–2191. doi:10.1093/annonc/mdy413
- Bailey P, Chang DK, Nones K, Johns AL, Patch AM, Gingras MC, Miller DK, Christ AN, Bruxner TJ, Quinn MC, et al. 2016. Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* **531**: 47–52. doi:10.1038/nature16965
- Biankin AV, Waddell N, Kassahn KS, Gingras MC, Muthuswamy LB, Johns AL, Miller DK, Wilson PJ, Patch AM, Wu J, et al. 2012. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature* **491**: 399–405. doi:10.1038/nature11547
- Boegel S, Löwer M, Schäfer M, Bukur T, de Graaf J, Boisguérin V, Türeci Ö, Diken M, Castle JC, Sahin U. 2013. HLA typing from RNA-seq sequence reads. *Genome Med* **4**: 102. doi:10.1186/gm403
- Bradley RM, Hashemi A, Aristizabal-Henao JJ, Stark KD, Duncan RE. 2022. PLAAT1 exhibits phosphatidylcholine:monolysocardiolipin transacylase activity. *Int J Mol Sci* **23**: 6714. doi:10.3390/ijms23126714
- Brocks D, Schmidt CR, Daskalakis M, Jang HS, Shah NM, Li D, Li J, Zhang B, Hou Y, Laudato S, et al. 2017. DNMT and HDAC inhibitors induce cryptic transcription start sites encoded in long terminal repeats. *Nat Genet* **49**: 1052–1060. doi:10.1038/ng.3889

- The Cancer Cell Line Encyclopedia Consortium, The Genomics of Drug Sensitivity in Cancer Consortium. 2015. Pharmacogenomic agreement between two cancer cell line data sets. *Nature* **528**: 84–87. doi:10.1038/nature15736
- Cao L, Huang C, Cui Zhou D, Hu Y, Lih TM, Savage SR, Krug K, Clark DJ, Schnaubelt M, Chen L, et al. 2021. Proteogenomic characterization of pancreatic ductal adenocarcinoma. *Cell* **184**: 5031–5052.e26. doi:10.1016/j.cell.2021.08.023
- Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**: i884–i890. doi:10.1093/bioinformatics/bty560
- Chiappinelli KB, Strissel PL, Desrichard A, Li H, Henke C, Akman B, Hein A, Rote NS, Cope LM, Snyder A, et al. 2015. Inhibiting DNA methylation causes an interferon response in cancer via dsRNA including endogenous retroviruses. *Cell* **162**: 974–986. doi:10.1016/j.cell.2015.07.011
- Chong C, Coukos G, Bassani-Sternberg M. 2022. Identification of tumor antigens with immunopeptidomics. *Nat Biotechnol* **40**: 175–188. doi:10.1038/s41587-021-01038-8
- Choudhary MNK, Quaid K, Xing X, Schmidt H, Wang T. 2023. Widespread contribution of transposable elements to the rewiring of mammalian 3D genomes. *Nat Commun* **14**: 634. doi:10.1038/s41467-023-36364-9
- Dbouk M, Katona BW, Brand RE, Chak A, Syngal S, Farrell JJ, Kastrinos F, Stoffel EM, Blackford AL, Rustgi AK, et al. 2022. The multicentric Cancer of Pancreas Screening study: impact on stage and survival. *J Clin Oncol* **40**: 3257–3266. doi:10.1200/JCO.22.00298
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21. doi:10.1093/bioinformatics/bts635
- Domingo-Almenara X, Siuzdak G. 2020. Metabolomics data processing using XCMS. *Methods Mol Biol* **2104**: 11–24. doi:10.1007/978-1-0716-0239-3\_2
- Dumesnil C, Vanharanta L, Prasanna X, Omrane M, Carpentier M, Bhopkar A, Enkavi G, Salo VT, Vattulainen I, Ikonen E, et al. 2023. Cholesterol esters form supercooled lipid droplets whose nucleation is facilitated by triacylglycerols. *Nat Commun* **14**: 915. doi:10.1038/s41467-023-36375-6
- Ely ZA, Kulstad ZJ, Gunaydin G, Addepalli S, Verzani EK, Casarrubios M, Clauser KR, Wang X, Lippincott IE, Louvet C, et al. 2025. Pancreatic cancer-restricted cryptic antigens are targets for T cell recognition. *Science* **388**: eadk3487. doi:10.1126/science.adk3487
- Elyada E, Bolisetty M, Laise P, Flynn WF, Courtois ET, Burkhart RA, Teinor JA, Belleau P, Biffi G, Lucito MS, et al. 2019. Cross-species single-cell analysis of pancreatic ductal adenocarcinoma reveals antigen-presenting cancer-associated fibroblasts. *Cancer Discov* **9**: 1102–1123. doi:10.1158/2159-8290.CD-19-0094
- Espinete E, Gu Z, Imbusch CD, Giese NA, Büscher M, Safavi M, Weisenburger S, Klein C, Vogel V, Falcone M, et al. 2021. Aggressive PDACs show hypomethylation of repetitive elements and the execution of an intrinsic IFN program linked to a ductal cell of origin. *Cancer Discov* **11**: 638–659. doi:10.1158/2159-8290.CD-20-1202
- Fujimoto T, Parton RG. 2011. Not just fat: the structure and function of the lipid droplet. *Cold Spring Harb Perspect Biol* **3**: a004838. doi:10.1101/cshperspect.a004838
- Ghandi M, Huang FW, Jané-Valbuena J, Kryukov GV, Lo CC, McDonald ER, Barretina J, Gelfand ET, Bielski CM, Li H, et al. 2019. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**: 503–508. doi:10.1038/s41586-019-1186-3
- Golczak M, Kiser PD, Sears AE, Lodowski DT, Blaner WS, Palczewski K. 2012. Structural basis for the acyltransferase activity of lecithin:retinol acyltransferase-like proteins. *J Biol Chem* **287**: 23790–23807. doi:10.1074/jbc.M112.361550
- Hayashi A, Hong J, Iacobuzio-Donahue CA. 2021. The pancreatic cancer genome revisited. *Nat Rev Gastroenterol Hepatol* **18**: 469–481. doi:10.1038/s41575-021-00463-z
- Hoenisch Gravel N, Nelde A, Bauer J, Mühlenbruch L, Schroeder SM, Neidert MC, Scheid J, Lemke S, Dubbelaar ML, Wacker M, et al. 2023. TOFIMS mass spectrometry-based immunopeptidomics refines tumor antigen identification. *Nat Commun* **14**: 7472. doi:10.1038/s41467-023-42692-7
- Huang D, Zhu X, Ye S, Zhang J, Liao J, Zhang N, Zeng X, Wang J, Yang B, Zhang Y, et al. 2024. Tumour circular RNAs elicit anti-tumour immunity by encoding cryptic peptides. *Nature* **625**: 593–602. doi:10.1038/s41586-023-06834-7
- Jacob F, Salinas RD, Zhang DY, Nguyen PTT, Schnoll JG, Wong SZH, Thokala R, Sheikh S, Saxena D, Prokop S, et al. 2020. A patient-derived glioblastoma organoid model and biobank recapitulates inter- and intra-tumoral heterogeneity. *Cell* **180**: 188–204.e22. doi:10.1016/j.cell.2019.11.036
- Jang HS, Shah NM, Du AY, Dailey ZZ, Pehrsson EC, Godoy PM, Zhang D, Li D, Xing X, Kim S, et al. 2019. Transposable elements drive widespread expression of oncogenes in human cancers. *Nat Genet* **51**: 611–617. doi:10.1038/s41588-019-0373-3
- Jun Y, Tam OH, Paniagua E, Hammell M. 2015. Tetrascripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics* **31**: 3593–3599. doi:10.1093/bioinformatics/btv422
- Jung Y, Han D, Marschall T. 2022. BWA-MEME: BWA-MEM emulated with a machine learning approach. *Bioinformatics* **38**: 2404–2413. doi:10.1093/bioinformatics/btac137
- Kang Y-J, Yang D-C, Kong L, Hou M, Meng Y-Q, Wei L, Gao G. 2017. CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Res* **45**: W12–W16. doi:10.1093/nar/gkx428
- Kessner D, Chambers M, Burke R, Agus D, Mallick P. 2008. ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics* **24**: 2534–2536. doi:10.1093/bioinformatics/btn323
- Khorkova O, Hsiao J, Wahlestedt C. 2015. Basic biology and therapeutic implications of lncRNA. *Adv Drug Deliv Rev* **87**: 15–24. doi:10.1016/j.addr.2015.05.012
- Kozak M. 2005. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene* **361**: 13–37. doi:10.1016/j.gene.2005.06.037
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754–1760. doi:10.1093/bioinformatics/btp324
- Li J, Gu D, Lee SSY, Song B, Bandyopadhyay S, Chen S, Konieczny SF, Ratliff TL, Liu X, Xie J, et al. 2016. Abrogating cholesterol esterification suppresses growth and metastasis of pancreatic cancer. *Oncogene* **35**: 6378–6388. doi:10.1038/onc.2016.168
- Li T-H, Wang Y-Y, Zhao B-B, Qin C, Li Z-R, Wang W-B. 2022. Phospholipase A/acyltransferase 4 is a prognostic biomarker and correlated with immune infiltrates in pancreatic cancer. *Heliyon* **8**: e10416. doi:10.1016/j.heliyon.2022.e10416
- Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923–930. doi:10.1093/bioinformatics/btt656
- Liu Y, Sun J, Zhao M. 2017. ONGene: a literature-based database for human oncogenes. *Journal of Genetics and Genomics* **44**: 119–121. doi:10.1016/j.jgg.2016.12.004
- Lizio M, Harshbarger J, Shimoji H, Severin J, Kasukawa T, Sahin S, Abugessaisa I, Fukuda S, Hori F, Ishikawa-Kato S, et al. 2015. Gateways to the FANTOM5 promoter level mammalian expression atlas. *Genome Biol* **16**: 22. doi:10.1186/s13059-014-0560-6
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550. doi:10.1186/s13059-014-0550-8
- Ma Y, Zhou X. 2022. Spatially informed cell-type deconvolution for spatial transcriptomics. *Nat Biotechnol* **40**: 1349–1359. doi:10.1038/s41587-022-01273-7
- Maurer C, Holmstrom SR, He J, Laise P, Su T, Ahmed A, Hibshoosh H, Chabot JA, Oberstein PE, Sepulveda AR, et al. 2019. Experimental microdissection enables functional harmonisation of pancreatic cancer subtypes. *Gut* **68**: 1034–1043. doi:10.1136/gutjnl-2018-317706
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**: 1297–1303. doi:10.1101/gr.107524.110
- McLean CY, Bristor D, Hiller M, Clarke SL, Schaaf BT, Lowe CB, Wenger AM, Bejerano G. 2010. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* **28**: 495–501. doi:10.1038/nbt.1630
- Moffitt RA, Marayati R, Flate EL, Volmar KE, Loeza SG, Hoadley KA, Rashid NU, Williams LA, Eaton SC, Chung AH, et al. 2015. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat Genet* **47**: 1168–1178. doi:10.1038/ng.3398
- Moncada R, Barkley D, Wagner F, Chiodin M, Devlin JC, Baron M, Hajdu CH, Simeone DM, Yanai I. 2020. Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nat Biotechnol* **38**: 333–342. doi:10.1038/s41587-019-0392-8
- Netanel D, Stern N, Laufer I, Shamir R. 2019. PROMO: an interactive tool for analyzing clinically-labeled multi-omic cancer datasets. *BMC Bioinformatics* **20**: 732. doi:10.1186/s12859-019-3142-5
- Newey A, Griffiths B, Michaux J, Pak HS, Stevenson BJ, Woolston A, Semiannikova M, Spain G, Barber LJ, Matthews N, et al. 2019. Immunopeptidomics of colorectal cancer organoids reveals a sparse HLA class I neoantigen landscape and no increase in neoantigens with interferon or MEK-inhibitor treatment. *J Immunother Cancer* **7**: 309. doi:10.1186/s40425-019-0769-8
- Nielsen M, Peters B, Paul S, Alvarez B, Reynisson B. 2020. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res* **48**: W449–W454. doi:10.1093/nar/gkaa379

- Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* **33**: 290–295. doi:10.1038/nbt.3122
- Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM. 2014. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res* **74**: 2913–2921. doi:10.1158/0008-5472.CAN-14-0155
- Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. 2014. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**: W187–W191. doi:10.1093/nar/gku365
- R Core Team. 2024. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/>.
- Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* **16**: 276–277. doi:10.1016/S0168-9525(00)02024-2
- Roulois D, Yau HL, Singhanian R, Wang Y, Danesh A, Shen SY, Han H, Liang G, Jones PA, Pugh TJ, et al. 2015. DNA-demethylating agents target colorectal cancer cells by inducing viral mimicry by endogenous transcripts. *Cell* **162**: 961–973. doi:10.1016/j.cell.2015.07.056
- Savvitska N, Heutink P, Bansal V. 2022. Transcription start site signal profiling improves transposable element RNA expression analysis at locus-level. *Front Genet* **13**: 1026847. doi:10.3389/fgene.2022.1026847
- Shah NM, Jang HJ, Liang Y, Maeng JH, Tzeng S-C, Wu A, Basri NL, Qu X, Fan C, Li A, et al. 2023. Pan-cancer analysis identifies tumor-specific antigens derived from transposable elements. *Nat Genet* **55**: 631–639. doi:10.1038/s41588-023-01349-3
- Shi X, Li Y, Yuan Q, Tang S, Guo S, Zhang Y, He J, Zhang X, Han M, Liu Z, et al. 2022. Integrated profiling of human pancreatic cancer organoids reveals chromatin accessibility features associated with drug sensitivity. *Nat Commun* **13**: 2169. doi:10.1038/s41467-022-29857-6
- Siegel RL, Giaquinto AN, Jemal A. 2024. Cancer statistics, 2024. *CA Cancer J Clin* **74**: 12–49. doi:10.3322/caac.21820
- Sun T, Xu Y, Xiang Y, Ou J, Soderblom EJ, Diaoy Y. 2023. Crosstalk between RNA m<sup>6</sup>A and DNA methylation regulates transposable element chromatin activation and cell fate in human pluripotent stem cells. *Nat Genet* **55**: 1324–1335. doi:10.1038/s41588-023-01452-5
- Tang K, Cheng L, Zhang C, Zhang Y, Zheng X, Zhang Y, Zhuang R, Jin B, Zhang F, Ma Y. 2017. Novel identified HLA-A\*0201-restricted Hantaan virus glycoprotein cytotoxic T-cell epitopes could effectively induce protective responses in HLA-A2.1/Kb transgenic mice may associate with the severity of hemorrhagic fever with renal syndrome. *Front Immunol* **8**: 1797. doi:10.3389/fimmu.2017.01797
- Tsugawa H, Ikeda K, Takahashi M, Satoh A, Mori Y, Uchino H, Okahashi N, Yamada Y, Tada I, Bonini P, et al. 2020. A lipidome atlas in MS-DIAL 4. *Nat Biotechnol* **38**: 1159–1163. doi:10.1038/s41587-020-0531-2
- Tu T, Zhang H, Xu H. 2023. Targeting sterol-O-acyltransferase 1 to disrupt cholesterol metabolism for cancer therapy. *Front Oncol* **13**: 1197502. doi:10.3389/fonc.2023.1197502
- Tyanova S, Temu T, Cox J. 2016. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* **11**: 2301–2319. doi:10.1038/nprot.2016.136
- Uyama T, Ichi I, Kono N, Inoue A, Tsuboi K, Jin X-H, Araki N, Aoki J, Arai H, Ueda N. 2012a. Regulation of peroxisomal lipid metabolism by catalytic activity of tumor suppressor H-rev107. *J Biol Chem* **287**: 2706–2718. doi:10.1074/jbc.M111.267575
- Uyama T, Ikematsu N, Inoue M, Shinohara N, Jin X-H, Tsuboi K, Tonai T, Tokumura A, Ueda N. 2012b. Generation of N-acylphosphatidylethanolamine by members of the phospholipase A/acyltransferase (PLA/AT) family. *J Biol Chem* **287**: 31905–31919. doi:10.1074/jbc.M112.368712
- Wang K, Li M, Hakonarson H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**: e164. doi:10.1093/nar/gkq603
- Wilkinson MD, Hayes DN. 2010. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26**: 1572–1573. doi:10.1093/bioinformatics/btq170
- Xiang X, Tao Y, DiRusso J, Hsu F-M, Zhang J, Xue Z, Pontis J, Trono D, Liu W, Clark AT. 2022. Human reproduction is regulated by retrotransposons derived from ancient Hominidae-specific viral infections. *Nat Commun* **13**: 463. doi:10.1038/s41467-022-28105-1
- Xie M, Hong C, Zhang B, Lowdon RF, Xing X, Li D, Zhou X, Lee HJ, Maire CL, Ligon KL, et al. 2013. DNA hypomethylation within specific transposable element families associates with tissue-specific enhancer landscape. *Nat Genet* **45**: 836–841. doi:10.1038/ng.2649
- Yadav M, Jhunjhunwala S, Phung QT, Lupardus P, Tanguay J, Bumbaca S, Franci C, Cheung TK, Fritsche J, Weinschenk T, et al. 2014. Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. *Nature* **515**: 572–576. doi:10.1038/nature14001
- Zadoorian A, Du X, Yang H. 2023. Lipid droplet biogenesis and functions in health and disease. *Nat Rev Endocrinol* **19**: 443–459. doi:10.1038/s41574-023-00845-0
- Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137. doi:10.1186/gb-2008-9-9-r137
- Zheng R, Wan C, Mei S, Qin Q, Wu Q, Sun H, Chen C-H, Brown M, Zhang X, Meyer CA, et al. 2019. Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis. *Nucleic Acids Res* **47**: D729–D735. doi:10.1093/nar/gky1094
- Zhuo X, Hsu S, Purushotham D, Kuntala PK, Harrison JK, Du AY, Chen S, Li D, Wang T. 2023. Comparing genomic and epigenomic features across species using the WashU Comparative Epigenome Browser. *Genome Res* **33**: 824–835. doi:10.1101/gr.277550.122

Received February 9, 2025; accepted in revised form November 24, 2025.