Intelligent Oncology 1 (2025) 233–243

Contents lists available at ScienceDirect

Intelligent Oncology

journal homepage: www.keaipublishing.com/en/intonc/

Full Length Article

# AI-driven network biology identifies SRC as a therapeutic target in metastatic pancreatic adenocarcinoma

Ayla Zhang[a,b], Jake Y. Chen[a,b,*]

[a] AlphaMind Club, LLC, Birmingham AL 35244, USA
[b] Systems Pharmacology AI Research Center, School of Medicine, The University of Alabama at Birmingham, Birmingham AL 35233, USA

ARTICLE INFO

*Keywords:*
Pancreatic adenocarcinoma
SRC kinase
ChatGPT-4o evaluation framework
Bioinformatics
Oncology drug development

ABSTRACT

Pancreatic adenocarcinoma (PAAD) is notorious for its limited treatment options and dismal prognosis, underscoring an urgent need for innovative therapeutic strategies. This study leverages advanced bioinformatics and a novel ChatGPT-4o evaluation framework to identify and validate SRC kinase as a significant drug target for PAAD. Initial bioinformatics screening illustrated SRC's critical involvement in essential signaling pathways associated with the disease. Subsequently, the ChatGPT-4o framework provided an unbiased, efficient evaluation, emphasizing SRC's role in disease mechanisms. Our findings reveal substantial overexpression of SRC in PAAD compared to normal pancreatic tissue, with expression levels escalating in line with tumor grade and stage. Protein expression analysis using the Human Protein Atlas dataset showed a statistically significant difference between PAAD samples and normal samples for SRC ($P < 1 \times 10^{-15}$). Survival analysis using the Kaplan-Meier Plotter dataset further demonstrates a strong association between high SRC expression and reduced survival, with hazard ratios of 1.45 ($P < 7.69 \times 10^{-3}$) for overall survival and 2.15 ($P < 9.19 \times 10^{-5}$) for disease free survival in advanced disease stages. SRC's involvement in key oncogenic pathways suggests its potential as part of combination therapies, enhancing the efficacy of existing treatments by targeting complementary mechanisms. This study shows the therapeutic relevance of SRC and demonstrates the effectiveness of combining innovative AI technologies with traditional bioinformatics to accelerate the discovery and validation of critical drug targets in oncology.

## 1. Introduction

Pancreatic adenocarcinoma (PAAD) is a highly aggressive solid tumor that constitutes over 90% of all pancreatic cancer cases.[1] Pancreatic ductal adenocarcinoma (PDAC) is a common subtype of PAAD, although these two terms are sometimes used interchangeably. PAAD is one of the most challenging malignancies to treat, with a 5-year survival rate of less than 10%,[2] and currently ranks third in cancer-related deaths in the United States,[3] with projections indicating it will become the second-deadliest cancer by 2030.[4]

PAAD is commonly characterized by its poor prognosis and ineffective treatment options. PAAD tumors progress rapidly without recognizable symptoms; at the time of diagnosis, over 80% of patients have experienced metastasis and spread of the disease, evidencing PAAD's poor prognosis and resulting in advanced tumors ineligible for surgery.[5] Standard treatments for PAAD, such as chemotherapy and immunotherapy, have failed to produce meaningful results measured by patient survival, since the immunosuppressive and intensely desmoplastic tumor microenvironment in PAAD often limits tumor exposure to chemotherapy and attack by immune cells.[6,7] For example, approved chemotherapies such as gemcitabine/nab-paclitaxel or FOLFIRINOX only improve survival by 2–4 months and are accompanied by severe toxicity and adverse effects.[8]

Current targeted approaches to PAAD drug development have been researched extensively with the availability of cancer genomic sequencing data. For example, KRAS, which is mutated in nearly 90% of PAAD patient samples,[9] has become a common drug target for investigational drugs. However, leading drugs involved in clinical trials for PAAD treatment, such as sotorasib, which targets the KRAS G12C mutation, have shown minimal survival improvements of around 3 months, and also commonly exhibit adverse effects.[10] These results may be attributed to the low frequency (1%–2%) of KRAS G12C mutations in

* Corresponding author.
*E-mail address:* jakechen@uab.edu (J.Y. Chen).

https://doi.org/10.1016/j.intonc.2025.06.004
Received 23 February 2025; Received in revised form 12 June 2025; Accepted 17 June 2025
Available online 2 July 2025
2950-2616/© 2025 The Author(s). Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

pancreatic cancer.[11] Additionally, KRAS mutations are early events present in the initial stages of PAAD[9] and may have reduced influence in advanced PAAD stages, which is when most PAAD patients are typically diagnosed.[2] Cancer progression may also lead to the accumulation of both driver and passenger mutations, resulting in genetic instability, especially at advanced stages.[12,13] Therefore, recent research regarding PAAD drug targets has shifted towards genomic tumor drivers that are predominantly involved in PAAD signaling pathways as well as differentially expressed in tumor tissues.[14–16]

In this paper, we describe the bioinformatics process of establishing SRC as a new PAAD drug target with a systematic, unbiased approach, by comparing it to existing targets and evaluating its relation to PAAD with regard to expression, significance in late stage tumors, and impact on patient survival. The approach that we developed has the following innovative characteristics. First, we performed a comprehensive pathway analysis designed to evaluate a drug target's relevance to PAAD progression and suitability as a novel target for PAAD treatment. Second, we leveraged ChatGPT-4o in our evaluation tasks to improve efficiency and demonstrate the utility of generative AI in de novo drug discovery. Unlike many prior studies, which often focus exclusively on either computational biology techniques or manual data interpretation, this combined approach leverages the strengths of both traditional bioinformatics tools and generative AI. This integration facilitates improved data processing, enhanced interpretation of complex biological networks, and streamlined validation of potential drug targets. ChatGPT-4o, the latest iteration of the Generative Pre-trained Transformer Large Language Model developed by OpenAI, exhibits improved understanding and generation of complex concepts and text, providing a potential tool to aid in de novo drug discovery through contributions in data interpretation, simulation, modeling, and more.[17] Finally, by integrating this ChatGPT-4o evaluation with bioinformatic pathway and survival analysis, we efficiently uncover a novel drug target candidate and streamline the drug discovery process. In this paper, we will demonstrate that this novel approach has been successfully validated by cohort analysis using SRC expression and its impact on patient survival.

## 2. Methods

Our methodology was structured into seven sequential steps to systematically evaluate and prioritize candidate drug targets for pancreatic cancer targeted therapy (Figure 1). First, we curated a comprehensive list of 252 candidate drug targets using Pathway Annotated Gene Enrichment Ranking (PAGER), a bioinformatics platform that integrates pathway-associated gene sets to identify pancreatic cancer-relevant genes and pathways. Second, we retrieved biomolecular pathways associated with our candidate gene set using PAGER. Third, we utilized a custom evaluation framework generated by ChatGPT-4o to score each pathway, leveraging the AI's capabilities to ensure an unbiased and efficient analysis of pathway relevance and drug target suitability. Fourth, we calculated a baseline pathway score by evaluating unrelated pathways with the same framework, using ChatGPT-4o and PAGER to establish benchmarks for pathway significance. Fifth, we integrated pathway scores and individual gene scores to calculate weighted final ranking scores for each candidate gene. Sixth, we conducted expression analysis to investigate differential gene expression and trends of gene expression across stages of PAAD progression. Seventh, and finally, we performed survival analysis to investigate the impact of SRC expression on patient survival outcomes. Next, we will describe the approaches in detail.

### 2.1. Candidate drug target curation

We first manually compiled an extensive list of candidates to establish a prioritized list of candidate drug targets. This initial compilation involved searching for pancreatic cancer-related pathways, annotated lists, and gene signatures (PAGs) using PAGER, which is a bioinformatics platform that identifies and analyzes PAGs associated with a user-provided gene set. PAGER has 113,830 PAGs from 37 data sources, ensuring comprehensive coverage of relevant pathways.[18] This extensive coverage as well as PAGER's prioritization algorithms allow it to uncover critical genes that other tools might overlook. Additionally, its ability to integrate diverse datasets ensures a more holistic understanding of pathway relevance in complex diseases such as pancreatic cancer. From this search, we identified 252 genes that exhibited high prioritization scores within the resulting PAGs, forming our comprehensive initial candidate list.

### 2.2. Curation of PAAD-affected biomolecular pathways

To effectively rank the candidate drug targets, we established a comprehensive evaluation process that assessed each candidate in relation to PAAD. Our approach went beyond direct assessment of the drug targets themselves; instead, we focused on retrieving and evaluating biomolecular pathways that are impacted in PAAD and pertinent to each candidate. This strategy was designed to ensure that candidates

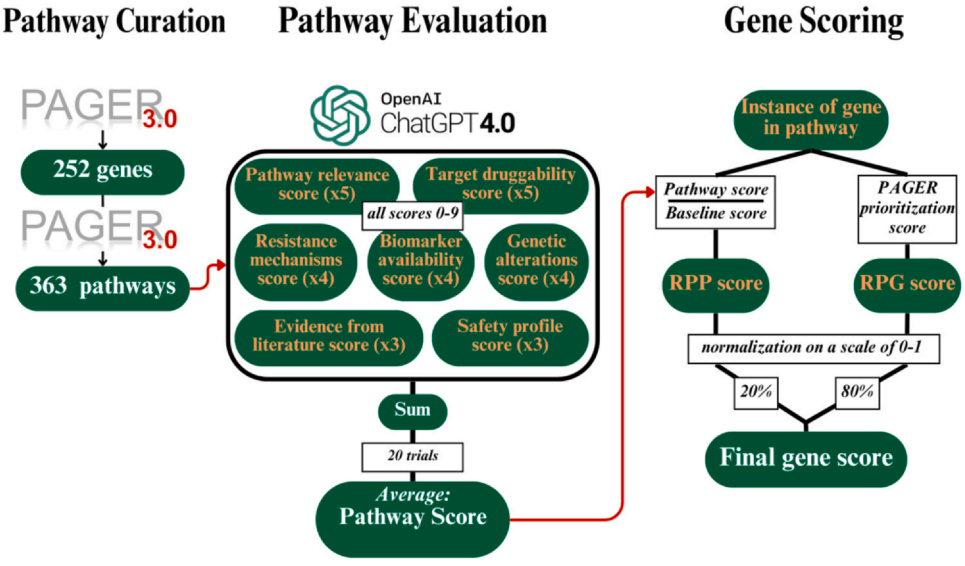



**Figure 1.** Flowchart of methodology.

with limited research literature in the context of PAAD were not overlooked, thereby facilitating the discovery of novel drug targets.[19–21]

We employed PAGER to source pathways associated with each candidate within the context of PAAD. We then queried our initial candidate list in PAGER and compiled the resulting pathway-type PAGs with at least five genes overlapping with the candidate list. The threshold of five overlapping genes was selected to balance comprehensiveness and relevance, ensuring that pathways with minimal overlap, which might lack sufficient biological context or statistical power, were excluded. This threshold aimed to focus the analysis on pathways more likely to play a significant role in PAAD progression while controlling the number of pathways included for downstream analysis. However, we acknowledge that this threshold could potentially overlook important pathways with fewer overlapping genes, particularly if those genes are highly significant to PAAD. To address this limitation, future analyses could include complementary methods, such as weighting gene overlap by functional relevance or incorporating additional pathway-level metrics, to capture pathways with fewer but highly impactful genes.

To focus on pathway databases that provide more experimentally validated, disease-specific, and actionable insights, we filtered this pathway list by data source to exclude MSigDB and PharmGKB. This resulted in 363 PAAD-related pathways. These detailed findings are presented in Supplementary File 1.

Additionally, while we included umbrella-level pathways (e.g., hsa05212 pancreatic cancer) in our evaluation, we acknowledge that core signaling pathways such as KRAS-MAPK, p53, and Rb/E2F often provide more precise mechanistic and therapeutic insights. However, umbrella pathways offer a curated, integrative view of multiple intersecting biological processes that are specific to PAAD and reflect higher-order coordination of oncogenic mechanisms. These composite pathways are particularly useful in identifying candidate genes that participate across multiple subsystems and are therefore more likely to serve as viable, high-impact therapeutic targets. Moreover, umbrella pathways from sources like KEGG often include experimental annotations that consolidate the contributions of canonical pathways in disease-specific contexts, providing a practical and interpretable basis for systematic scoring. Our approach balances both perspectives by scoring umbrella pathways while retaining core pathway signals through gene level scoring (Rank Prioritization Gene [RPG]) and inclusion of pathways like PI3K-Akt, p53, and MAPK in our final ranking. This ensures comprehensive yet therapeutically relevant prioritization.

### *2.3. Evaluation of pathway contribution to PAAD progression using ChatGPT*

In our study, we utilized ChatGPT-4o to generate a comprehensive evaluation framework to assess the significance and influence of specific biomolecular pathways in the context of pancreatic cancer. While traditional platforms such as PAGER provide robust influence scoring based on statistical enrichment and gene overlap metrics, they are limited in synthesizing multidimensional biological context, clinical relevance, and druggability simultaneously. We chose ChatGPT-4o for its ability to integrate complex and heterogeneous biomedical data into a structured, multi-criteria scoring framework. ChatGPT-4o enabled us to systematically evaluate pathways not only based on gene content but also their literature-supported roles in cancer biology, known drug interactions, biomarker availability, and resistance mechanisms—factors that are not easily quantified in conventional tools. The primary advantage of using ChatGPT lies in its ability to extract contextual knowledge from diverse sources and apply human-like reasoning to prioritize targets, as well as its efficiency and scalability. Additionally, previous research has demonstrated ChatGPT's consistent performance in evaluation tasks compared to human evaluators and its reliability in medical and biological contexts,[17,22–24] making it an efficient and accurate tool for our analysis.

Ranking biomolecular pathways based on their relevance to specific diseases is a well-established bioinformatics and systems biology approach. Various studies have proposed scoring frameworks integrating multiple criteria, such as pathway relevance, genetic mutations, clinical biomarkers, and therapeutic potential. For example, Subramanian et al.[25] introduced the Gene Set Enrichment Analysis (GSEA) framework, which evaluates pathway significance based on differential expression patterns across gene sets. Similarly, Kanehisa et al.[26] used KEGG pathways to prioritize pathway significance in cancer by examining functional annotations and mutation frequencies.

To further augment traditional frameworks, we integrated ChatGPT-4o into our pathway evaluation process. ChatGPT-4o synthesizes and contextualizes information from diverse datasets, providing a streamlined approach to multi-criteria scoring. By using structured prompts to evaluate each pathway across defined categories, ChatGPT leverages existing literature and datasets to generate objective scores.[27–29]

We employed a numerical scoring system to minimize potential errors, such as hallucinations and ambiguities commonly associated with AI, fully detailed in Supplementary File 2. The ChatGPT-4o-generated framework systematically evaluated each pathway across seven defined categories to assess its role in PAAD progression and its potential as a drug target. These categories were selected based on established criteria in cancer drug discovery frameworks, including KEGG druggability annotations, the Cancer Genome Atlas (TCGA) mutation and expression data, and biomarker-guided therapy models. To evaluate each pathway, the pathway name and scoring system were inputted as prompts for ChatGPT-4o, which then provided the evaluation scores. These prompts were slightly varied between trials to ensure robustness and reduce the influence of prompt-specific biases, thereby improving the reliability of the evaluation scores. These varied prompts can be found in Supplementary File 2.

Each pathway was given a subscore on a scale from 0 to 9 in each of seven categories, with 0 indicating no relevance or evidence and 9 indicating strong, well-supported significance. These subscores for each category were multiplied by specific weights, from 1 to 5, assigned to each category based on their relative importance in determining a pathway's therapeutic potential.

Pathway relevance to PAAD was assigned the highest weight of 5 because it directly evaluates the pathway's involvement in PAAD-specific processes such as tumor initiation, progression, and therapeutic resistance. These processes are central to understanding the disease's biology and identifying actionable drug targets. A pathway that demonstrates strong evidence of being integral to PAAD biology is inherently more likely to yield meaningful therapeutic opportunities. Pathway relevance was evaluated by analyzing the pathway's documented involvement in key biological processes such as tumor initiation, progression, metastasis, and therapeutic resistance.

Target druggability, weighted at 5 to emphasize the practical feasibility of translating pathway findings into therapeutic strategies, assessed the presence of well-characterized, actionable drug targets within the pathway and the availability of drugs targeting these components. The high weight correlates with the importance of pathways containing actionable components that can be targeted with existing drugs or drug-like compounds. Pathways with multiple approved drugs or drugs in advanced clinical trials received higher scores, while those with hypothetical or preclinical targets received lower scores.

Genetic alterations in the pathway were given a weight of 4, due to its role in indicating the fundamental molecular drivers of PAAD and providing a foundation for stratifying and targeting patient subgroups. The slightly lower weight compared to relevance and druggability reflects the fact that not all genetic alterations are immediately actionable or relevant to late-stage disease. Genetic alterations were evaluated by examining the frequency and impact of mutations in pathway-associated genes among PAAD patients. Pathways containing genes with high mutation prevalence, such as KRAS or CDKN2A, received higher scores, especially if these mutations were linked to tumor progression or poor prognosis.

Biomarker availability was weighted at 4 because it is essential for tailoring treatments and monitoring therapeutic responses. The weighting acknowledges its importance while recognizing that some highly relevant pathways may not yet have validated biomarkers but still hold therapeutic promise. It assessed whether the pathway was associated with reliable markers for patient stratification or treatment monitoring. Pathways with clinically validated biomarkers, such as CA19–9, received higher scores, while those with experimental or unvalidated biomarkers scored lower.

Resistance mechanisms, crucial for overcoming therapeutic challenges and improving clinical outcomes, were weighted at 4 and focused on the pathway's involvement in resistance to existing therapies for PAAD, such as gemcitabine or 5-FU. Pathways with strong evidence from clinical or preclinical studies linking them to therapy resistance scored higher, while those with minimal or inconclusive evidence scored lower.

Literature evidence helps contextualize findings but is inherently dependent on research focus, which may bias representation; therefore, it was weighted at 3. Literature evidence considered the depth and breadth of research supporting the pathway's relevance to PAAD. Pathways with extensive support from clinical trials or high-impact publications received higher scores, while those with limited or speculative evidence received lower scores.

Lastly, the safety profile, weighted at 3 since it may be less determinative in early-stage target identification than other metrics, evaluated the feasibility of targeting the pathway based on known or anticipated side effects. Pathways associated with well-tolerated therapeutic interventions scored higher, while those with significant toxicity concerns or off-target effects received lower scores.

We now provide an explicit example of the structured prompt used in the evaluation process. The prompts were explicitly designed to reflect the rubric categories and simulate domain-expert reasoning while limiting overfitting. We monitored outputs for consistency, and by averaging across multiple independently sampled runs, we reduced susceptibility to AI hallucinations or overconfidence. This design allows our framework to remain interpretable, reproducible, and aligned with existing benchmarks in gene and pathway prioritization. For each pathway, we input the following prompt into ChatGPT-4o:

"For each pathway I provide, apply the defined scoring system tailored to pancreatic cancer research. Score each subcategory as a biology expert would, limiting your literature review to a randomly chosen 15% of sources for accurate and concise scoring. Output each pathway in a tab-separated format without headers or a final score, so I can directly copy the results. Ensure complete scores for each pathway, with attention to accuracy in the context of pancreatic cancer."

As an example, for the EGF/EGFR signaling pathway, ChatGPT returned the following sub-scores based on our 7-category framework: 9 for Pathway Relevance to PAAD, 8 for Target Druggability, 7 for Genetic Alterations, 7 for Biomarker Availability, 8 for Resistance Mechanisms, 9 for Literature Evidence, and 7 for Safety Profile. Each category score was multiplied by its assigned weight (see Supplementary File 2), and the final pathway score (e.g., 221 for EGF/EGFR) was the weighted sum. These weighted sums were used to rank pathways according to their overall potential as targetable pathways in PAAD.

This process was repeated across all 363 PAAD-related pathways and 129 unrelated pathways, with each score averaged over 20 trials to reduce prompt variability. To ensure robustness, we evaluated the scoring framework's sensitivity to changes in category weights. Pathway relevance and druggability weights varied by ±10%, but the overall rankings remained consistent for the top 10 pathways, demonstrating stability in scoring.

### 2.4. Establishment of pathway baseline score

To provide a reference point for pathway evaluation, we generated a list of random genes not included in our candidate gene list using ChatGPT-4o. This gene list, "unrelated genes," was subsequently queried in PAGER to identify 129 pathways unrelated to PAAD, termed "unrelated pathways."

The ChatGPT-4o scoring system was applied to each unrelated pathway using the same methodology employed for PAAD-related pathways. Each pathway was evaluated 20 times, and the average score across all trials was calculated to determine the final ranking score for each unrelated pathway. The final scores of all unrelated pathways were then averaged to establish the pathway baseline score, which served as a benchmark for assessing the significance of pathways in the context of PAAD.

To ensure robustness, we evaluated the scoring framework's sensitivity to changes in category weights. Pathway relevance and druggability weights were varied by 10%, and we recalculated final pathway scores and re-ranked the full list. The overall rankings remained the same for the top 10 pathways, demonstrating the stability of the scoring algorithm.

### 2.5. Integration of pathway evaluation and gene prioritization scores

Each candidate gene's membership in the 363 previously evaluated pathways was systematically recorded. This was accomplished by examining the "PAG Members" table associated with each pathway in the PAGER database. Two distinct scores were assigned for every instance of a gene being part of a pathway: the Rank Prioritization Pathway (RPP) score and the RPG score. The RPP score indicates the standardized pathway significance relative to its baseline, derived by dividing the pathway ranking score, obtained through ChatGPT-4o's evaluation, by the pathway baseline score. The RPG score corresponds to the rank prioritization score of the gene within the specific pathway, as provided by the PAGER database. This score directly reflects the gene's importance or contribution within the pathway. Both the RPP and RPG scores were subsequently normalized on a scale from 0 to 1 using the following formula:

$$z_i = (x_i - \min(x)) / (\max(x) - \min(x))$$

The final score for each gene-pathway combination was calculated using a weighted formula, with the RPP score contributing 20% and the RPG score contributing 80%. This weighting emphasizes the gene-specific contribution (RPG), reflecting its functional importance within the pathway, while incorporating the pathway-level significance (RPP) as a contextual factor. To consolidate all instances of a gene across PAAD-related pathways, the average of the final score for all instances was calculated to derive the gene's overall ranking score. This comprehensive approach ensures a balanced assessment of pathway-level and gene-specific contributions, effectively prioritizing genes that demonstrate significance within the most impactful pathways.

### 2.6. Expression analysis in normal and PAAD tissue samples

To further identify suitable drug targets after our initial AI-powered evaluation, we collected comprehensive expression data for the top 20 candidates. We examined expression levels in normal and PAAD tissue samples to assess differential expression patterns. This data was primarily sourced from the Human Protein Atlas, utilizing protein relative expression data from the Clinical Proteomic Tumor Analysis Consortium, which provides measurements in transcripts per million. The Human Protein Atlas was selected because it allows for comparison of expression levels between normal and tumor tissues, providing insights into cancer-specific expression dynamics.

Additionally, we examined gene expression throughout the progression of PAAD to assess expression patterns. For this purpose, we utilized data from UALCAN, an online bioinformatics platform that incorporates RNA-seq data from TCGA. UALCAN was chosen because it contains expression levels across different stages of PAAD, allowing for the exploration of SRC's clinical relevance in pancreatic cancer.[30]

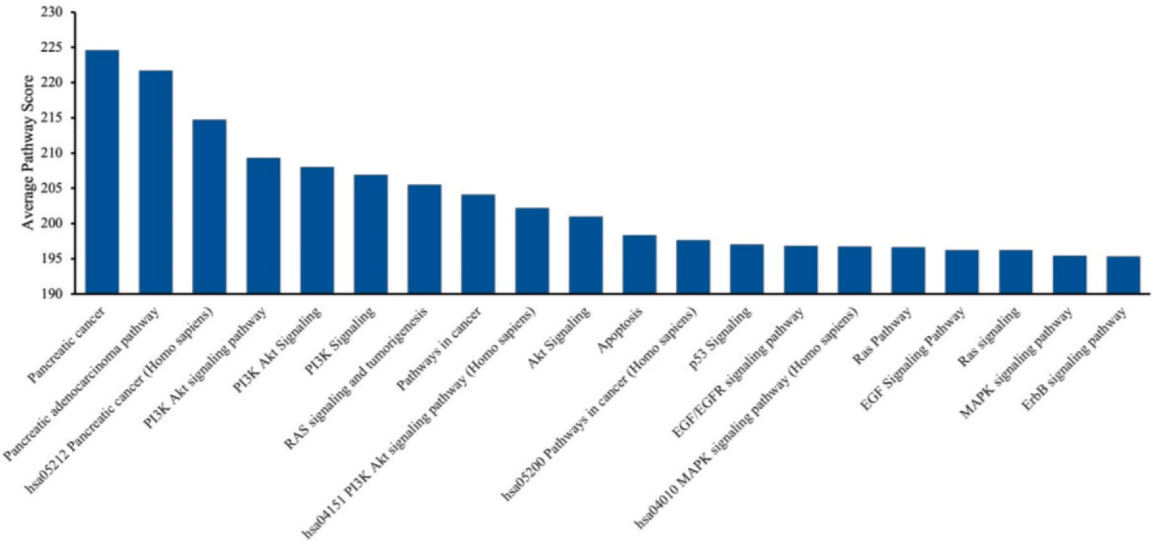



**Figure 2.** Top 20 pancreatic cancer–associated pathways ranked by average pathway score using our AI-assisted evaluation framework. Pathways were scored across seven categories using ChatGPT-4o with 20 iterations per pathway, and weighted following their importance in PDAC. All scores shown are the average composite scores after weighting. See Methods (Section 2.3) for scoring methodology and Supplementary File 2 for full score breakdowns.

### *2.7. Survival analysis using gene expression as a biomarker*

We utilized the Kaplan-Meier Plotter tool to conduct survival analyses of SRC in pancreatic cancer. Kaplan-Meier Plotter is an online tool that enables survival analysis using gene expression and clinical data from multiple cancer studies. This tool was selected for its ability to integrate datasets from diverse sources, including 12 Gene Expression Omnibus and 4 International Cancer Genome Consortium Data Portal cohorts, which provide a large, heterogeneous dataset ideal for robust survival analysis (detailed in Supplementary File 3). The platform's authors applied stringent preprocessing to ensure consistency and reduce redundancy. Specifically, duplicate measurements or multiple samples from the same tumor were removed, and gene annotations were standardized using the HGNC symbol checker. For genes measured by multiple probes, the probe with the highest total expression was retained. Expression values were quantile normalized (via the preprocessCore R package) and then scaled such that the average expression per sample equaled 1000. Only genes with expression data in $\geq$50% of samples were retained, and samples without clinical metadata were excluded. These steps ensured the final dataset was deduplicated, batch-adjusted, and clinically annotated for robust survival modeling. Our analysis included both overall survival and disease-free survival, focusing exclusively on samples with specified tumor stages (n = 473).

To determine the optimal threshold for distinguishing between low and high SRC expression groups, we used the "auto-select best cutoff" feature available in the Kaplan-Meier Plotter tool. This feature systematically evaluates all potential cutoff points within the interquartile range (between the 25th and 75th percentiles) of gene expression values. The cutoff that demonstrated the highest statistical significance (lowest *P* value) and the lowest false discovery rate was selected as the optimal threshold. In instances where multiple cutoffs showed equal significance, the threshold with the highest hazard ratio (HR) was chosen to ensure the strongest association with survival outcomes. These thresholds provided a robust basis for stratifying patients into low and high SRC expression groups for survival analysis.

However, certain limitations of this approach should be acknowledged. The cohort composition, drawn from publicly available datasets, may introduce biases, such as overrepresentation of specific demographics or treatment protocols that do not generalize across all patient populations. Additionally, while the "auto-select best cutoff" feature minimizes arbitrary threshold selection, it is prone to overfitting within the given dataset, which may limit reproducibility in independent cohorts. To address the potential overfitting associated with the "auto-select best cutoff" approach, we performed a sensitivity analysis using a range of predefined SRC expression thresholds spanning the 25th to 75th percentiles. As shown in Supplementary Figure 1, hazard ratios remained consistently $>1$ across all tested cutoffs, with the most significant associations observed between the 25th–35th percentiles. In contrast, median-based cutoffs did not reach statistical significance, suggesting that survival differences are more apparent at the lower-to-intermediate expression thresholds. This analysis supports the robustness of the prognostic association while mitigating concerns of overfitting. Finally, the analyses lack external validation using entirely independent datasets, which is critical for confirming the findings' generalizability.

## 3. Results

The results section presents a detailed analysis of pathways and genes associated with PAAD. By leveraging a robust scoring framework, this study aims to identify and rank pathways and candidate genes based on their relevance to PAAD progression and therapeutic potential. Key findings highlight the significance of specific pathways and genes in PAAD-related mechanisms, as described below.

### *3.1. Comparison of PAAD-related and unrelated pathways*

Figure 2 illustrates the top 20 pathways identified through our pathway evaluation, ranked by their average final scores. The top three pathways, "Pancreatic cancer" (224.6), which refers to the KEGG umbrella pathway that integrates multiple molecular alterations frequently observed in PDAC, "Pancreatic adenocarcinoma pathway" (221.7), and "hsa05212 Pancreatic cancer (Homo sapiens)" (214.7), underscoring the strong association of these pathways with PAAD progression and therapeutic relevance. These pathways integrate key signaling mechanisms, such as PI3K/Akt and MAPK signaling, which are also independently ranked highly in the pathway evaluation. The prominence of pancreatic cancer-specific pathways in the rankings validates the scoring framework and emphasizes the biological relevance of these pathways in driving PAAD progression, survival, and resistance mechanisms. A full list of pathways and their scores is provided in Supplementary File 1.

We conducted a two-sample independent *t*-test to compare the mean scores between 129 unrelated pathways (mean = 95.69, standard

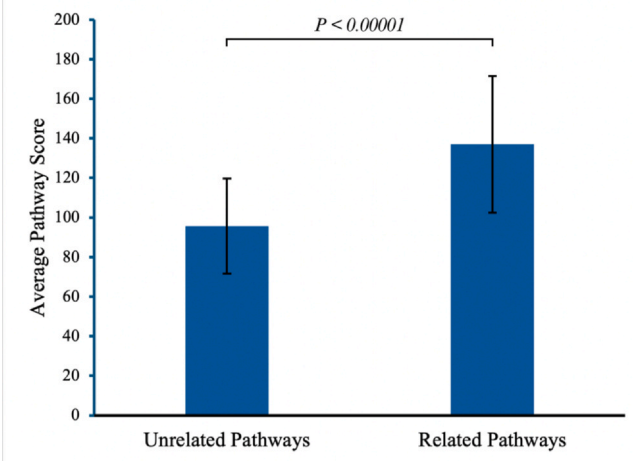



**Figure 3.** Comparison of PAAD-unrelated and related pathway scores. This bar chart compares the average pathway scores between pathways unrelated (n = 129) and related (n = 363) to PAAD. The average pathway score for PAAD-related pathways is significantly higher than that for unrelated pathways, with $P < 0.00001$, indicating a highly significant difference. Error bars represent standard deviation.

deviation [SD] = 22.67) and 363 PAAD-related pathways (mean = 137.00, SD = 36.48). The *t*-test yielded a significant difference between the two groups ($t(489) = -12.06$, $P < 0.00001$), indicating that the PAAD-related pathways had a significantly higher mean score compared to the unrelated pathways. The calculated pooled variance ($s_p^2$) was 1,117.2. Figure 3 illustrates the differences in mean scores and variability between the two groups, indicating the significantly higher scores for PAAD-related pathways. The full rankings of PAAD-related pathways are provided in Supplementary File 1, and unrelated pathways in Supplementary File 3.

To quantify the magnitude of this difference, we calculated Hedges' g statistic, which is an effect size that accounts for the two groups' unequal sample sizes ($N_{unrelated} = 129$, $N_{related} = 363$). Hedge's g statistic showed a large effect size of 1.24, indicating a greater relevance of PAAD-related pathways compared to unrelated pathways.

### 3.2. Ranking of candidate genes in PAAD-related pathways

Following the scoring and ranking process for the 252 candidate genes, the integration of RPP and RPG scores allowed for a comprehensive evaluation of gene significance within PAAD-related pathways. Figure 4 shows the top 20 candidates, representing the genes with the most substantial and consistent contributions to high-scoring pathways critical to PAAD progression and therapeutic resistance. A complete list of ranked genes is provided in Supplementary File 4.

The top-ranked gene, AURKB, achieved an average final score of 0.4336, indicating its critical involvement in high-priority pathways linked to PAAD progression. Other highly ranked genes include POLA1 (0.3900), RRM2 (0.3843), and TOP2A (0.3791), all of which play key roles in cell proliferation and genomic stability, processes heavily implicated in PAAD tumorigenesis. Additionally, several genes with less explored roles in PAAD, such as SSTR4 (0.3749), KCNH2 (0.3676), and SRC (0.3641), emerged as high-ranking candidates, suggesting novel avenues for future research.

The gene ranking results closely align with our pathway analyses, supporting the biological significance of these findings. For instance, EGFR, a highly ranked gene with an average final score of 0.3595, is part of the well-characterized PI3K/AKT signaling pathway, which ranked among the top pathways in our evaluation. Genes such as AKT1 (0.3428) and PIK3CA (0.3297) were closely tied to the top-ranking PI3K Akt signaling pathway (209.3), supporting their role in cell proliferation and survival. Furthermore, VEGFA (0.3480), critical for

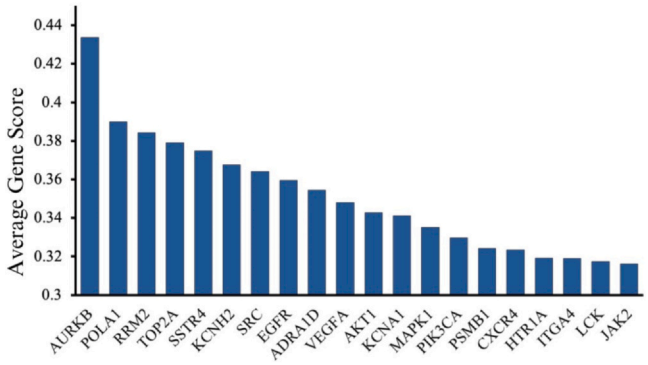



**Figure 4.** Top 20 candidate genes ranked by average final score. Each gene's final score is a weighted combination of two components: the RPP score, reflecting the average pathway scores from pathways in which the gene appears, and the RPG score, derived from the PAGER 3.0 gene prioritization index. These scores were normalized on a 0–1 scale and combined using an 80%:20% weighting scheme (RPG:RPP) to balance data-driven relevance with pathway-level context.

angiogenesis, aligns with pathways including Ras signaling (196.2) and MAPK signaling pathway (195.4), known to drive tumor progression. Additionally, MAPK1 (0.3351) strongly correlates with the MAPK signaling pathway, further reinforcing its role in oncogenic signaling. These associations underscore the integration of gene and pathway analyses in identifying biologically relevant targets for PAAD, providing a robust framework for prioritizing therapeutic interventions.

Sensitivity analysis confirmed that minor weight changes did not significantly impact pathway rankings. For example, a 10% reduction in the weight of pathway relevance shifted rankings by 3.87% (Supplementary File 5), demonstrating the reliability of the scoring framework.

### 3.3. Drug target candidate expression in normal and PAAD tumor samples

For each of the top 20 candidate genes from our ChatGPT-4o evaluation, we retrieved its protein expression in PAAD samples and normal samples from the Human Protein Atlas. Out of the 20 candidates, 4 showed a statistically significant difference in normalized Relative Protein Expression (nRPX) values between the two groups ($P < 0.0001$): SRC, ITGA4, TOP2A, and RRM2 (Figure 5).

Protein expression data from the Human Protein Atlas further supports SRC's critical role in PAAD. SRC exhibits a highly significant difference in protein expression between pancreatic cancer and normal tissues, with a $P$ value $< 0.0001$. This expression pattern is among the most pronounced of the evaluated candidates and highlights SRC's potential involvement in driving tumor progression and metastasis. By contrast, other clinically explored candidates, such as EGFR and VEGFA, show no significant expression differences ($P \geq 0.5$), indicating their limited disease-specific relevance. This combination of significant expression changes and relative novelty in clinical applications positions SRC as a target for further research.

Furthermore, we analyzed the trends of SRC expression in PAAD progression using data from TCGA. Figure 6A illustrates the expression of SRC across different tumor grades. SRC expression progressively increases from Grade 1 to Grade 3, with statistically significant differences observed between Grade 1 and Grade 2 ($P = 0.020$), and between Grade 1 and Grade 3 ($P = 0.036$). These findings suggest an upward trend in SRC expression across tumor grades, particularly in early-to-intermediate transitions, therefore indicating that SRC may be involved in tumor progression and could potentially serve as a marker of tumor aggressiveness. Figure 6B details SRC expression across early cancer stages. While SRC expression appeared visually higher in later stages, no statistically significant differences were observed between stages (all $P$ values $> 0.1$), suggesting the trend may not be robust. However, we

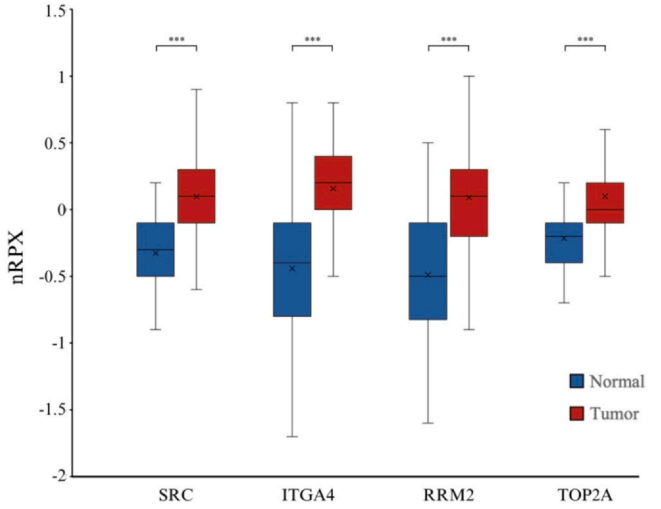



**Figure 5.** Genes with significant differential protein expression in normal vs. PAAD tissues. Box-and-whisker plots showing the normalized Relative Protein Expression (nRPX) values of tumor (red) and normal (blue) samples for SRC, ITGA4, RRM2, and TOP2A in tumor (red, n = 141) and normal (blue, n = 74) samples. The data reveal a statistically significant upregulation of protein expression in tumor samples compared to normal samples for all four genes, with $P$ values less than $1\times10^{-15}$, $7\times10^{-12}$, $7\times10^{-10}$, and $4\times10^{-8}$, respectively. The nRPX values were obtained from the Human Protein Atlas and calculated as log2(intensity), derived from global protein abundance measurements. These measurements were performed through mass spectrometry experiments using isobaric tandem mass tags by the Clinical Proteomic Tumor Analysis Consortium. *** denotes $P < 0.0001$.

acknowledge that caution should be exercised when interpreting these results given the limited number of normal and stage 1 tissues. This increasing trend of SRC expression with tumor grade and stage underlines its potential importance in PAAD pathogenesis and merits further investigation into its role as a therapeutic target.

### *3.4. Validation of expression level as biomarker in survival analysis*

The validation of SRC expression levels as a biomarker in survival analysis is presented through a series of survival curves derived from the Kaplan-Meier Plotter tool. Figure 7A shows the impact of SRC expression on overall survival, reinforcing the potential of SRC as an important factor in PAAD patient survival. The cutoff between high and low groups, optimized using the "auto-select best cutoff" feature of Kaplan-Meier Plotter, was set at 891. This panel illustrates a significant difference in survival between the groups, with an HR of 1.45 (95% confidence interval [CI]: 1.1–1.92, log-rank $P = 0.00769$).

Figure 7B further supports the relevance of SRC expression with a disease-free survival analysis, demonstrating a significant association between high SRC levels and more aggressive and recurrent forms of PAAD. The expression groups were defined using an optimized cutoff value of 1457. This survival curve shows an HR of 2.15 (95% CI: 1.47–3.16) and a log-rank $P$ of $9.19\times10^{-5}$.

## 4. Discussion

In this study, we developed a novel framework for prioritizing therapeutic targets in PAAD, focusing on biologically relevant pathways rather than relying solely on genetic mutation prevalence. This framework incorporates human-designed metrics, such as pathway relevance, target druggability, and biomarker availability, alongside AI-driven evaluations to ensure a systematic and reproducible assessment of potential drug targets. By integrating ChatGPT-4o into this process, we significantly enhanced the speed and consistency of pathway evaluations, leveraging AI's computational power to analyze complex datasets and streamline the discovery process. At the same time, we mitigated AI's inherent biases through iterative scoring, curated datasets, and human oversight.

Our framework emphasizes statistical rigor and transparency in scoring pathways. Each category, such as pathway relevance to PAAD or target druggability, was carefully weighted to reflect its importance in determining therapeutic potential. A key innovation in this framework is the use of ChatGPT-4o, which systematically evaluates pathways across multiple categories, accelerating the process of target identification. ChatGPT-4o enabled systematic evaluation of pathways by synthesizing diverse sources of information into actionable insights. Furthermore, we addressed potential biases in AI outputs by repeatedly scoring pathways, incorporating diverse datasets (e.g., TCGA, UALCAN), and ensuring human-designed metrics were the foundation of AI prompts. These measures helped make the framework less susceptible to biases arising from incomplete training data or model oversimplification.

The iterative nature of the scoring process allowed us to statistically validate the results, ensuring consistency and reproducibility.[31] For

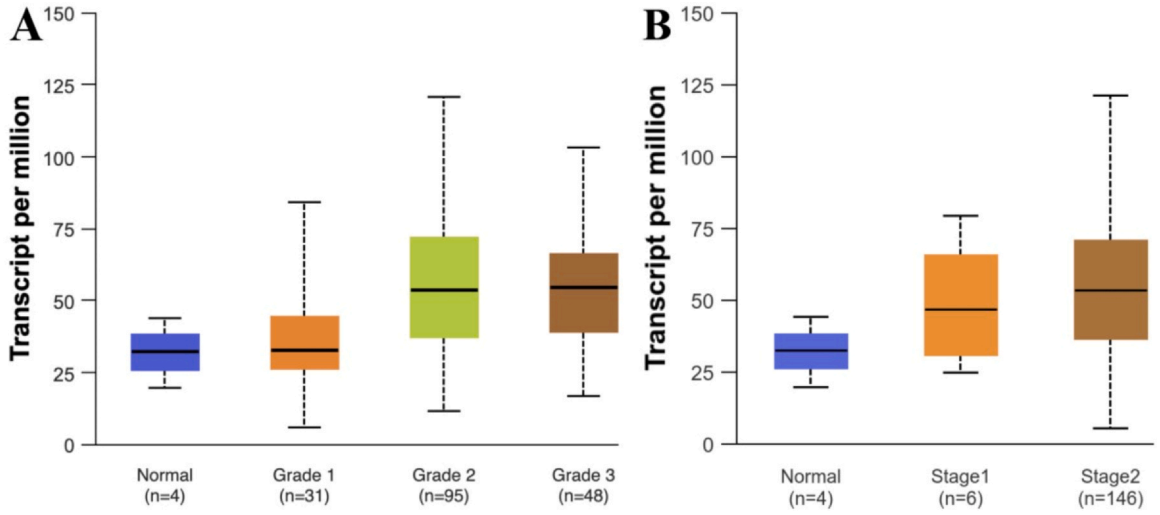



**Figure 6.** SRC expression trends in normal and PAAD tumor samples. This figure illustrates the variability in SRC expression measured in transcripts per million across normal and PAAD tumor tissues, using data sourced from UALCAN based on TCGA datasets. Figure 6**A**. Expression by tumor grade. Pairwise comparisons were performed, and significant differences were observed between Grade 1 vs. Grade 2 ($P = 0.020$), and Grade 1 vs. Grade 3 ($P = 0.036$). However, the difference between Grades 2 and 3 was not significant ($P = 0.918$). Figure 6**B**. Expression by individual cancer stage.

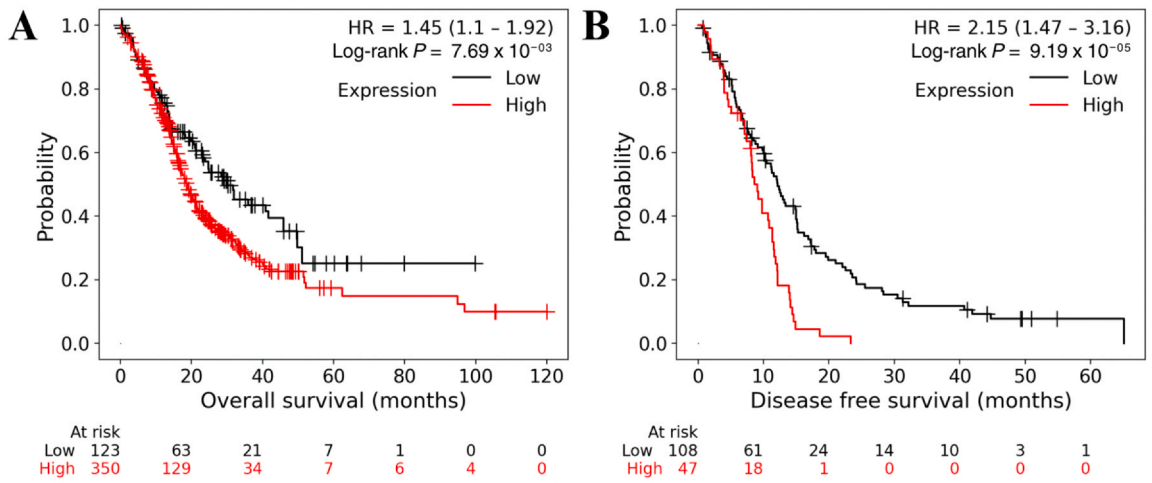



**Figure 7.** Effect of SRC expression level on PAAD patient survival. Hazard ratios and log-rank *P* values are indicated on the plots. These analyses were performed using the Kaplan–Meier Plotter platform. High and low expression groups were stratified by SRC expression using an optimized cutoff determined via maximally selected rank statistics. Figure 7**A**. Kaplan-Meier overall survival curve using a cohort size of 473. Figure 7**B**. Kaplan-Meier disease-free survival analysis for PAAD patients, performed on a sample size of 155.

example, multiple scoring iterations were performed to identify pathways with consistently high rankings. Cross-referencing ChatGPT-4o results with literature and experimental evidence further ensured accuracy, reducing the risk of AI-generated hallucinations.[32] This rigorous approach demonstrates the potential for AI to complement traditional bioinformatics methods in drug target discovery.

In comparison to established pathway and gene scoring systems such as GSEA, KEGG-based enrichment tools, PAGER's prioritization scores, and mutation-based ranking frameworks, our ChatGPT-4o-driven framework introduces a novel layer of interpretability and multi-criteria integration. Traditional systems including GSEA and PAGER primarily rely on statistical enrichment, gene overlap, and predefined pathway databases, which make them robust for high-throughput analysis but less informative when evaluating therapeutic relevance across clinical, pharmacological, and functional domains simultaneously. For instance, GSEA effectively identifies gene sets with differential expression but does not account for druggability, survival relevance, or biomarker availability. PAGER's influence score prioritizes gene-pathway relationships but is primarily based on gene connectivity and does not incorporate therapeutic resistance mechanisms or safety profiles.

In contrast, our AI-based framework evaluates each pathway using a predefined, weighted scoring rubric across seven biologically and clinically relevant dimensions: pathway relevance to PAAD, target druggability, mutation frequency, biomarker availability, resistance mechanisms, literature support, and safety. This design enables interpretability by producing subscores that reflect specific criteria, allowing researchers to assess which attributes most influence a pathway's ranking. Additionally, objectivity is supported by applying the same structured prompts and weights across all pathways, while reproducibility is enhanced through 20 scoring iterations per pathway and averaging to reduce stochastic prompt variability. Finally, the framework demonstrates strong integration with biological data—including datasets from TCGA, the Human Protein Atlas, and literature—allowing the AI to contextualize scores based on expression profiles, survival outcomes, and clinical relevance.

While conventional tools remain indispensable for statistical validation and early-stage screening, our method complements these approaches by enabling higher-order reasoning across datasets and clinical variables. We view this hybrid model, combining statistical rigor from traditional tools with contextual depth from AI, as a valuable advancement in therapeutic target prioritization, particularly in complex diseases like PAAD.

The integration of ChatGPT-4o into our evaluation framework underscores the practical utility of AI in systematically identifying and prioritizing drug targets in PAAD. ChatGPT-4o facilitated the efficient analysis of large, complex datasets, including TCGA and UALCAN, enabling a structured assessment of biologically relevant targets such as SRC.[33,34] This AI-driven approach allowed for rapid evaluation across predefined metrics, such as pathway relevance, druggability, and biomarker availability, substantially accelerating the target discovery process.

The AI-driven approach used in this study holds significant potential for application to other cancers and complex diseases. By integrating diverse datasets such as TCGA, UALCAN, and curated literature, alongside human-designed metrics, this framework provides a systematic and reproducible method for identifying therapeutic targets. Similar to PAAD, other malignancies such as colorectal, breast, or lung cancers involve complex molecular pathways and resistance mechanisms that can benefit from AI's ability to synthesize and analyze multidimensional data. Furthermore, the iterative scoring process, which evaluates pathway relevance, target druggability, and biomarker availability, can be adapted to reflect disease-specific priorities, such as immune checkpoint pathways in melanoma or angiogenesis in glioblastoma. Beyond oncology, the methodology could also be extended to other complex diseases, such as autoimmune disorders or neurodegenerative conditions, where multifactorial interactions and pathway-driven mechanisms play critical roles. This adaptability underscores the versatility of the AI-driven framework and its potential to accelerate target discovery and therapeutic innovation across diverse medical challenges.

However, the limitations of AI-driven methodologies must be acknowledged to contextualize their findings. The accuracy of AI outputs is inherently tied to the quality, diversity, and representativeness of its training data. While this study utilized curated datasets such as TCGA and UALCAN to reduce biases, the underrepresentation of specific pathways or populations—such as those of rare subtypes of PAAD or from non-Western demographics—may still influence results.[33] Furthermore, AI frameworks often simplify complex biological interactions.[35] For example, while the weighted scoring approach used here allowed for a structured evaluation of pathways, it did not fully capture the intricate interplay of tumor microenvironment factors, such as

stromal-epithelial signaling, or the dynamics of therapy-induced resistance mechanisms. This reveals the need for additional validation.

Human oversight played a critical role in mitigating these limitations. In this study, AI-generated findings were systematically cross-referenced with empirical data, such as survival analyses, and verified against high-impact literature to ensure their reliability.[32] For example, while ChatGPT-4o flagged SRC as a high-priority target, further validation through survival trends (Figures. 3 and 6) confirmed its association with decreased survival in both early and late tumor stages. Additionally, AI hallucinations, where plausible but incorrect outputs are generated, were reduced by anchoring the evaluation framework in established datasets and human-designed metrics.[34] This dual approach of leveraging AI efficiency while relying on human validation helped ensure the outputs were actionable and scientifically sound.

Ultimately, the application of AI in this study emphasizes the need for robust experimental validation and clinical trials to bridge the gap between computational predictions and actionable therapies.[16,36] For SRC, its established role in PAAD progression and the availability of existing inhibitors provide a strong starting point for clinical translation. However, addressing resistance mechanisms, such as SRC reactivation through compensatory signaling pathways, will require a detailed exploration of its downstream effects and combination strategies. As PAAD's genetic landscape evolves, ongoing research will be essential to optimize SRC-targeted therapies and validate their efficacy in diverse patient populations.

We recognize that SRC is not an unexplored target in PDAC. Multiple clinical trials involving SRC inhibitors such as dasatinib and saracatinib have reported limited efficacy, both as monotherapies and in combination with chemotherapeutics like gemcitabine. The modest clinical outcomes in these studies raise valid concerns about SRC's standalone translational potential. However, our findings suggest that these prior efforts may have been limited by a lack of patient stratification, insufficient biomarker integration, and incomplete understanding of SRC's role in advanced disease stages. Our framework addresses these gaps by re-evaluating SRC in the context of differential expression, survival impact, and pathway interconnectivity, highlighting its overexpression in high-grade tumors and association with poor prognosis. This multidimensional evidence suggests that while SRC alone may not be sufficient as a monotherapy target, it remains a promising candidate for inclusion in rational combination therapies, particularly when paired with agents targeting co-activated or compensatory pathways. Our AI-powered prioritization provides a refined, biologically contextual case for SRC's reconsideration as part of a precision medicine strategy, rather than as a universal or first-line monotherapy.

This study lays a foundation for future research, combining computational efficiency with experimental rigor to advance the development of impactful therapies for PAAD. While AI tools like ChatGPT-4o have demonstrated their utility in accelerating drug discovery processes, further refinements could enhance their performance. Integrating more diverse datasets, including data from underrepresented populations, rare cancer subtypes, and real-world clinical evidence, would improve the generalizability of AI-generated insights. Additionally, context-specific learning could be enhanced by training AI models on datasets enriched with cancer biology-specific annotations, enabling a more nuanced understanding of complex interactions within tumor microenvironments. Advances in explainable AI could also facilitate better interpretation of model outputs, providing researchers with actionable insights rather than opaque predictions. Lastly, incorporating AI tools into experimental workflows, such as optimizing experimental designs or predicting potential off-target effects of drugs, could further streamline and refine the drug discovery pipeline.

Our research also underscores the potential of SRC as a critical drug target for PAAD treatment. Through comprehensive evaluations using ChatGPT-4o, coupled with statistical analyses such as survival analyses and *t*-tests, we have established SRC's substantial role in influencing PAAD progression and metastasis. Notably, SRC's involvement in aggressive forms of PAAD, as indicated by survival trends across various tumor stages and grades (Figure 3), suggests its suitability as a therapeutic target, potentially in combination therapies for advanced PAAD.

Overall, we can derive the following key biological insights regarding SRC's role in PAAD. First, SRC is implicated in key processes driving tumor aggressiveness, including enhanced cell migration, proliferation, invasion, and resistance to apoptosis[37] by playing a significant role in pathways conducive to PAAD progression. Such pathways include RAS/RAF/MEK/ERK, PI3K/AKT, and JAK/STAT, which are pivotal for cell survival and tumor growth.[38–43] Second, SRC's frequent overexpression and activation in PAAD highlight its pivotal role despite gain-of-function mutations being relatively rare in PAAD patients (less than 2%).[44] Third, as a non-receptor tyrosine kinase part of the SRC family kinases, SRC interacts with key cellular components like receptor tyrosine kinases (e.g., MET, EGFR, PDGFR), which play a crucial role in several signaling pathways fostering PAAD growth and metastasis, including PI3K, STAT3, YAP, and FAK signaling.[45–47] Fourth, SRC's overexpression contributes significantly to PAAD pathology through mechanisms such as inflammation, angiogenesis, and desmoplasia, enhancing the tumor's aggressive characteristics.[45]

The identification of SRC as a high-priority therapeutic target carries substantial clinical implications for advancing PAAD treatment. SRC's correlation with poor survival outcomes demonstrates its potential as both a druggable target and a biomarker for stratifying patients and guiding therapeutic interventions. SRC's role in tumor progression, metastasis, and resistance to conventional therapies underscores its value in developing targeted strategies, particularly in advanced-stage PAAD where treatment options are limited. While survival trends suggest SRC's involvement at multiple stages of the disease, further studies using additional statistical models, such as the Cox proportional hazards model, could delineate its impact more precisely, incorporating tumor stage, expression levels, and interaction terms as covariates.

Although this study centers on SRC due to its distinct clinical and expression profile, several other high-ranking genes identified by our framework, such as AURKB, POLA1, and RRM2, also exhibit compelling biological relevance in PDAC. AURKB, for example, is a mitotic kinase associated with genomic instability and is overexpressed in several aggressive cancers. POLA1 and RRM2 are critical for DNA replication and repair and have also been implicated in chemoresistance. The emergence of these targets supports the concept that PDAC, given its marked heterogeneity and complex signaling environment, may benefit from multitargeted strategies rather than reliance on a single therapeutic agent.

Moreover, our pathway-level analysis revealed significant overlap among key oncogenic pathways such as PI3K/Akt, cell cycle regulation, and DNA damage response, suggesting that co-targeting parallel or compensatory pathways may improve treatment efficacy and prevent adaptive resistance. For example, combined inhibition of SRC and AURKB could concurrently disrupt invasive signaling and proliferative cycling, especially in tumors with high-grade features. Similarly, co-targeting DNA replication machinery (e.g., POLA1 or RRM2) with established chemotherapy could sensitize tumors to DNA-damaging agents.

These findings highlight the potential of our AI-driven framework not just for single target prioritization, but also for rational design of combination therapies informed by network biology. Future studies could build on this analysis by integrating synthetic lethality screens or modeling co-dependencies between top-ranked genes, ultimately enabling precision-guided polytherapeutic approaches for PDAC.

SRC inhibitors such as dasatinib, an FDA-approved drug for chronic myeloid leukemia and Philadelphia chromosome-positive acute lymphoblastic leukemia, represent a promising avenue for repurposing in PAAD. Dasatinib functions as a multi-kinase inhibitor targeting SRC family kinases and other tyrosine kinases involved in oncogenic

signaling. Preclinical studies have demonstrated that dasatinib effectively inhibits tumor cell migration, invasion, and proliferation, particularly in cancers characterized by SRC overactivation.[48] Its repurposing for PAAD offers an expedited path to clinical application, bypassing the early stages of drug development. Moreover, SRC inhibitors like dasatinib could play a critical role in overcoming therapy resistance by targeting SRC-driven compensatory pathways that contribute to tumor cell survival.[43,49] This could be particularly impactful in PAAD, where resistance to conventional therapies often undermines long-term treatment success.

Combining dasatinib with established chemotherapeutic regimens such as gemcitabine or FOLFIRINOX could further enhance therapeutic outcomes. Gemcitabine, a nucleoside analog and standard first-line treatment for advanced PAAD, targets rapidly dividing cells but is often limited by resistance mechanisms and short-lived responses.[50] SRC overactivation has been implicated in gemcitabine resistance by promoting pro-survival and repair pathways in tumor cells.[51] By co-administering dasatinib, SRC-driven resistance pathways could be disrupted, enhancing gemcitabine's efficacy and extending patient response duration.[52] This combination approach may also provide a means to counteract tumor heterogeneity, a hallmark of PAAD that contributes to variable treatment responses.[53,54] By integrating SRC inhibition into gemcitabine regimens, patients with tumors exhibiting high SRC activation could achieve better outcomes, effectively stratifying patient populations based on tumor biology.

Similarly, FOLFIRINOX, a combination of 5-FU, irinotecan, leucovorin, and oxaliplatin, could benefit from the addition of SRC inhibitors to mitigate resistance and improve outcomes in advanced-stage disease.[55] SRC inhibitors could disrupt resistance mechanisms inherent to this regimen, such as stromal-mediated drug resistance, while potentially amplifying the cytotoxic effects of the chemotherapeutic agents. The ability to stratify patients based on SRC activation levels or related biomarkers could further refine the use of FOLFIRINOX, ensuring that only those likely to benefit are exposed to this aggressive regimen. This approach aligns with precision medicine goals, where treatments are tailored to specific molecular and cellular characteristics of the tumor.

## 5. Conclusions

In conclusion, our integrated approach combining bioinformatics, AI-driven evaluations, and statistical analyses presents a potentially powerful model for advancing drug discovery in oncology. By focusing on biologically significant targets including SRC and EGFR, and leveraging advanced AI tools, we can streamline the drug development process, potentially bringing more effective therapies to clinical trials more swiftly. This strategy not only accelerates the identification of viable drug targets but also enhances our understanding of their roles in cancer biology, paving the way for more informed and strategic therapeutic interventions in PAAD.

## Ethical Statement

Hereby, Ayla Zhang and Dr. Jake Chen consciously assure that for "AI-Driven Network Biology Identifies SRC as a Therapeutic Target in Metastatic Pancreatic Adenocarcinoma" the following is fulfilled:

1) This material is the authors' own original work, which has not been previously published elsewhere.
2) The paper is not currently being considered for publication elsewhere.
3) The paper reflects the authors' own research and analysis in a truthful and complete manner.
4) The paper properly credits the meaningful contributions of co-authors and co-researchers.
5) The results are appropriately placed in the context of prior and existing research.
6) All sources used are properly disclosed (correct citation). Literally copying of text must be indicated as such by using quotation marks and giving proper reference.
7) All authors have been personally and actively involved in substantial work leading to the paper, and will take public responsibility for its content.

## CRediT authorship contribution statement

**Ayla Zhang:** Writing – original draft, Validation, Investigation, Formal analysis, Data curation. **Jake Y. Chen:** Supervision, Methodology, Conceptualization, Writing – review & editing, Visualization.

## Data availability

The original data and supplementary information presented in the study are openly available in Zenodo at https://doi.org/10.5281/zenodo.15676716.

## Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Ayla Zhang receives research training from Jake Y. Chen through AlphaMind Club, LLC, of which Chen is an owner. Jake Chen reports financial support was provided by National Institute of Health. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to improve the clarity and readability of the manuscript. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

## Acknowledgments

We thank Adrian Gu of AlphaMind Club for preparing part of the drug candidate dataset. This work is partly supported by the National Institute of Health Center grant awards (U54TR001005) to Dr. Jake Y. Chen, in which he serves as a co-investigator, a research start-up fund provided by the University of Alabama at Birmingham (UAB) to Dr. Chen and the UAB Systems Pharmacology AI Research Center grant to Dr. Chen. We also thank Dr. Xiang Ling for generously proofreading the statistics section for accuracy.

## References

1. Wang S, Zheng Y, Yang F, et al. The molecular biology of pancreatic adenocarcinoma: translational challenges and clinical perspectives. *Signal Transduct Target Ther.* 2021;6(1):249.
2. Rawla P, Sunkara T, Gaduputi V. Epidemiology of pancreatic cancer: Global trends, etiology and risk factors. *World J Oncol.* 2019;10(1):10–27.
3. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA Cancer J Clin.* 2024;74(1):12–49.
4. Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res.* 2014;74(11):2913–2921.
5. Hidalgo M, Cascinu S, Kleeff J, et al. Addressing the challenges of pancreatic cancer: Future directions for improving outcomes. *Pancreatology.* 2015;15(1):8–18.
6. Ho WJ, Jaffee EM, Zheng L. The tumour microenvironment in pancreatic cancer — clinical challenges and opportunities. *Nat Rev Clin Oncol.* 2020;17(9):527–540.
7. Provenzano PP, Cuevas C, Chang AE, Goel VK, Von Hoff DD, Hingorani SR. Enzymatic targeting of the stroma ablates physical barriers to treatment of pancreatic ductal adenocarcinoma. *Cancer Cell.* 2012;21(3):418–429.
8. Morrison AH, Byrne KT, Vonderheide RH. Immunotherapy and prevention of pancreatic cancer. *Trends Cancer.* 2018;4(6):418–428.

9. Luo J. KRAS mutation in Pancreatic Cancer. *Semin Oncol.* 2021;48(1):10–18.
10. Strickler JH, Satake H, George TJ, et al. Sotorasib in KRAS p.G12C–Mutated Advanced Pancreatic Cancer. *N Engl J Med.* 2023;388(1):33–43.
11. Yousef A, Yousef M, Chowdhury S, et al. Impact of KRAS mutations and co-mutations on clinical outcomes in pancreatic ductal adenocarcinoma. *Npj Precis Oncol.* 2024;8(1):1–13.
12. McFarland CD, Yaglom JA, Wojtkowiak JW, et al. The Damaging Effect of Passenger Mutations on Cancer Progression. *Cancer Res.* 2017;77(18):4763–4772.
13. Wodarz D, Newell AC, Komarova NL. Passenger mutations can accelerate tumour suppressor gene inactivation in cancer evolution. *J R Soc Interface.* 2018;15(143):20170967.
14. Kolbeinsson HM, Chandana S, Wright GP, Chung M. Pancreatic Cancer: A Review of Current Treatment and Novel Therapies. *J Invest Surg.* 2023;36(1):2129884.
15. Gordon ER, Wright CA, James M, Cooper SJ. Transcriptomic and functional analysis of ANGPTL4 overexpression in pancreatic cancer nominates targets that reverse chemoresistance. *BMC Cancer.* 2023;23(1):524.
16. Du X, Yi X, Zou X, et al. PCDH1, a poor prognostic biomarker and potential target for pancreatic adenocarcinoma metastatic therapy. *BMC Cancer.* 2023;23(1):1102.
17. Dave T, Athaluri SA, Singh S. ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. *Front Artif Intell.* 2023;6:1169595.
18. Yue Z, Kshirsagar MM, Nguyen T, et al. PAGER: constructing PAGs and new PAG–PAG relationships for network biology. *Bioinformatics.* 2015;31(12):i250–i257.
19. Chowbina S, Deng Y, Ai J, et al. A new approach to construct pathway connected networks and its application in dose responsive gene expression profiles of rat liver regulated by 2,4DNT. *BMC Genomics.* 2010;11(Suppl 3):S4.
20. Sun L, Li HM, Seufferheld MJ, et al. Systems-Scale Analysis Reveals Pathways Involved in Cellular Response to Methamphetamine. *PLOS ONE.* 2011;6(4):e18215.
21. Yue Z, Arora I, Zhang EY, Laufer V, Bridges SL, Chen JY. Repositioning drugs by targeting network modules: a Parkinson's disease case study. *BMC Bioinformatics.* 2017;18(Suppl 14):532.
22. Zhang F, Chen JY. Discovery of pathway biomarkers from coupled proteomics and systems biology methods. *BMC Genom.* 2010;11(2):S12.
23. Lin CC, Akuhata-Huntington Z, Hsu CW. Comparing ChatGPT's ability to rate the degree of stereotypes and the consistency of stereotype attribution with those of medical students in New Zealand in developing a similarity rating test: a methodological study. *J Educ Eval Health Prof.* 2023;20:17.
24. Walker HL, Ghani S, Kuemmerli C, et al. Reliability of Medical Information Provided by ChatGPT: Assessment Against Clinical Guidelines and Patient Information Quality Instrument. *J Med Internet Res.* 2023;25:e47479.
25. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA.* 2005;102(43):15545–15550.
26. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res.* 2021;49(D1):D545–D551.
27. Zack T, Lehman E, Suzgun M, et al. Assessing the potential of GPT-4 to perpetuate racial and gender biases in health care: a model evaluation study. *Lancet Digit Health.* 2024;6(1):e12–e22.
28. Wang ZP, Bhandary P, Wang Y, Moore JH. Using GPT-4 to write a scientific review article: a pilot evaluation study. *BioData Min.* 2024;17(1):16.
29. Hirosawa T, Harada Y, Mizuta K, Sakamoto T, Tokumasu K, Shimizu T. Evaluating ChatGPT-4's Accuracy in Identifying Final Diagnoses Within Differential Diagnoses Compared With Those of Physicians: Experimental Study for Diagnostic Cases. *JMIR Form Res.* 2024;8:e59267.
30. Chandrashekar DS, Bashel B, Balasubramanya SAH, et al. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia N Y N.* 2017;19(8):649–658.
31. Kumar A, Aelgani V, Vohra R, et al. Artificial intelligence bias in medical system designs: a systematic review. *Multimed Tools Appl.* 2024;83(6):18005–18057.
32. Salvagno M, Taccone FS, Gerli AG. Artificial intelligence hallucinations. *Crit Care.* 2023;27(1):180.
33. Kurapati SS, Yaghy A, Shukla AG. Data bias: ethical considerations for understanding diversity in medical artificial intelligence. *AI and Ethics.* 2025;5(3):3399–3405.
34. Nazer LH, Zatarah R, Waldrip S, et al. Bias in artificial intelligence algorithms and recommendations for mitigation. *PLOS Digit Health.* 2023;2(6):e0000278.
35. Xie T, Huang A, Yan H, Ju X, Xiang L, Yuan J. Artificial intelligence: illuminating the depths of the tumor microenvironment. *J Transl Med.* 2024;22(1):799.
36. Patne AY, Dhulipala SM, Lawless W, Prakash S, Mohapatra SS, Mohapatra S. Drug Discovery in the Age of Artificial Intelligence: Transformative Target-Based Approaches. *Int J Mol Sci.* 2024;25(22):12233.
37. Nagaraj NS, Smith JJ, Revetta F, Washington MK, Merchant NB. Targeted Inhibition of Src Kinase Signaling Attenuates Pancreatic Tumorigenesis. *Mol Cancer Ther.* 2010;9(8):23222332.
38. Agell L, Hernández S, Salido M, et al. PI3K signaling pathway is activated by PIK3CA mRNA overexpression and copy gain in prostate tumors, but PIK3CA, BRAF, KRAS and AKT1 mutations are infrequent events. *Mod Pathol.* 2011;24(3):443–452.
39. Horn RDV, Chu S, Fan L, et al. Cdk1 Activity Is Required for Mitotic Activation of Aurora A during G2/M Transition of Human Cells. *J Biol Chem.* 2010;285(28):21849–21857.
40. Ma X, Zhang L, Song J, et al. Characterization of the Src-regulated kinome identifies SGK1 as a key mediator of Src-induced transformation. *Nat Commun.* 2019;10(1):296.
41. Parkin A, Man J, Timpson P, Pajic M. Targeting the complexity of Src signalling in the tumour microenvironment of pancreatic cancer: from mechanism to therapy. *Febs J.* 2019;286(18):3510–3539.
42. Parsons SJ, Parsons JT. Src family kinases, key regulators of signal transduction. *Oncogene.* 2004;23(48):7906–7909.
43. Roskoski Jr R. Src protein-tyrosine kinase structure, mechanism, and small molecule inhibitors. *Pharmacol Res.* 2015;94:9–25.
44. Alcalá S, Mayoral-Varo V, Ruiz-Cañas L, et al. Targeting SRC Kinase Signaling in Pancreatic Cancer Stem Cells. *Int J Mol Sci.* 2020;21(20):7437.
45. Poh AR, Ernst M. Functional roles of SRC signaling in pancreatic cancer: Recent insights provide novel therapeutic opportunities. *Oncogene.* 2023;42(22):1786–1801.
46. Pelaz SG, Tabernero A. Src: coordinating metabolism in cancer. *Oncogene.* 2022;41(45):4917–4928.
47. Rozengurt E, Eibl G. Crosstalk between KRAS, SRC and YAP Signaling in Pancreatic Cancer: Interactions Leading to Aggressive Disease and Drug Resistance. *Cancers.* 2021;13(20):5126.
48. Montero JC, Seoane S, Ocaña A, Pandiella A. Inhibition of Src Family Kinases and Receptor Tyrosine Kinases by Dasatinib: Possible Combinations in Solid Tumors. *Clin Cancer Res.* 2011;17(17):5546–5552.
49. Leclercq G, Haegel H, Schneider A, et al. Src/lck inhibitor dasatinib reversibly switches off cytokine release and T cell cytotoxicity following stimulation with T cell bispecific antibodies. *J Immunother Cancer.* 2021;9(7):e002582.
50. Klein AP. Pancreatic cancer epidemiology: understanding the role of lifestyle and inherited risk factors. *Nat Rev Gastroenterol Hepatol.* 2021;18(7):493–502.
51. Duxbury MS, Ito H, Zinner MJ, Ashley SW, Whang EE. Inhibition of Src Tyrosine Kinase Impairs Inherent and Acquired Gemcitabine Resistance in Human Pancreatic Adenocarcinoma Cells. *Clin Cancer Res.* 2004;10(7):2307–2318.
52. Bartscht T, Rosien B, Rades D, et al. Dasatinib blocks transcriptional and promigratory responses to transforming growth factor-beta in pancreatic adenocarcinoma cells through inhibition of Smad signalling: implications for in vivo mode of action. *Mol Cancer.* 2015;14(1):199.
53. Grünwald BT, Devisme A, Andrieux G, et al. Spatially confined sub-tumor microenvironments in pancreatic cancer. *Cell.* 2021;184(22):5577–5592.e18.
54. Sherman MH, Beatty GL. Tumor Microenvironment in Pancreatic Cancer Pathogenesis and Therapeutic Resistance. *Annu Rev Pathol.* 2023;18:123–148.
55. Zhao S, Zhang H, Yang N, Yang J. A narrative review about CDK4/6 inhibitors in the setting of drug resistance: updates on biomarkers and therapeutic strategies in breast cancer. *Transl Cancer Res.* 2023;12(6):1617–1634.