Article

Identification of Cancer Associated Fibroblasts Related Genes Signature to Facilitate Improved Prediction of Prognosis and Responses to Therapy in Patients with Pancreatic Cancer

Yong Zhou [1], Yanxi Lu [1], Franziska Czubayko [1], Jisheng Chen [1], Shuwen Zheng [1], Huaqing Mo [1], Rui Liu [1], Georg F. Weber [1,2,3], Robert Grützmann [1,2,3], Christian Pilarsky [1,*] and Paul David [1,*]

[1] Department of Surgery, University Hospital Erlangen, Friedrich-Alexander Universität Erlangen-Nürnberg (FAU), 91054 Erlangen, Germany; yong.zhou@fau.de (Y.Z.); yanxi.lu@uk-erlangen.de (Y.L.); franziska.czubayko@uk-erlangen.de (F.C.); jisheng.chen@extern.uk-erlangen.de (J.C.); shuwen.zheng@extern.uk-erlangen.de (S.Z.); huaqing.mo@fau.de (H.M.); rui.liu@fau.de (R.L.); georg.weber@uk-erlangen.de (G.F.W.); robert.gruetzmann@uk-erlangen.de (R.G.)

[2] Deutsches Zentrum für Immuntherapie, Universitätsklinikum Erlangen, Friedrich-Alexander Universität Erlangen-Nürnberg (FAU), 91054 Erlangen, Germany

[3] Bavarian Cancer Research Center (BZKF), 91052 Erlangen, Germany

* Correspondence: christian.pilarsky@uk-erlangen.de (C.P.); paul.david@uk-erlangen.de (P.D.)

Abstract: Pancreatic cancer (PC) is highly aggressive, with a 5-year survival rate of 12.8%, making early detection vital. However, non-specific symptoms and precursor lesions complicate diagnosis. Existing tools for the early detection of PC are limited. CAFs are crucial in cancer progression, invasion, and metastasis, yet their role in PC is poorly understood. This study analyzes mRNA data from PC samples to identify CAF-related genes and drugs for PC treatment using algorithms like EPIC, xCell, MCP-counter, and TIDE to quantify CAF infiltration. Weighted gene co-expression network analysis (WGCNA) identified 26 hub genes. Our analyses revealed eight prognostic genes, leading to establishing a six-gene model for assessing prognosis. Correlation analysis showed that the CAF risk score correlates with CAF infiltration and related markers. We also identified six potential drugs, observing significant differences between high-CAF and low-CAF risk groups. High CAF risk scores were associated with lower responses to immunotherapy and higher tumor mutation burdens. GSEA indicated that these scores are enriched in tumor microenvironment pathways. In summary, these six model genes can predict overall survival and responses to chemotherapy and immunotherapy for pancreatic cancer, offering valuable insights for future clinical strategies.

Keywords: pancreatic cancer; cancer associated fibroblast; WGCNA; prognosis; chemotherapy; immunotherapy; molecular docking

1. Introduction

Pancreatic cancer (PC) is one of the most aggressive and deadly cancers, with current treatment options largely ineffective. It ranks among the fifteen most common cancers globally, particularly affecting Western countries like Europe and North America [1]. PC is characterized by a desmoplastic reaction, leading to a dense, fibrous tumor stroma that creates high intertumoral pressure. This physical barrier hampers therapy by compressing blood vessels and contributing to chemoresistance [2]. The tumor microenvironment (TME) is crucial in cancer progression, affecting metastasis and treatment responses. It consists of

various cells, extracellular matrix components, and signaling molecules that promote tumor growth and therapeutic resistance. Notably, cancer-associated fibroblasts (CAFs) comprise over 70% of the TME and actively interact with cancer cells and immune components. Understanding this complex relationship is essential for addressing challenges in cancer treatment [3–5]. Activated pancreatic stellate cells (PSCs) also contribute significantly to PDAC, creating a complex TME that promotes growth, invasiveness, and immunosuppression [6]. Single-cell RNA sequencing reveals that about 35% of cells in PDAC tumors are cancerous, while 26% are fibroblasts and PSCs. Despite similarities across studies, there is substantial variability in cellular distribution among patients [7–11]. This heterogeneity increases the likelihood of therapy failure in clinical trials and underscores the need for more sophisticated models during the in vitro research stage [12].

Cancer-associated fibroblasts (CAFs) are known to affect the TME through remodeling the extracellular matrix (ECM) and facilitating neoplastic cell invasion. Their interaction with cancer cells and other stromal components happens via secreting growth factors, cytokines, and chemokines, thereby shaping the TME and influencing tumor progression [13]. Identifying specific fibroblast markers has enabled the live cell sorting of CAF subpopulations, facilitating in-depth in vivo mechanistic studies [14]. Mechanistically, CAFs promote tumor migration through the mechanisms of contraction and proteolysis that remodel the ECM [13]. CAFs release substantial amounts of growth factors and proinflammatory cytokines, such as transforming growth factor-β (TGF-β), interleukin-6 (IL-6), and CC-chemokine ligand 2 (CCL2), which recruit immune cells, particularly immunosuppressive cells, into the tumor stroma, promoting immune evasion [15,16]. CAFs contribute to chemotherapy resistance through multiple mechanisms, such as CAFs upregulating glutathione levels in prostate cancer and inhibit reactive oxygen species (ROS) production [17], facilitating natural immunity to RAF inhibitors in cancers with BRAF mutations [18], expressing *CD10* and *GPR77* exhibit intrinsic resistance to chemotherapy [14], and regulating the NF-κB signaling pathway via secreting interleukin-6 (IL-6) and interleukin-8 (IL-8) in CD10$^+$GPR77$^+$ CAFs.

Weighted gene co-expression network analysis (WGCNA) [19] is a systems biology approach that directly identifies gene clusters (modules) based on their co-expression patterns across multiple samples [20]. This method facilitates the detection of highly correlated gene modules, summarizes them using module eigengenes or intramodular hub genes, and establishes relationships between modules and external sample traits through eigengene network analysis. Additionally, WGCNA allows for the assessment of module membership, providing insights into gene significance within specific biological contexts. It has been widely used to identify key biomarkers in various cancers. For instance, it has facilitated the discovery of Epstein–Barr virus (EBV) markers in gastric cancer [21], *CENPM* as a metastasis-related gene in adrenocortical carcinoma [22], and prognostic as well as diagnostic biomarkers for colorectal cancer [23]. It has also been applied to identify circulating microRNA signatures for the early detection of pancreaticobiliary cancer [24] and to explore the role of *HLF* in anaplastic thyroid cancer immunotherapy [25]. Similarly, CAF markers have been identified and utilized in multiple malignancies, including breast cancer [26], gastric cancer [27], intrahepatic cholangiocarcinoma [28], high-grade serous ovarian carcinoma [29], colorectal cancer progression [30], bladder cancer [31], and prostate cancer [27] among others.

This study aims to identify a gene signature associated with cancer-associated fibroblasts to enhance the prediction of prognosis and therapeutic responses in patients diagnosed with pancreatic cancer.

2. Results

2.1. Higher CAF Infiltrations Are Related to Worse Overall Survival in Pancreatic Cancer

The synopsis of the workflow is illustrated in Figure 1A. A total of 185 pancreatic cancer samples from TCGA-PAAD and 182 tumor tissues merged from GSE183795 and GSE78229 were analyzed, with 112 non-tumor and non-survival data samples being excluded. Four algorithms were evaluated: EPIC, xCell, MCPcounter, and the estimated algorithm. The results of the CAF infiltration and stromal score are shown in Supplementary Table S1 (TCGA-PAAD) and Supplementary Table S2 (GSE183795 and GSE78229). Subsequently, the Kaplan–Meier curve was utilized to analyze the association between high and low CAF scores with overall survival (OS), as depicted in (Figure 1B,C; Supplementary Figure S1 and Table S3). The weighted gene co-expression network analysis (WGCNA) was performed based on the MCPcounter scores. The data from GSE183795 and GSE78229 and TCGA-PAAD were utilized, and genes were ranked according to their fold change. The top 5000 most differentially expressed genes (DEGs) were retained from the GEO dataset. The optimal soft threshold power (sft) was determined to be 8 (scale-free $R2 = 0.996$) (Figure 2A), while the TCGA exhibited a value of 6 (scale-free $R2 = 0.972$) (Figure 2B). The topology plot (GSE183795 and GSE78229) (Supplementary Figure S2A) and TCGA-PAAD (Supplementary Figure S2B) demonstrate the rationality of our subsequent clustering. Concurrently, the clustering of module eigengenes yielded a cut-off value < 0.2 in GSE183795 and GSE78229 (Figure 2C) and TCGA-PAAD (Figure 2D). The construction of the co-expression network [32] from the dendrogram-module color map reveals 19 co-expressions in GSE183795 and GSE78229 (Supplementary Figure S3A) and 20 co-expression modules in TCGA-PAAD (Supplementary Figure S3B). Module traits display that the strongest positive correlation with the CAF proportion (Cor = 0.92, $p = 6 \times 10^{-73}$ and stromal score (Cor = 0.91, $p = 1 \times 10^{-71}$) was the magenta module in GSE183795 and GSE78229 (Figure 2E). In contrast, the brown module exhibited the strongest positive correlation with the CAF proportion (Cor = 0.95, $p = 1 \times 10^{-92}$) and stromal score (Cor = 0.79, $p = 8 \times 10^{-39}$) was brown in TCGA-PAAD (Figure 2F). Following the identification of modules with the most significant relationship, magenta hub genes were selected from GSE183795 and GSE78229 (Supplementary Table S4, Sheet S1) and brown (Supplementary Table S4, Sheet S2) from TCGA-PAAD. The GEO-magenta module's scatter plot results demonstrated a substantial correlation between gene significance (GS) and module membership (MM) in CAF (Cor = 0.93, $p = 7.2 \times 10^{-106}$) (Figure 2G) and stromal scores (Cor = 0.92, $p = 3.3 \times 10^{-99}$) (Figure 2H). A similar correlation was observed in the TCGA-brown module, MM and GS for CAF (Cor = 0.98, $p < 1 \times 10^{-200}$) (Figure 2I) and stromal scores (Cor = 0.89, $p = 2.7 \times 10^{-171}$) (Figure 2J) demonstrated a high degree of correlation. Utilizing gene Significance filter >0.4 and model Significance filter >0.8 as benchmark standards, 55 genes in the GSE183795 and GSE78229 black model and 195 in the TCGA-PAAD brown module were identified as hub genes, exhibiting a high correlation with CAF and stromal scores.

2.2. Intersection Genes Enrichment Related to ECM

The Venn diagram was used to visualize the intersection of two hub gene sets, which resulted in the identification of 26 genes (Figure 3A, Supplementary Table S4, Sheet S3). Subsequently, gene ontology (GO) analysis was performed on these 26 genes. The analysis revealed that the regulation of neuron projection development exhibited significant enrichment in the biological process (BP) terms. The collagen-containing extracellular matrix and extracellular matrix structural constituents were identified as the primary enriched cell component (CC) and molecular function (MF) terms, respectively, as illustrated by the barplot (Supplementary Figure S4A) and bubble plot (Supplementary Figure S4C).

Additionally, KEGG pathway enrichment analysis revealed significant enrichment in the cytoskeleton in muscle cells (Supplementary Figure S4B,D).
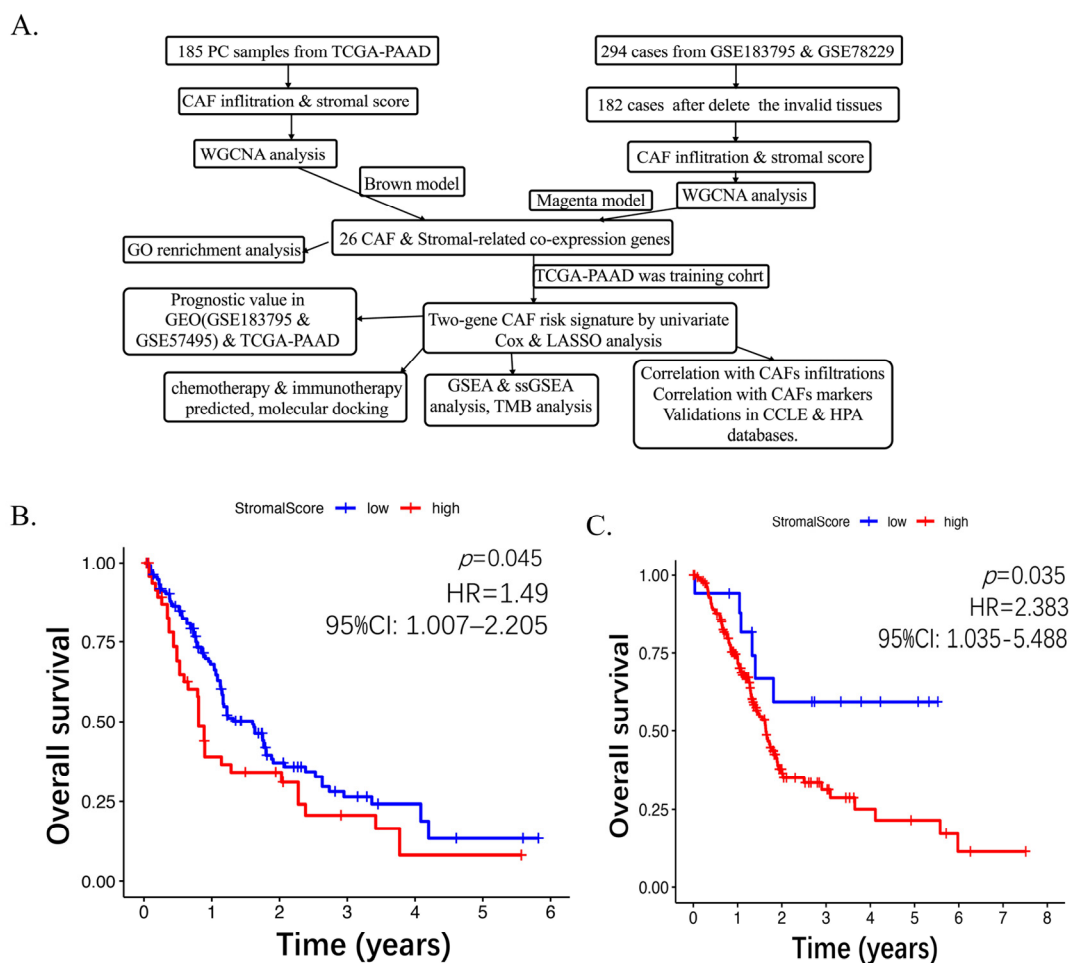

Figure 1. Workflow of the research on TCGA-PAAD and GSE183795 and GSE78229 (**A**). Overall survival analysis by log-rank. K-M curves show the pancreatic cancer patients with higher stromal-score related to bad overall survival in GSE183795 and GSE78229 (**B**) and TCGA-PAAD (**C**).

2.3. Establishing the Prognostic Risk Model Genes from the Stromal CAF-Based

A total of 185 cases from TCGA-PAAD were utilized as the training cohort due to the more substantial data available, while 182 pancreatic cancer samples from the GEO dataset (GSE183795 and GSE78229) were employed as the validation group. Eight core genes were screened from the 26 intersection genes, resulting in the creation of a favorable survival-related forest plot (Figure 3B). Subsequently, a univariate Cox regression on 26 shared hub genes, resulting in the identification of eight genes with an OS-related (*p*-value < 0.05). The coefficient profiles of these genes were then analyzed using the least absolute shrinkage and selection operator (LASSO) Cox regression (Supplementary Figure S4E). The adjustment factor (lambda), as depicted in Supplementary Figure S4F, was determined using partial likelihood variance through tenfold cross-validation. To create the CAF risk model, we identified six specific genes: the CAF risk score = *FSTL1* expression $* (-0.051) + CTSK$ expression $* (-0.93) + PRRX1$ expression $* 0.44 + GFPT2$ expression $* (-0.14) + MMP2$ expression $* 0.68 + MFAP5$ expression $* 0.13$ (Supplementary Table S4, Sheet S4).
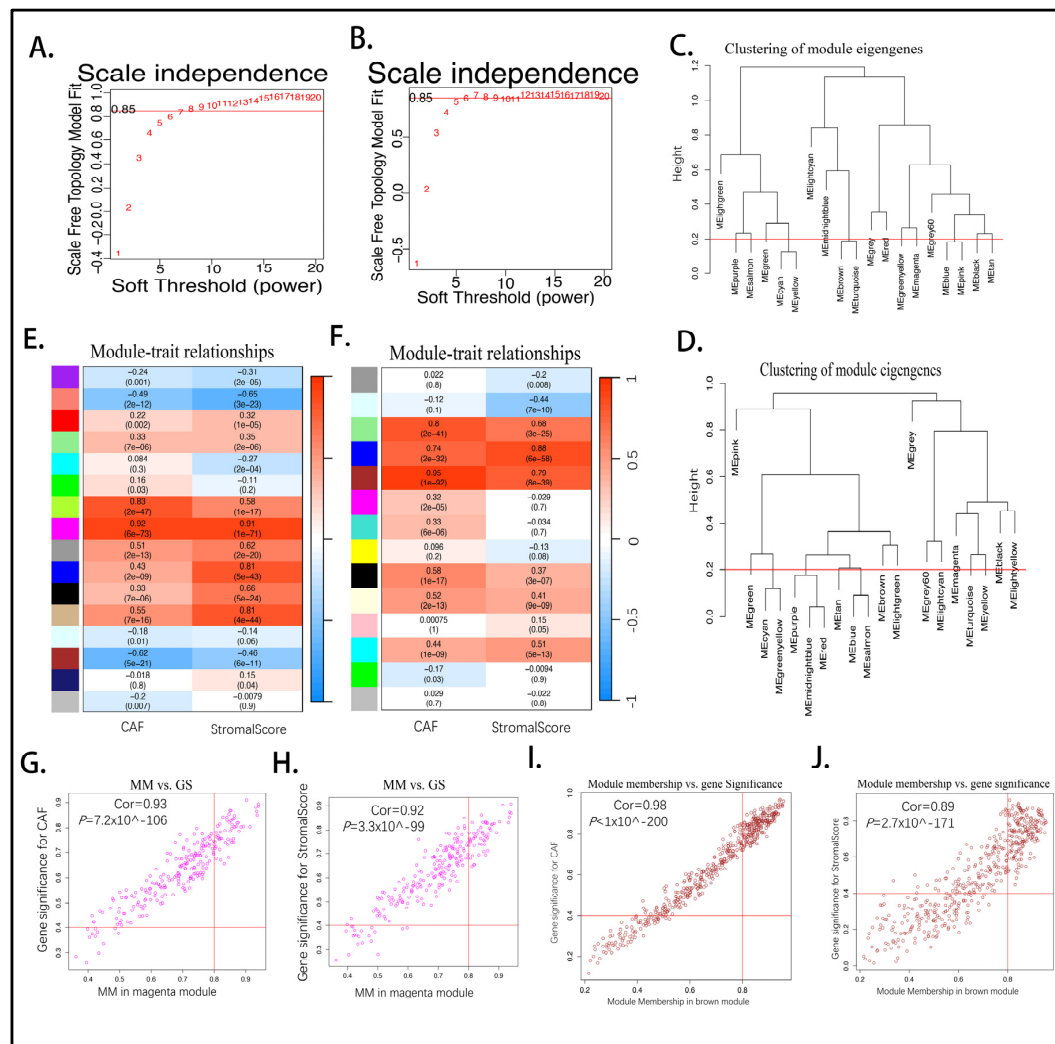
Figure 2. Co-express network of the WGCNA. Based on the scale topology tree in GSE183795 and GSE78229 (**A**) and TCGA-PAAD (**B**), the soft threshold power was decided. The cluster gene tree shows the depth of cut needed <0.2 displayed in GSE183795 and GSE78229 (**C**) and TCGA-PAAD (**D**). Module-trait relationships show different module colors' relation values in GSE183795 and GSE78229 (**E**) and TCGA-PAAD (**F**). Scatter graphs of the module membership and gene significance of each gene in the magenta module in GSE183795 and GSE78229 for CAF (**G**) and for StromalScore (**H**); the brown module in brown TCGA-PAAD for CAF (**I**) and for StromalScore (**J**) (MM: module membership, GS: gene significance). The horizontal axis is the correlation between gene and co-expression modules, and the vertical axis is about the gene and phenotype.

2.4. High CAF Infiltrations and Markers Can Be Shown from CAF Signature Genes

To enhance the reliability of the CAF model as a forecasting tool, Spearman's correlation analysis was conducted to establish a relationship between CAF risk and stromal score, as well as CAF abundances, as indicated by EPIC, alongside xCell, MCP-counter, and TIDE. It was observed that there was a positive correlation between the CAF risk score and the multi-estimated CAF infiltrations along with the stromal score in both GSE183795 and GSE78229 (Figure 4A) and TCGA-PAAD (Figure 4B) groups. Furthermore, the heatmap, which revealed the expression patterns of CAF markers (Supplementary Table S4, Sheet S5), identified six CAF module genes associated with the CAF risk score in GSE183795 and GSE78229 (Figure 4C) and TCGA-PAAD (Figure 4D) cohorts. Additionally, we observed a high and positive correlation between the CAF risk

score and the expression levels of the two genes and the collected CAF markers in both GSE183795 and GSE78229 (Figure 4E) and TCGA-PAAD (Figure 4F) cohorts.
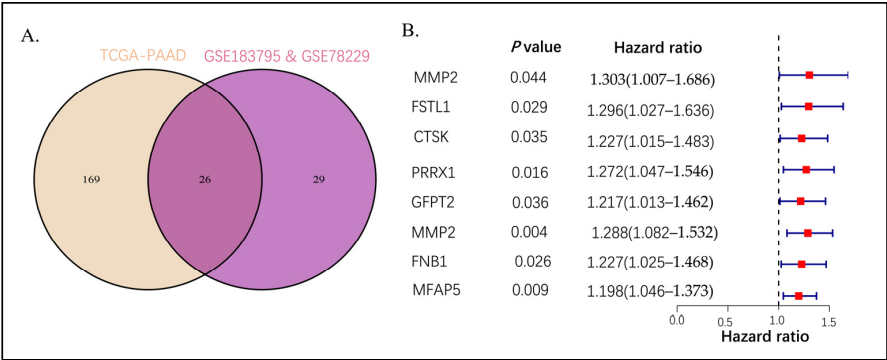

Figure 3. (**A**) The Venn diagram presented the intersection of TCGA-PAAD brown and GSE183795 and GSE78229 magenta module genes. (**B**) Univariate Cox analysis for the screening of overall survival-associated genes in TCGA-PAAD.
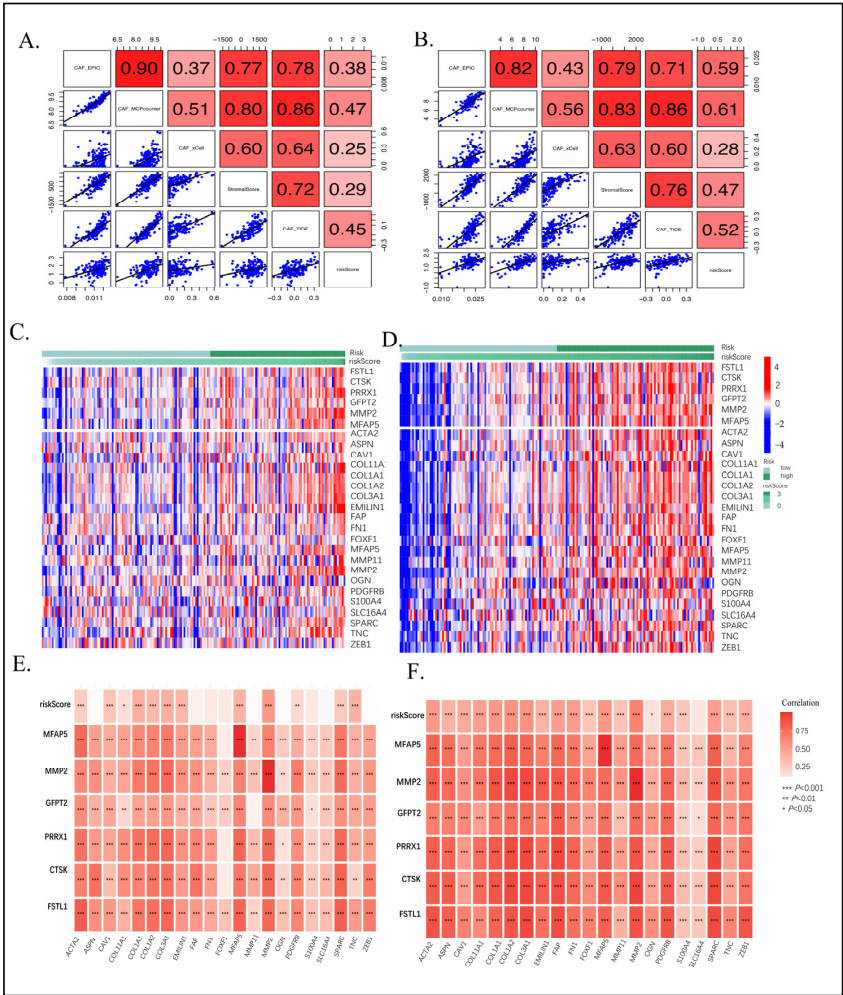

Figure 4. Spearman's correlation analyses revealed the CAF risk score was strongly and positively correlated with stromal scores and multi-estimated CAF infiltrations in GSE183795 and GSE78229 (**A**) and TCGA-PAAD (**B**) cohorts. The heatmap revealing the expression patterns of CAF markers identified two CAF genes with the CAF risk score in GSE183795 and GSE78229 (**C**) and TCGA-PAAD (**D**) cohorts. The CAF risk score and two signature genes were positively correlated with the literature that reported CAF markers in GSE183795 and GSE78229 (**E**) and TCGA-PAAD (**F**) cohorts. (* means $p < 0.05$, ** means $p < 0.01$, *** means $p < 0.001$).

2.5. Chemo-Immuno-Therapy Sensitivity Prediction and Molecular Docking

Pancreatic cancer is recognized as the leading cancer within the digestive system. Surgical intervention often yields suboptimal outcomes, underscoring the necessity for adjuvant treatment modalities, such as chemotherapy and immunotherapy, to enhance clinical outcomes. In this study, we employed the GDSC online database (accessed on 2 March 2025) to screen chemotherapy drugs based on their association with CAF-risk groups. This analysis identified 13 sensitivity drugs (Supplementary Table S5, Sheet S1) that satisfied the $p < 0.001$ in high-/low-risk groups in TCGA-PAAD and 30 in GSE183795 and GSE78229 (Supplementary Table S5, Sheet S2). Six intersection drugs, including staurosporine, dasatinib, OTX015, BMS-536924, Luminespib, and IGF1R, were shown in both the GSE183795 and GSE78229 (Figure 5A) and TCGA-PAAD cohorts (Figure 5B). Immunotherapy using immune checkpoint inhibitors has generated optimism among PC patients. To assess whether the CAF risk score could serve as an immunotherapy predictor for pancreatic cancer patients, we employed the TIDE method. In Figure 5C, bar charts showed that the low-CAF group exhibited a higher immunotherapy response (58%) than the higher-CAF group (30%), which was significant ($p < 0.001$) in GSE183795 and GSE78229. The low-CAF group exhibited a higher immune-therapy response (58%) than the higher-CAF group (21%), which is significant ($p < 0.001$) in TCGA-PAAD. In Figure 5D, violin plots illustrate that the low-risk group receiving a TIDE score, and the high-risk group receiving a higher score between GSE183795, GSE78229, and TCGA-PAAD. Additionally, the AUC figures of 0.725 (95%CI: 0.650–0.797) in GSE183795 and GSE78229 (Figure 5E). Furthermore, the AUC figures of 0.773 (95% CI: 0.696–0.844) in TCGA-PAAD (Figure 5F) demonstrated the superior efficacy of our CAF model in forecasting immunotherapy responses. The 3D structure of the chemical drugs was obtained from PubChem (https://pubchem.ncbi.nlm.nih.gov/) (accessed on 2 March 2025), and the 3D structure of the module genes was obtained from the Protein Data Bank (https://www.rcsb.org/) (accessed on 2 March 2025). Utilizing the CB-DOCK2 online platform (http://cadd.labshare.cn/cb-dock2/index.php) (accessed on 2 March 2025). The Vina score of these interactions was found below (Table 1), indicating robust binding characteristics. We show their 3D binding spatial structure of the least binding energies between the gene and drugs, including MMP2 with OTX015 (Figure 6A), FSTL1 with OTX015 (Figure 6B), GFPT2 with Luminespib (Figure 6C), and CTSK with staurosporine (Figure 6D). The potential drugs for expanding the drug repertoire for pancreatic cancer are enhanced by these findings.

Table 1. Molecular docking result of available protein and drugs.

| Target	PDB ID	Drug	PubChem ID	Binding Energy (kcal/mol)
MMP2	1RTG	Staurosporine	44259	−7
		Dasatinib	3062316	−8.2
		OTX015	9936746	−8.4
		BMS-536924	135440466	−8
		Luminespib	135539077	−8.1
FSTL1	6JZA	Staurosporine	44259	−7.4
		Dasatinib	3062316	−7.3
		OTX015	9936746	−7.6
		BMS-536924	135440466	−7
		Luminespib	135539077	−7
GFPT2	7NUT	Staurosporine	44259	−8
		Dasatinib	3062316	−8.5
		OTX015	9936746	−8.5
		BMS-536924	135440466	−9
		Luminespib	135539077	−9.1
CTSK	8V57	Staurosporine	44259	−9
		Dasatinib	3062316	−7.6
		OTX015	9936746	−8.5
		BMS-536924	135440466	−7.9
		Luminespib	135539077	−7.5
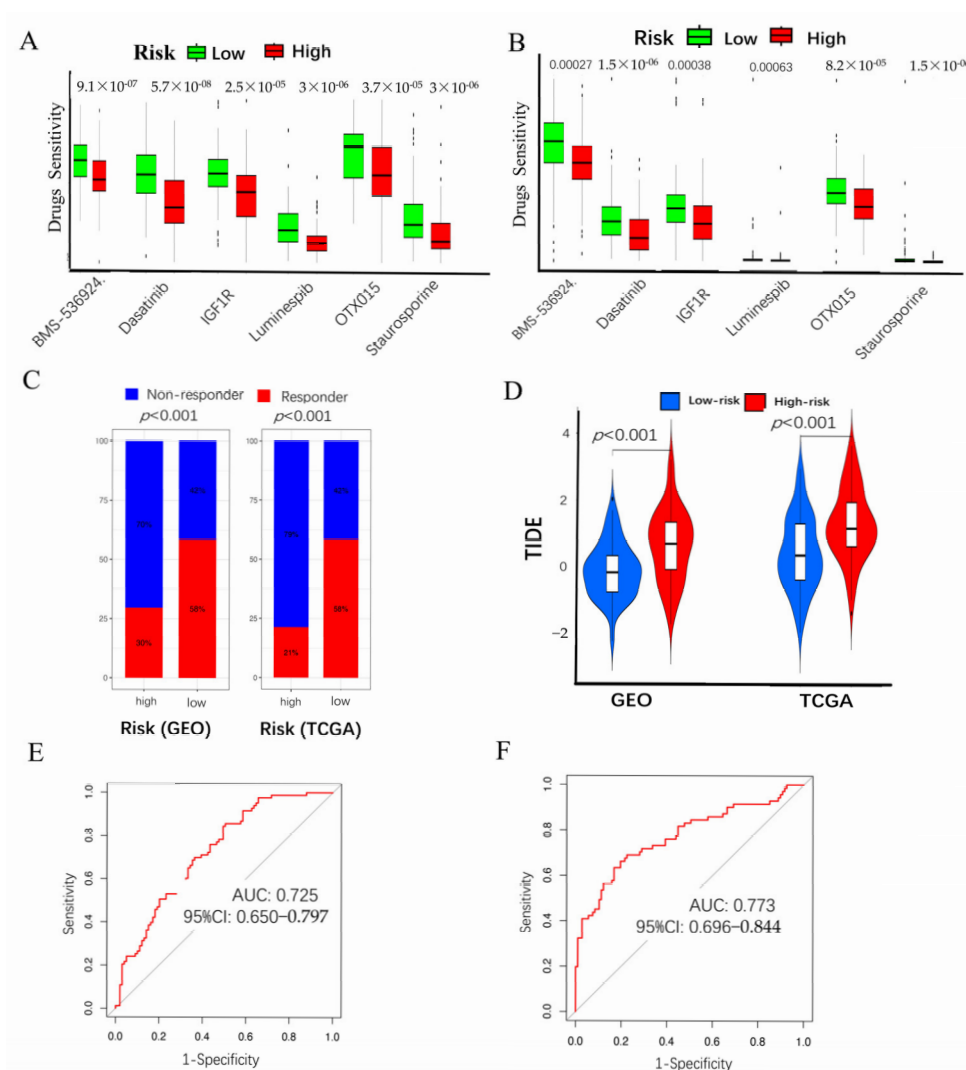
Figure 5. Sensitivity drugs: about the GSE183795 and GSE78229 (**A**) and TCGA-PAAD (**B**) including Staurosporine, Dasatinib, OTX015, BMS-536924, Luminespib, IGF1R. Immunotherapy responses of the high-/low-risk groups based on TIDE scores. The box plots (**C**) in GSE183795 and GSE78229 and TCGA-PAAD, violin plots (**D**) in GSE183795 and GSE78229 and TCGA-PAAD, and the curve showing the sensitivity of different methods in GSE183795 and GSE78229 (**E**) and TCGA-PAAD (**F**).

2.6. High CAF Risk Is Correlated with Tumor Mutation Burden

After this observation, it was noted that there is a discrepancy in tumor mutation burden between the CAF high-risk group (83.5%, Supplementary Figure S5A) and low-risk group (81.7%, Supplementary Figure S5B). The relationship between high- and low-risk groups and TMB is demonstrated in Supplementary Figure S5C. Spearman's correlation analysis revealed a relationship between CAF-risk score and TMB in Supplementary Figure S5D. The correlation between different score algorithms and TMB is illustrated in Supplementary Figure S5E.
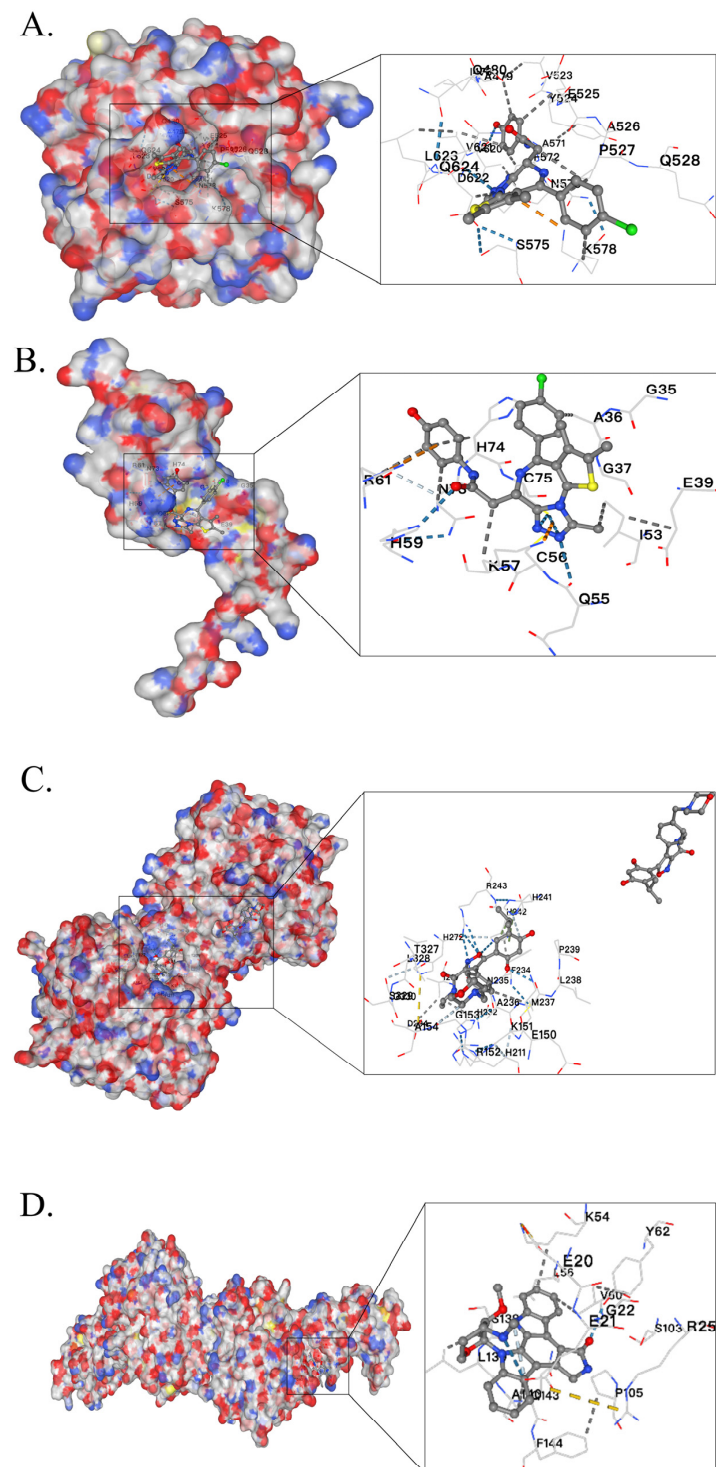
Figure 6. Molecular docking graphs MMP2 binding with OTX015 (**A**), FSTL1 with OTX015 (**B**), GFPT2 with Luminespib (**C**), and CTSK with Staurosporine (**D**).

2.7. GSEA and ssGSEA of the Six Genes of the CAF Signature

Gene Set Enrichment Analysis (GSEA) was utilized to compare high and low CAF-infiltration pancreatic cancer in GSE183795 and GSE78229 cohorts, as well as the TCGA-PAAD cohorts. The high-risk group in GEO datasets (Figure 7A) and TCGA datasets (Supplementary Figure S6A) exhibited significant enrichment of genes belonging to the extracellular matrix (ECM) structural constituent category, along with a concomitant association with tumor evasion and immigration. This finding underscores the pivotal role of CAFs in modulating the tumor microenvironment. In contrast, the low-risk group in GEO datasets

(Figure 7B) and TCGA datasets (Supplementary Figure S6B) demonstrated a predominant enrichment in cell metabolism and signal transduction pathways, which are more intimately linked to basic cellular function. Furthermore, single-sample gene set enrichment analysis (ssGSEA) revealed consistent enrichment of five Gene Ontology (GO) biological processes: bone morphogenesis (GEO datasets in Figure 7C, TCGA dataset in Supplementary Figure S6C), cell fate specification involved in pattern specification (GEO datasets in Figure 7D, TCGA dataset in Supplementary Figure S6D), embryonic neurocranium morphogenesis (GEO datasets in Figure 7E, TCGA dataset in Supplementary Figure S6E), and fibroblast growth factor receptor signaling pathway (GEO datasets in Figure 7F, TCGA dataset in Supplementary Figure S6F), mesenchyme development (GEO datasets in Figure 7G, TCGA dataset in Supplementary Figure S6G). The correlation coefficients (R^2) for both datasets ranged over 0.5, and the $p < 2 \times 10^{-13}$ GSE183795 and GSE78229, and TCGA-PAAD. These findings further support the valuable role of CAFs in ECM components and tumor invasion and immigration in pancreatic cancer. In a similar vein, the KEGG pathway analysis revealed an enrichment of cytokine receptor and extracellular matrix reshape and invasion in the high-risk group in GSE183795 and GSE78229 (Supplementary Figure S7A) and TCGA-PAAD (Supplementary Figure S8A). In contrast, the low-risk groups exhibited a significant association with basic cell biophysical processes in GSE183795 and GSE78229 (Supplementary Figure S7B) and TCGA-PAAD (Supplementary Figure S8B).
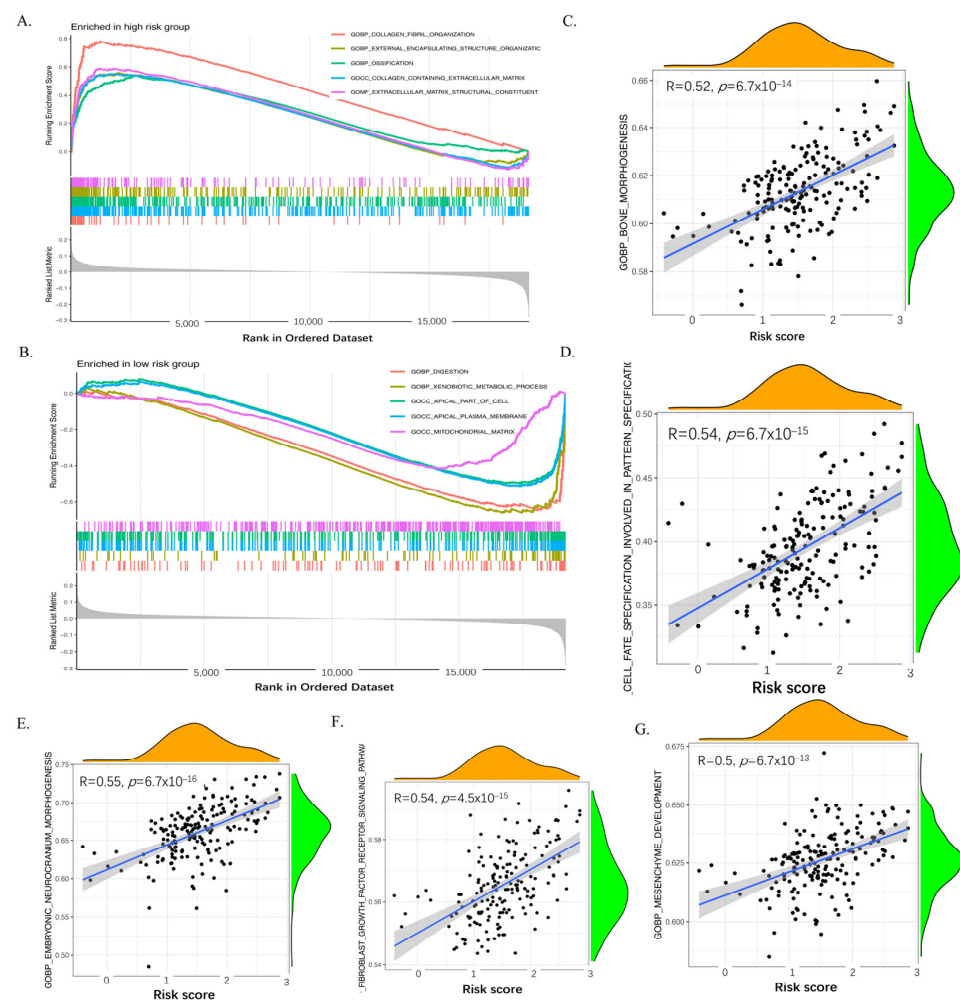

Figure 7. GSEA and ssGSEA enriched GO biological function in GEO datasets. Gene set enrichment analysis (GSEA) sets between high-CAF risk groups (**A**) and low risk group (**B**) in GSE183795 and GSE78229. ssGSEA results showed that the CAF risk score was positively correlated with bone morphogenesis (**C**),

cell fate specification involved in pattern (**D**), embryonic neurocranium morphogenesis (**E**), fibroblast growth factor receptor signaling pathway (**F**), and mesenchyme development (**G**) in GSE183795 and GSE78229.

2.8. Validation of Key Genes in CCLE and HPA Databases

Utilizing the CCLE database, it was confirmed that fibroblast cell lines exhibited elevated mRNA levels of the six module genes (GFPT2, MFAP5, CTSK, MMP2, FSTL1, and PRRX1) (accessed on 2 March 2025) compared to pancreatic cancer cell lines (Figure 8A). The Wilcoxon analysis (Figure 8B) shows the FSTL1 ($p < 2.22 \times 10^{-16}$), CTSK ($p < 2.22 \times 10^{-16}$), PRRX1 ($p < 2.22 \times 10^{-16}$), GFPT2 ($p = 0.00015$), MMP2 ($p < 2.22 \times 10^{-16}$), and MFAP5 (4.1×10^{-13}) genes exhibited higher expression in fibroblasts than in the pancreas. To further substantiate the protein expression traits of six CAF signature genes, the HPA database's (accessed on 2 March 2025) IHC images were obtained to illustrate higher expression in CAFs than in pancreatic cancer (Figure 8C). These confirmations suggested the possibility of these genes being markers specific to CAFs.
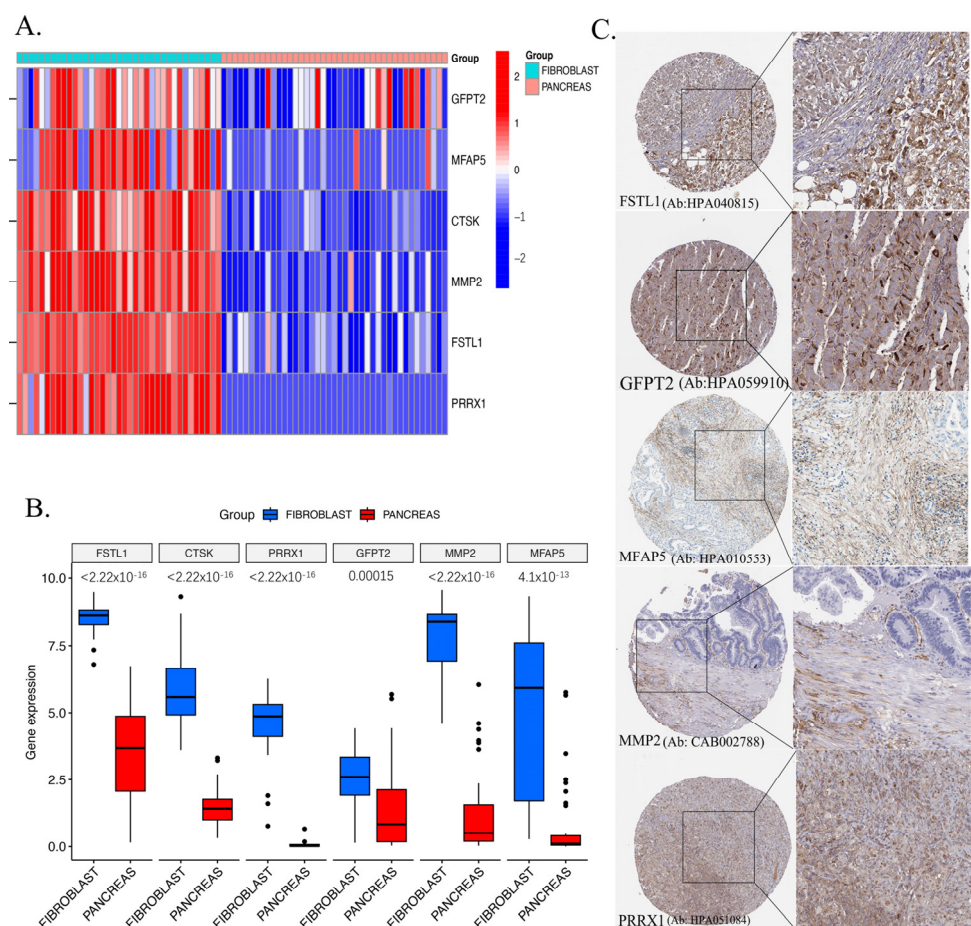

Figure 8. The mRNA expression levels of the two CAF genes in the fibroblasts and pancreatic cancer cell lines were illustrated in the heat map (**A**) and compared by Wilcoxon analysis (**B**). Protein expression in pancreatic cancer tissues (**C**).

3. Discussion

Cancer-associated fibroblasts (CAFs), as the predominant component of the tumor microenvironment (TME), play a vital role in pancreatic cancer, influencing tumor initiation, metastasis, drug sensitivity, immune evasion, and prognosis [33]. Similarly, our research

found that a high CAF score was associated with poor overall survival. To the best of our knowledge, our research is pioneering in methodically disclosing the function of CAFs in pancreatic cancer through WGCNA co-expression analysis. As illustrated in the workflow (Figure 1A), two cohorts were analyzed: GSE183795 and GSE78229 from GEO and TCGA-PAAD from TCGA. Utilizing sophisticated computational algorithms, including WGCNA, univariate Cox regression, Spearman's correlation, and LASSO regression, we identified six targeted genes (GFPT2, MFAP5, CTSK, MMP2, FSTL1, and PRRX1) that exhibited a strong overall survival (HR > 1). These findings substantiate the validity and robustness of the CAF-associated model genes in pancreatic cancer. The GSDC database (accessed on 2 March 2025) was then queried to identify six intersection drugs that exhibited higher sensitivity in the high-risk groups compared to the low-risk group. The 3D structures of four gene proteins and five drugs were obtained from a designated website. Subsequent molecular docking analyses revealed that binding of these drugs was highly stable, as evidenced by their binding structures and energies. Notably, GFPT2 demonstrated the lowest binding energy of −9.1 kcal/mol when bound to Luminespib, which may be a potential drug for high GFPT2 expression PC patients. The application of the online TIDE algorithm revealed that the group with higher TIDE scores exhibited an increased rate of non-responsiveness to immunotherapy in both the GSE183795 and GSE78229 and TCGA-PAAD groups. The ROC analysis further validated the robustness of our findings.

In the high CAF-infiltration group, GSEA revealed enrichment in extracellular matrix (ECM) structural constituents and related to tumor evasion and immigration, thereby highlighting the role of CAFs in modulating the tumor microenvironment. Conversely, the low CAF-infiltration group exhibited enrichment in cell metabolism and signal transduction, which are more connected to basic function. Furthermore, the ssGSEA findings revealed a direct relationship between epithelial–mesenchymal transition (EMT) and ECM, which plays a part in the tumor microenvironment in pancreatic cancer. CAFs reshape and secrete the ECM [34]. CAF-mediated collagen-induced ECM alterations depend not solely on the binding and activating discoidin domain receptors (DDRs) [35]. The enhanced expression of DDR1 and DDR2 collagen receptors in melanoma cells was observed when grown on 3D decellularized ECM from melanoma-related CAFs, in contrast to matrices from regular dermal fibroblasts, highlighting the tumor-enhancing function of CAF-derived ECM [36]. In addition to the receptor activation, CAFs secrete matrix-remodeling enzymes, such as lysyl oxidase (LOX), which induce ECM stiffness, thereby promoting cancer cell proliferation, invasion, and epithelial-to-mesenchymal transition (EMT) in oral squamous cell carcinoma (SCC) in vitro [37] and facilitating gastric cancer-derived liver metastasis in vivo [38]. Furthermore, CAFs regulate the mechanical properties of ECM, including mechanical forces [39], stiffness [40], rigidity [41], shear stress, and viscosity [42,43]. These biomechanical changes can further drive tumor progression and therapy resistance. Additionally, CAF-derived ECM contributes to macrophage infiltration, influencing tumor matrix remodeling and poor prognosis in ovarian cancer patients [44]. Furthermore, CAF-derived matrix proteins have been associated with a desmoplastic ECM signature, which is characteristic of aggressive tumor phenotypes [45]. In addition, the impact of CAFs and ECM components on dendritic cell (DC)-mediated immune responses in various cancers [46]. These components can regulate DC phenotypes, thereby inducing immune tolerance and consequently, facilitating immune evasion [47].

We assessed CAF infiltration in pancreatic cancer to ensure robustness and prevent overfitting. The EPIC method built the model, while the xCell, MCP-counter, and TIDE methods validated it. Our study confirms the established role of CAFs as crucial components of the tumor microenvironment (TME), influencing various aspects of tumor behavior, such as initiation, metastasis, and immune evasion. Similar studies, like those

by Kaluri et.al. [48] have highlighted the importance of CAFs in tumor progression across different cancers, reinforcing our observations. The CCLE dataset revealed elevated expression of six module genes (GFPT2, MFAP5, CTSK, MMP2, FSTL1, and PRRX1) associated with poor overall survival in patients with pancreatic cancer. This aligns with findings from other studies that have also discovered specific gene signatures linked to CAF activity and patient prognosis. For instance, a recent study by Yu et al., [49] also demonstrated CAF gene signatures in pancreatic cancer correlating with survival outcomes.

MMP2, a matrix metalloproteinase, is a promising early marker for PDAC. It plays a role in tissue remodeling, inflammation, and cancer progression [50–62].

PRRX1, a member of the paired-related homeobox family of transcription factors, has been implicated in the development of acinar-ductal metaplasia (ADM) and regeneration. Additionally, it impacts gemcitabine sensitivity via its role in epithelial–mesenchymal transition (EMT) [63–65]. Our findings confirm that PRRX1 is more expressed in CAFs than in normal pancreatic tissue, and its suppression may improve chemotherapy effectiveness.

MFAP5 is associated with poor prognosis in cancers like ovarian and intrahepatic cholangiocarcinoma, and high MFAP5 expression in CAFs correlates with unfavorable outcomes in pancreatic cancer [66,67].

We utilized molecular docking to evaluate drug interactions with target proteins, particularly highlighting GFPT2's significant binding with Luminespib as a potential therapeutic option. This technique has been similarly employed in studies such as that by Chen et al. [68] underlining the utility of computational methods in understanding drug efficacy associated with specific biomarkers. Our method of assessing CAF infiltration using tools such as EPIC and xCell, corroborated by the TIDE method, is further supported by methodologies in other studies that also aim to validate CAF-related models, such as Zhu et al. [68], reinforcing the rigor of our analytical approach.

In summary, while our research contributes novel insights into the functions of CAFs in pancreatic cancer and establishes a specific gene signature, it also aligns well with the existing literature, highlighting a growing consensus on the multifaceted roles of CAFs in tumor biology and therapy resistance. These comparisons not only validate our findings but also enhance the understanding of CAFs in the context of cancer research as a whole.

Limitations: Despite the confidence in the functionality of pancreatic cancer genes and the docking of target genes with drug molecules, several limitations must be acknowledged. Firstly, the present bioinformatic analysis is retrospective, drawn from two publicly available gene expression datasets, and lacking multi-center data. Second, there is a deficiency in real-world study data concerning CAFs in pancreatic cancer, and the CAF marker has not been verified at the molecular, cellular, and animal levels. Nonetheless, our study offers novel insights into the relationship between CAF infiltration and pancreatic cancer.

4. Materials and Methods

4.1. Accumulating and Processing Data

Data on RNA sequencing and clinical traits of patients with pancreatic cancer were obtained from two groups for thorough examination. The fragments per kilobase of transcript per million mapped reads (FPKM) formatted RNA-seq data and corresponding clinical prognostic records (including overall survival time and vital status) for 185 pancreatic adenocarcinoma (TCGA-PAAD) samples were retrieved from The Cancer Genome Atlas (TCGA) Genomic Data Commons (GDC) portal (https://portal.gdc.cancer.gov/ (accessed on 2 March 2025), with data freeze date on 10 February 2025. Tumor mutation burden (TMB) data were simultaneously acquired through the TCGA Mutation Annotation Format

(MAF) files. Normalized FPKM values were converted to transcripts per million (TPM) using the following formula:

$$TPM = FPKM \times 10^6 / \sum FPKM$$

Followed by log2 (TPM + 1) transformation to approximate normality [69,70]. We extracted the normalized expression matrix and clinical metadata of 295 pancreatic cancer samples (Accession: GSE183795 and GSE78229) [71,72] were downloaded from the Gene Expression Omnibus (GEO) repository (https://www.ncbi.nlm.nih.gov/geo/) accessed on 2 March 2025. Samples with >20% missing expression values or no survival data were excluded for both cohorts. Finally, we included 185 samples from TCGA-PAAD and 182 samples from GSE183795 and GSE78229. There is no ethical issue because our study utilized the publicly available datasets online.

4.2. Quantification of Cancer-Associated Fibroblasts (CAFs)

To comprehensively characterize CAF infiltration within the tumor microenvironment (TME), we integrated four established computational algorithms complemented by stromal scoring analysis: EPIC (Estimating the Proportion of Immune and Cancer cells) [73], xCell [74], MCP-counter (Microenvironment Cell Populations-counter) [75], and TIDE (Tumor Immune Dysfunction and Exclusion) [76,77]. The four algorithms EPIC (v 1.1.7), (https://github.com/GfellerLab/EPIC, accessed on 2 March 2025), xCell (v 1.1.0) (https://github.com/dviraran/xCell, accessed on 2 March 2025), and MCP-counter (v 1.2.0) (https://github.com/ebecht/MCPcounter, accessed on 2 March 2025) analyses were based on the "limma ()" R package (version "4.4") [78] with default parameters. TIDE scores were obtained through the official web portal (http://tide.dfci.harvard.edu/, accessed on 2 March 2025) using pre-processed log2 (TPM + 1) expression matrices. The ESTIMATE algorithm [79] was additionally applied via the estimate R package (v 1.0.13) to calculate stromal scores reflecting global stromal infiltration levels. This stromal index served as orthogonal validation for CAF quantification. We normalized the different methods of CAF scores to generate a consensus CAF infiltration index and reduce algorithm biases.

4.3. CAF and Stromal Co-Expression Network Module

The WGCNA R package (v 1.73) was utilized to conduct a weighted gene co-expression network analysis (WGCNA) [19], aiming to build co-expression networks and pinpoint key genes linked to the infiltration of cancer-related fibroblasts (CAF) and their stromal ratings. WGCNA identifies gene modules based on similarity in expression patterns and detects hub genes highly connected within relevant modules. Gene expression profiles from the TCGA-PAAD and GSE183795 and GSE78229 cohorts were analyzed. The leading 5000 genes exhibiting the most significant median absolute deviation (MAD) were chosen to construct the network. A Pearson correlation similarity matrix was computed and transformed using a soft-thresholding power to approximate scale-free topology. The adjacency matrix was converted into a topological overlap matrix (TOM), followed by hierarchical clustering based on TOM-derived dissimilarity. The identification of gene modules was achieved through the dynamic tree-cut algorithm, requiring a minimum of 30 genes, and selected 1000 genes for visualization. The initial principal component of each module, known as module eigengenes (MEs), showed a correlation with CAF infiltration levels and stromal scores estimated by EPIC. The most relevant module was selected for further analysis. Core genes were identified based on gene significance (GS > 0.4) and module membership (MM > 0.8), indicating a strong association with the trait and module. Intersection hub genes were considered as our final hub genes in these two cohorts.

4.4. Kaplan–Meier Curve and Log-Rank Analysis

Cohorts included from the GSE183795, GSE78229, and TCGA-PAAD were grouped into high-/low-CAF scores, and overall survival (OS) was determined using Kaplan–Meier curves and checked by log-rank analysis based on a scoring algorithm. Subsequently, we also verified our constructed model genes.

4.5. Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) Analysis

The enrichment analysis of Gene Ontology (GO) and KEGG pathways was performed to delineate the biological importance of the pinpointed central genes. Functional categories, including biological processes (BPs), molecular functions (MFs), and cellular components (CCs), were analyzed using the "clusterProfiler" R package (version 3.20) [80,81]. $p < 0.05$ was considered statistically significant. The hub genes enrichment results were visualized using bar and bubble plots.

4.6. CAF-Associated Model Construction and Validation

To further explore the biological implications of CAF-related hub genes, a CAF-associated model was developed based on the TCGA-PAAD cohort, along with the GSE183795 and GSE78229 datasets for validation. Univariate analysis was performed to identify genes significantly correlated with CAF infiltration. Significant genes were subjected to the least absolute shrinkage and selection operator (LASSO) regression with 1000 iterations to refine the model, using the "glmnet" R package (version 4.1.8) [82]. The CAF-related model was constructed using the formula: CAF score = $\sum (\beta\_i * Exp\_i)$, where β_i represents the LASSO coefficient of the i-th gene, and Exp_i denotes its expression value. Patients were stratified into high- and low-CAF-score groups based on the median score, and differences in CAF-associated characteristics between the groups were analyzed. The model's robustness was validated in the TCGA-PAAD cohort to ensure reproducibility and biological relevance.

4.7. Sensitivity Drugs Prediction and Molecular Docking

To evaluate the potential clinical implications of the CAF-associated model, drug sensitivity and immunotherapy response predictions were conducted. The chemotherapeutic response was estimated based on the Genomics of Drug Sensitivity in Cancer (GDSC) database (https://www.cancerrxgene.org/) (accessed on 2 March 2025) [83]. Additionally, the Tumour Immune Dysfunction and Exclusion (TIDE) algorithm (http://tide.dfci.harvard.edu/) (accessed on 2 March 2025) was employed to predict responses to immune checkpoint blockade (ICB) therapy [76,77]. The chi-squared test was used to evaluate the variance in reaction rates among groups with high and low CAF scores. Subsequently, an assessment of the CAF-associated model's forecasting accuracy was conducted using receiver operating characteristic (ROC) curves and their respective area under the curve (AUC) values. This analysis provided insight into the model's potential effectiveness in directing tailored treatment approaches. Then, we employed the module genes with sensitive drugs to binding, which conducted molecular docking to assess the binding energies and interaction patterns between candidate drugs and their targets. Sensitivity drugs' structural data were sourced from the PubChem Compound Database (https://pubchem.ncbi.nlm.nih.gov/) (accessed on 2 March 2025) [84] and downloaded in SDF format. Potential protein structural data were downloaded from the Protein Data Bank (https://www.rcsb.org/) (accessed on 2 March 2025) in PDB format. The top essential drugs and the proteins encoded by the respective target genes were subjected to molecular docking using the computerized Protein Ligand Docking online tools (http://cadd.labshare.cn/cb-dock2/index.php) (accessed on 2 March 2025) [85]. The in-

teraction graph was downloaded, and the Vina scores were saved. By identifying ligands that exhibit high binding affinity and beneficial interaction patterns, we can prioritize drug targets for additional experimental validation and refine the design of prospective candidate drugs.

4.8. Gathering and Analyzing Data on Somatic Changes

Somatic mutation data for the TCGA-PAAD cohort were obtained from the TCGA GDC (https://portal.gdc.cancer.gov/) (accessed on 2 March 2025). The 20 most frequently mutated genes in high- and low-CAF-score groups were identified and visualized using the "maftools" R package (3.20) [86]. The tumor mutation burden (TMB) is recognized as a potential indicator of the efficacy of immuno-therapy [87], and it was calculated for each PAAD sample using the "tmb" function in the "maftools" package. The relationship between TMB and CAF risk scores was assessed using Spearman's correlation analysis to explore potential associations between mutational load and CAF-related tumor characteristics.

4.9. GSEA and ssGSEA Enrichment Analyses

A gene set enrichment analysis (GSEA) was conducted to identify key and KEGG pathway gene clusters that exhibited variation in enrichment between groups with high and low CAF scores in the GSE183795 and GSE78229 and TCGA-PAAD cohorts. This analysis was conducted using the enrichplot and "clusterProfiler" R packages, utilizing gene sets "c2.cp.kegg.Hs.symbols.gmt" and "c5.go.Hs.symbols.gmt" from the Molecular Signatures Database (MSigDB) [88]. Additionally, single-sample GSEA (ssGSEA, R package version 3.20) was applied to quantify the enrichment scores of hallmark gene sets related to dilated cardiomyopathy, focal adhesion, hypertrophic cardiomyopathy HCM, and TGF-beta signaling pathway [89]. Spearman's correlation analysis evaluated associations between CAF risk scores and gene set enrichment scores, providing insights into the functional pathways linked to CAF activity.

4.10. Validation Via Cancer Cell Line Encyclopedia (CCLE) and Human Protein Atlas (HPA) Databases

To validate the findings at the cellular level, the mRNA expression data of the identified markers in 37 fibroblasts and 41 pancreatic cell lines were retrieved from the Cancer Cell Line Encyclopedia (CCLE) database (https://portals.broadinstitute.org/ccle) (accessed on 2 March 2025) [90]. Differential expressions between fibroblasts and GC cell lines were analyzed using Wilcoxon tests and visualized with heat maps. At the protein level, immunohistochemical (IHC) staining images of these markers in pancreatic cancer tissues were obtained from the Human Protein Atlas (HPA) database (https://www.proteinatlas.org/) (accessed on 2 March 2025) [91]. Protein localization and expression patterns in pancreatic cancer tissues were visually inspected to corroborate the findings further.

4.11. Statistical Analysis

All statistical evaluations were conducted using R software (version 4.4.2, available on 31 October 2024; https://www.r-project.org/) (accessed on 2 March 2025). Pancreatic cancer patients were categorized into high and low-CAF-risk categories based on the median CAF risk score, which served as the threshold value. The Wilcoxon rank-sum test was employed to conduct paired analyses within each group. The Kaplan–Meier survival curves were employed to evaluate survival variances among sub-groups, and the log-rank test was used to determine statistical significance through the "survival" (v 3.8.3) and "survminer" (v 0.5.0) R packages. A $p < 0.05$ was deemed to hold statistical significance. Correlation analyses utilized Spearman's rank correlation test. A multivariate Cox cregression analysis was conducted to identify distinct prognostic elements, and the CAF risk score's forecasting precision was assessed through receiver operating characteristic (ROC) curves and the

area under the curve (AUC) metrics. The generation of data visualization, encompassing heatmaps, scatter plots, and survival curves, was facilitated by R software, GraphPad Prism (version 8.4.0, launched on 20 February 2020), Adobe Photoshop (version 2024), and Microsoft Office 365 to ensure efficient management of data, comprehensive statistical evaluation, and the generation of figures.

5. Conclusions

In conclusion, our study successfully identified a module of six key genes—GFPT2, MFAP5, CTSK, MMP2, FSTL1, and PRRX1—through Weighted Gene Co-expression Network Analysis (WGCNA) to assess cancer-associated fibroblast (CAF) infiltration in pancreatic cancer (PC). The robustness of these module genes highlights their potential as significant biomarkers for evaluating prognosis and predicting responses to both chemotherapy and immunotherapy. The findings underscore the critical role of the tumor microenvironment (TME) in pancreatic cancer progression and treatment efficacy. Our results suggest that further exploration of the TME's interactions and the implications of CAF infiltration might yield valuable insights into the clinical management of PC. This opens avenues for future research aimed at optimizing therapeutic strategies and enhancing patient outcomes through targeted approaches that consider CAF-related dynamics within the TME. Ultimately, a deeper understanding of these molecular markers and their pathways could lead to more personalized and effective treatment protocols in the fight against pancreatic cancer.

Abbreviations

The following abbreviations are used in this manuscript:

PC	Pancreatic cancer
CAFs	Cancer associated fibroblasts
GEO	Gene Expression Omnibus
TCGA	the Cancer Genome Atlas
WGCNA	weighted gene co-expression network analysis
TMB	tumor mutation burden
GSEA	Gene set enrichment analysis
OS	overall survival

TME	tumor microenvironment
PSCs	pancreatic stellate cells
ECM	extracellular matrix
TGF-β	transforming growth factor-β
IL-6	interleukin-6
CCL2	CC-chemokine ligand 2
ROS	reactive oxygen species
EBV	Epstein–Barr virus
LASSO	Least Absolute Shrinkage and Selection Operator
DEGs	differentially expressed genes
GS	gene significance
MM	module membership
GO	gene ontology
BP	biological process
CC	cell component
MF	molecular function
KEGG	kyoto encyclopedia of genes and genomes
GSEA	Gene Set Enrichment Analysis
ssGSEA	single-sample gene set enrichment analysis
ADM	acinar-ductal metaplasia
FPKM	per million mapped reads
MAF	Mutation Annotation Format
TPM	transcripts per million
EPIC	Estimating the Proportion of Immune and Cancer cells
MCP-counter	Microenvironment Cell Populations-counter
TIDE	Tumor Immune Dysfunction and Exclusion
MAD	median absolute deviation
TOM	topological overlap matrix
MEs	Module eigengenes
GDSC	the Genomics of Drug Sensitivity in Cancer
ICB	immune checkpoint blockade
AUC	area under the curve
CCLE	Cancer Cell Line Encyclopedias
HPA	Human Protein Atlas
ROC	receiver operating characteristic

References

1. Bray, F.; Laversanne, M.; Sung, H.; Ferlay, J.; Siegel, R.L.; Soerjomataram, I.; Jemal, A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA A Cancer J. Clin.* **2024**, *74*, 229–263. [CrossRef]
2. Zhao, Z.; Li, T.; Sun, L.; Yuan, Y.; Zhu, Y. Potential mechanisms of cancer-associated fibroblasts in therapeutic resistance. *Biomed. Pharmacother.* **2023**, *166*, 115425. [CrossRef] [PubMed]
3. Liot, S.; Balas, J.; Aubert, A.; Prigent, L.; Mercier-Gouy, P.; Verrier, B.; Bertolino, P.; Hennino, A.; Valcourt, U.; Lambert, E. Stroma Involvement in Pancreatic Ductal Adenocarcinoma: An Overview Focusing on Extracellular Matrix Proteins. *Front. Immunol.* **2021**, *12*, 612271. [CrossRef] [PubMed]
4. Chen, X.; Song, E. Turning foes to friends: Targeting cancer-associated fibroblasts. *Nat. Rev. Drug Discov.* **2019**, *18*, 99–115. [CrossRef]
5. Glaviano, A.; Lau, H.S.-H.; Carter, L.M.; Lee, E.H.C.; Lam, H.Y.; Okina, E.; Tan, D.J.J.; Tan, W.; Ang, H.L.; Carbone, D.; et al. Harnessing the tumor microenvironment: Targeted cancer therapies through modulation of epithelial-mesenchymal transition. *J. Hematol. Oncol.* **2025**, *18*, 6.
6. Verloy, R.; Privat-Maldonado, A.; Smits, E.; Bogaerts, A. Cold Atmospheric Plasma Treatment for Pancreatic Cancer-The Importance of Pancreatic Stellate Cells. *Cancers* **2020**, *12*, 2782. [CrossRef]
7. Peng, J.; Sun, B.-F.; Chen, C.-Y.; Zhou, J.-Y.; Chen, Y.-S.; Chen, H.; Liu, L.; Huang, D.; Jiang, J.; Cui, G.-S.; et al. Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. *Cell Res.* **2019**, *29*, 725–738. [CrossRef]

8. Wang, Y.; Liang, Y.; Xu, H.; Zhang, X.; Mao, T.; Cui, J.; Yao, J.; Wang, Y.; Jiao, F.; Xiao, X.; et al. Single-cell analysis of pancreatic ductal adenocarcinoma identifies a novel fibroblast subtype associated with poor prognosis but better immunotherapy response. *Cell Discov.* **2021**, *7*, 36. [CrossRef]

9. Yu, Y.; Yang, G.; Huang, H.; Fu, Z.; Cao, Z.; Zheng, L.; You, L.; Zhang, T. Preclinical models of pancreatic ductal adenocarcinoma: Challenges and opportunities in the era of precision medicine. *J. Exp. Clin. Cancer Res.* **2021**, *40*, 8. [CrossRef]

10. Werba, G.; Weissinger, D.; Kawaler, E.A.; Zhao, E.; Kalfakakou, D.; Dhara, S.; Wang, L.; Lim, H.B.; Oh, G.; Jing, X.; et al. Single-cell RNA sequencing reveals the effects of chemotherapy on human pancreatic adenocarcinoma and its tumor microenvironment. *Nat. Commun.* **2023**, *14*, 797. [CrossRef]

11. Ye, B.; Wang, Q.; Zhu, X.; Zeng, L.; Luo, H.; Xiong, Y.; Li, Q.; Zhu, Q.; Zhao, S.; Chen, T.; et al. Single-cell RNA sequencing identifies a novel proliferation cell type affecting clinical outcome of pancreatic ductal adenocarcinoma. *Front. Oncol.* **2023**, *13*, 1236435. [CrossRef] [PubMed]

12. Cros, J.; Raffenne, J.; Couvelard, A.; Poté, N. Tumour Heterogeneity in Pancreatic Adenocarcinoma. *Pathobiology* **2018**, *85*, 64–71. [CrossRef] [PubMed]

13. Gaggioli, C.; Hooper, S.; Hidalgo-Carcedo, C.; Grosse, R.; Marshall, J.F.; Harrington, K.; Sahai, E. Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells. *Nat. Cell Biol.* **2007**, *9*, 1392–1400. [CrossRef]

14. Su, S.; Chen, J.; Yao, H.; Liu, J.; Yu, S.; Lao, L.; Wang, M.; Luo, M.; Xing, Y.; Chen, F.; et al. CD10⁺GPR77⁺ Cancer-Associated Fibroblasts Promote Cancer Formation and Chemoresistance by Sustaining Cancer Stemness. *Cell* **2018**, *172*, 841–856.e16. [CrossRef]

15. Flavell, R.A.; Sanjabi, S.; Wrzesinski, S.H.; Licona-Limón, P. The polarization of immune cells in the tumour environment by TGFbeta. *Nat. Rev. Immunol.* **2010**, *10*, 554–567. [CrossRef]

16. Yang, X.; Lin, Y.; Shi, Y.; Li, B.; Liu, W.; Yin, W.; Dang, Y.; Chu, Y.; Fan, J.; He, R. FAP Promotes Immunosuppression by Cancer-Associated Fibroblasts in the Tumor Microenvironment via STAT3-CCL2 Signaling. *Cancer Res.* **2016**, *76*, 4124–4135. [CrossRef]

17. Cheteh, E.H.; Augsten, M.; Rundqvist, H.; Bianchi, J.; Sarne, V.; Egevad, L.; Bykov, V.J.; Östman, A.; Wiman, K.G. Human cancer-associated fibroblasts enhance glutathione levels and antagonize drug-induced prostate cancer cell death. *Cell Death Dis.* **2017**, *8*, e2848. [CrossRef]

18. Straussman, R.; Morikawa, T.; Shee, K.; Barzily-Rokni, M.; Qian, Z.R.; Du, J.; Davis, A.; Mongare, M.M.; Gould, J.; Frederick, D.T.; et al. Tumour micro-environment elicits innate resistance to RAF inhibitors through HGF secretion. *Nature* **2012**, *487*, 500–504. [CrossRef]

19. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **2008**, *9*, 559. [CrossRef]

20. Zhu, Y.; Yang, J.; Liu, X.; Sun, T.; Zhao, Y.; Xiang, F.; Chen, F.; He, H. Transcriptome Analysis Reveals Coexpression Networks and Hub Genes Involved in Papillae Development in *Lilium auratum*. *Int. J. Mol. Sci.* **2024**, *25*, 2436. [CrossRef]

21. Zhou, H.; Jing, S.; Liu, Y.; Wang, X.; Duan, X.; Xiong, W.; Li, R.; Peng, Y.; Ai, Y.; Fu, D.; et al. Identifying the key genes of Epstein-Barr virus-regulated tumour immune microenvironment of gastric carcinomas. *Cell Prolif.* **2023**, *56*, e13373. [CrossRef] [PubMed]

22. Zou, C.; Zhang, Y.; Liu, C.; Li, Y.; Lin, C.; Chen, H.; Hou, J.; Gao, G.; Liu, Z.; Yan, Q.; et al. Identification of CENPM as a key gene driving adrenocortical carcinoma metastasis via physical interaction with immune checkpoint ligand FGL1. *Clin. Transl. Med.* **2025**, *15*, e70182. [CrossRef]

23. Mahajan, M.; Dhabalia, S.; Dash, T.; Sarkar, A.; Mondal, S. A comprehensive multi-omics study reveals potential prognostic and diagnostic biomarkers for colorectal cancer. *Int. J. Biol. Macromol.* **2025**, *303*, 140443. [CrossRef]

24. Xu, Z.; Wang, J.; Wang, G. Weighted gene co-expression network analysis for hub genes in colorectal cancer. *Pharmacol. Rep.* **2024**, *76*, 140–153. [CrossRef]

25. Pengping, L.; Kexin, Y.; Yuwei, X.; Ke, S.; Rongguo, L.; Zhenyu, W.; Haigang, J.; Shaowen, W.; Yuqing, H. Construction of an anaplastic thyroid cancer stratification signature to guide immune therapy selection and validation of the pivotal gene HLF through in vitro experiments. *Front. Immunol.* **2024**, *15*, 1478904. [CrossRef]

26. Lin, S.; Zhou, M.; Cheng, L.; Shuai, Z.; Zhao, M.; Jie, R.; Wan, Q.; Peng, F.; Ding, S. Exploring the association of POSTN(+) cancer-associated fibroblasts with triple-negative breast cancer. *Int. J. Biol. Macromol.* **2024**, *268 Pt 1*, 131560. [CrossRef]

27. Zheng, H.; Liu, H.; Li, H.; Dou, W.; Wang, X. Weighted Gene Co-expression Network Analysis Identifies a Cancer-Associated Fibroblast Signature for Predicting Prognosis and Therapeutic Responses in Gastric Cancer. *Front. Mol. Biosci.* **2021**, *8*, 744677. [CrossRef]

28. Zhong, Y.-J.; Luo, Z.-M.; Liu, F.; He, Z.-Q.; Yang, S.-Q.; Ma, W.-J.; Wang, J.-K.; Dai, Y.-S.; Zou, R.-Q.; Hu, Y.-F.; et al. Integrative analyses of bulk and single-cell transcriptomics reveals the infiltration and crosstalk of cancer-associated fibroblasts as a novel predictor for prognosis and microenvironment remodeling in intrahepatic cholangiocarcinoma. *J. Transl. Med.* **2024**, *22*, 422. [CrossRef]

29. Li, Y.; Tian, R.; Liu, J.; Li, J.; Tan, H.; Wu, Q.; Fu, X. Deciphering the immune landscape dominated by cancer-associated fibroblasts to investigate their potential in indicating prognosis and guiding therapeutic regimens in high grade serous ovarian carcinoma. *Front. Immunol.* **2022**, *13*, 940801. [CrossRef]

30. Zhang, Z.; Liu, S.; Wang, Z.; Wang, S.; Jiang, L.; Wang, X.; Li, J.; Shen, L. Influential upregulation of KCNE4: Propelling cancer associated fibroblasts-driven colorectal cancer progression. *Cancer Cell Int.* **2024**, *24*, 103. [CrossRef]

31. Du, Y.; Jiang, X.; Wang, B.; Cao, J.; Wang, Y.; Yu, J.; Wang, X.; Liu, H. The cancer-associated fibroblasts related gene CALD1 is a prognostic biomarker and correlated with immune infiltration in bladder cancer. *Cancer Cell Int.* **2021**, *21*, 283. [CrossRef]

32. Bailey, P.; Chang, D.K.; Nones, K.; Johns, A.L.; Patch, A.-M.; Gingras, M.-C.; Miller, D.K.; Christ, A.N.; Bruxner, T.J.C.; Quinn, M.C.; et al. Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* **2016**, *531*, 47–52. [CrossRef] [PubMed]

33. Niu, N.; Shen, X.; Wang, Z.; Chen, Y.; Weng, Y.; Yu, F.; Tang, Y.; Lu, P.; Liu, M.; Wang, L.; et al. Tumor cell-intrinsic epigenetic dysregulation shapes cancer-associated fibroblasts heterogeneity to metabolically support pancreatic cancer. *Cancer Cell* **2024**, *42*, 869–884.e9. [CrossRef] [PubMed]

34. Arpinati, L.; Carradori, G.; Scherz-Shouval, R. CAF-induced physical constraints controlling T cell state and localization in solid tumours. *Nat. Rev. Cancer* **2024**, *24*, 676–693. [CrossRef] [PubMed]

35. Levi-Galibov, O.; Lavon, H.; Wassermann-Dozorets, R.; Pevsner-Fischer, M.; Mayer, S.; Wershof, E.; Stein, Y.; Brown, L.E.; Zhang, W.; Friedman, G.; et al. Heat Shock Factor 1-dependent extracellular matrix remodeling mediates the transition from chronic intestinal inflammation to colon cancer. *Nat. Commun.* **2020**, *11*, 6245. [CrossRef]

36. Berestjuk, I.; Lecacheur, M.; Carminati, A.; Diazzi, S.; Rovera, C.; Prod'homme, V.; Ohanna, M.; Popovic, A.; Mallavialle, A.; Larbret, F.; et al. Targeting Discoidin Domain Receptors DDR1 and DDR2 overcomes matrix-mediated tumor cell adaptation and tolerance to BRAF-targeted therapy in melanoma. *EMBO Mol. Med.* **2022**, *14*, e11814. [CrossRef]

37. Zhang, J.Y.; Zhu, W.-W.; Wang, M.-Y.; Zhai, R.-D.; Wang, Q.; Shen, W.-L.; Liu, L.-K. Cancer-associated fibroblasts promote oral squamous cell carcinoma progression through LOX-mediated matrix stiffness. *J. Transl. Med.* **2021**, *19*, 513. [CrossRef]

38. Li, Q.; Zhu, C.-C.; Ni, B.; Zhang, Z.-Z.; Jiang, S.-H.; Hu, L.-P.; Wang, X.; Zhang, X.-X.; Huang, P.-Q.; Yang, Q.; et al. Lysyl oxidase promotes liver metastasis of gastric cancer via facilitating the reciprocal interactions between tumor cells and cancer associated fibroblasts. *EBioMedicine* **2019**, *49*, 157–171. [CrossRef]

39. Li, C.; Qiu, S.; Liu, X.; Guo, F.; Zhai, J.; Li, Z.; Deng, L.; Ge, L.; Qian, H.; Yang, L.; et al. Extracellular matrix-derived mechanical force governs breast cancer cell stemness and quiescence transition through integrin-DDR signaling. *Signal Transduct. Target Ther.* **2023**, *8*, 247. [CrossRef]

40. Hupfer, A.; Brichkina, A.; Koeniger, A.; Keber, C.; Denkert, C.; Pfefferle, P.; Helmprobst, F.; Pagenstecher, A.; Visekruna, A.; Lauth, M. Matrix stiffness drives stromal autophagy and promotes formation of a protumourigenic niche. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2105367118. [CrossRef]

41. Malik, R.; Luong, T.; Cao, X.; Han, B.; Shah, N.; Franco-Barraza, J.; Han, L.; Shenoy, V.; Lelkes, P.; Cukierman, E. Rigidity controls human desmoplastic matrix anisotropy to enable pancreatic cancer cell spread via extracellular signal-regulated kinase 2. *Matrix Biol.* **2019**, *81*, 50–69. [CrossRef] [PubMed]

42. Bera, K.; Kiepas, A.; Godet, I.; Li, Y.; Mehta, P.; Ifemembi, B.; Paul, C.D.; Sen, A.; Serra, S.A.; Stoletov, K.; et al. Extracellular fluid viscosity enhances cell migration and cancer dissemination. *Nature* **2022**, *611*, 365–373. [CrossRef] [PubMed]

43. Kaur, A.; Ecker, B.L.; Douglass, S.M.; Kugel, C.H.; Webster, M.R.; Almeida, F.V.; Somasundaram, R.; Hayden, J.; Ban, E.; Ahmadzadeh, H.; et al. Remodeling of the Collagen Matrix in Aging Skin Promotes Melanoma Metastasis and Affects Immune Cell Motility. *Cancer Discov.* **2019**, *9*, 64–81. [CrossRef]

44. Puttock, E.H.; Tyler, E.J.; Manni, M.; Maniati, E.; Butterworth, C.; Ramos, M.B.; Peerani, E.; Hirani, P.; Gauthier, V.; Liu, Y.; et al. Extracellular matrix educates an immunoregulatory tumor macrophage phenotype found in ovarian cancer metastasis. *Nat. Commun.* **2023**, *14*, 2514. [CrossRef]

45. Rafaeva, M.; Jensen, A.R.D.; Horton, E.R.; Zornhagen, K.W.; Strøbech, J.E.; Fleischhauer, L.; Mayorca-Guiliani, A.E.; Nielsen, S.R.; Grønseth, D.S.; Kuś, F.; et al. Fibroblast-derived matrix models desmoplastic properties and forms a prognostic signature in cancer progression. *Front. Immunol.* **2023**, *14*, 1154528. [CrossRef]

46. Huang, T.X.; Tan, X.-Y.; Huang, H.-S.; Li, Y.-T.; Liu, B.-L.; Liu, K.-S.; Chen, X.; Chen, Z.; Guan, X.-Y.; Zou, C.; et al. Targeting cancer-associated fibroblast-secreted WNT2 restores dendritic cell-mediated antitumour immunity. *Gut* **2022**, *71*, 333–344. [CrossRef]

47. Cheng, J.T.; Deng, Y.-N.; Yi, H.-M.; Wang, G.-Y.; Fu, B.-S.; Chen, W.-J.; Liu, W.; Tai, Y.; Peng, Y.-W.; Zhang, Q. Hepatic carcinoma-associated fibroblasts induce IDO-producing regulatory dendritic cells through IL-6-mediated STAT3 activation. *Oncogenesis* **2016**, *5*, e198. [CrossRef]

48. Kalluri, R.; Zeisberg, M. Fibroblasts in cancer. *Nat. Rev. Cancer* **2006**, *6*, 392–401. [CrossRef]

49. Yu, C.; You, M.; Zhang, P.; Zhang, S.; Yin, Y.; Zhang, X. A five-gene signature is a prognostic biomarker in pan-cancer and related with immunologically associated extracellular matrix. *Cancer Med.* **2021**, *10*, 4629–4643. [CrossRef]

50. Gross, J.; Lapiere, C.M. Collagenolytic activity in amphibian tissues: A tissue culture assay. *Proc. Natl. Acad. Sci. USA* **1962**, *48*, 1014–1022. [CrossRef]

51. Montoya Mira, J.L.; Quentel, A.; Patel, R.K.; Keith, D.; Sousa, M.; Minnier, J.; Kingston, B.R.; David, L.; Esener, S.C.; Sears, R.C.; et al. Early detection of pancreatic cancer by a high-throughput protease-activated nanosensor assay. *Sci. Transl. Med.* **2025**, *17*, eadq3110. [CrossRef] [PubMed]

52. Page-McCaw, A.; Ewald, A.J.; Werb, Z. Matrix metalloproteinases and the regulation of tissue remodelling. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 221–233. [CrossRef] [PubMed]

53. McQuibban, G.A.; Gong, J.-H.; Tam, E.M.; McCulloch, C.A.G.; Clark-Lewis, I.; Overall, C.M. Inflammation dampened by gelatinase A cleavage of monocyte chemoattractant protein-3. *Science* **2000**, *289*, 1202–1206. [CrossRef]

54. Parks, W.C.; Wilson, C.L.; López-Boado, Y.S. Matrix metalloproteinases as modulators of inflammation and innate immunity. *Nat. Rev. Immunol.* **2004**, *4*, 617–629. [CrossRef]

55. Egeblad, M.; Werb, Z. New functions for the matrix metalloproteinases in cancer progression. *Nat. Rev. Cancer* **2002**, *2*, 161–174. [CrossRef]

56. Li, X.; Bao, C.; Ma, Z.; Xu, B.; Ying, X.; Liu, X.; Zhang, X. Perfluorooctanoic acid stimulates ovarian cancer cell migration, invasion via ERK/NF-κB/MMP-2/-9 pathway. *Toxicol. Lett.* **2018**, *294*, 44–50. [CrossRef]

57. Agrawal, S.; Anderson, P.; Durbeej, M.; van Rooijen, N.; Ivars, F.; Opdenakker, G.; Sorokin, L.M. Dystroglycan is selectively cleaved at the parenchymal basement membrane at sites of leukocyte extravasation in experimental autoimmune encephalomyelitis. *J. Exp. Med.* **2006**, *203*, 1007–1019. [CrossRef]

58. Littlepage, L.E.; Sternlicht, M.D.; Rougier, N.; Phillips, J.; Gallo, E.; Yu, Y.; Williams, K.; Brenot, A.; Gordon, J.I.; Werb, Z. Matrix metalloproteinases contribute distinct roles in neuroendocrine prostate carcinogenesis, metastasis, and angiogenesis progression. *Cancer Res.* **2010**, *70*, 2224–2234. [CrossRef]

59. Langenskiöld, M.; Holmdahl, L.; Falk, P.; Ivarsson, M.-L. Increased plasma MMP-2 protein expression in lymph node-positive patients with colorectal cancer. *Int. J. Color. Dis.* **2005**, *20*, 245–252. [CrossRef]

60. Nakamura, E.S.; Koizumi, K.; Kobayashi, M.; Saiki, I. Inhibition of lymphangiogenesis-related properties of murine lymphatic endothelial cells and lymph node metastasis of lung cancer by the matrix metalloproteinase inhibitor MMI270. *Cancer Sci.* **2004**, *95*, 25–31. [CrossRef]

61. Wong, J.C.; Chan, S.K.; Schaeffer, D.F.; Sagaert, X.; Lim, H.J.; Kennecke, H.; Owen, D.A.; Suh, K.W.; Kim, Y.-B.; Tai, I.T. Absence of MMP2 expression correlates with poor clinical outcomes in rectal cancer, and is distinct from MMP1-related outcomes in colon cancer. *Clin. Cancer Res.* **2011**, *17*, 4167–4176. [CrossRef]

62. Kessenbrock, K.; Plaks, V.; Werb, Z. Matrix metalloproteinases: Regulators of the tumor microenvironment. *Cell* **2010**, *141*, 52–67. [CrossRef] [PubMed]

63. Reichert, M.; Takano, S.; von Burstin, J.; Kim, S.-B.; Lee, J.-S.; Ihida-Stansbury, K.; Hahn, C.; Heeg, S.; Schneider, G.; Rhim, A.D.; et al. The Prrx1 homeodomain transcription factor plays a central role in pancreatic regeneration and carcinogenesis. *Genes Dev.* **2013**, *27*, 288–300. [CrossRef] [PubMed]

64. Feldmann, K.; Maurer, C.; Peschke, K.; Teller, S.; Schuck, K.; Steiger, K.; Engleitner, T.; Öllinger, R.; Nomura, A.; Wirges, N.; et al. Mesenchymal Plasticity Regulated by Prrx1 Drives Aggressive Pancreatic Cancer Biology. *Gastroenterology* **2021**, *160*, 346–361.e24. [CrossRef]

65. Takano, S.; Reichert, M.; Bakir, B.; Das, K.K.; Nishida, T.; Miyazaki, M.; Heeg, S.; Collins, M.A.; Marchand, B.; Hicks, P.D.; et al. Prrx1 isoform switching regulates pancreatic cancer invasion and metastatic colonization. *Genes Dev.* **2016**, *30*, 233–247. [CrossRef]

66. Leung, C.S.; Yeung, T.-L.; Yip, K.-P.; Pradeep, S.; Balasubramanian, L.; Liu, J.; Wong, K.-K.; Mangala, L.S.; Armaiz-Pena, G.N.; Lopez-Berestein, G.; et al. Calcium-dependent FAK/CREB/TNNC1 signalling mediates the effect of stromal MFAP5 on ovarian cancer metastatic potential. *Nat. Commun.* **2014**, *5*, 5092. [CrossRef]

67. Li, J.H.; Zhu, X.-X.; Li, F.-X.; Huang, C.-S.; Huang, X.-T.; Wang, J.-Q.; Gao, Z.-X.; Li, S.-J.; Xu, Q.-C.; Zhao, W.; et al. MFAP5 facilitates the aggressiveness of intrahepatic Cholangiocarcinoma by activating the Notch1 signaling pathway. *J. Exp. Clin. Cancer Res.* **2019**, *38*, 476. [CrossRef]

68. Zhu, L.; Chen, X.; Zhu, Y.; Qin, J.; Niu, T.; Ding, Y.; Xiao, Y.; Jiang, Y.; Liu, K.; Lu, J.; et al. Dihydroartemisinin Inhibits the Proliferation of Esophageal Squamous Cell Carcinoma Partially by Targeting AKT1 and p70S6K. *Front. Pharmacol.* **2020**, *11*, 587470. [CrossRef]

69. Lee, S.; Seo, C.H.; Lim, B.; Yang, J.O.; Oh, J.; Kim, M.; Lee, S.; Lee, B.; Kang, C.; Lee, S. Accurate quantification of transcriptome from RNA-Seq data by effective length normalization. *Nucleic Acids Res.* **2011**, *39*, e9. [CrossRef]

70. Zhao, Y.; Li, M.-C.; Konaté, M.M.; Chen, L.; Das, B.; Karlovich, C.; Williams, P.M.; Evrard, Y.A.; Doroshow, J.H.; McShane, L.M. TPM, FPKM, or Normalized Counts? A Comparative Study of Quantification Measures for the Analysis of RNA-seq Data from the NCI Patient-Derived Models Repository. *J. Transl. Med.* **2021**, *19*, 269.

71. Yang, S.; Tang, W.; Azizian, A.; Gaedcke, J.; Ströbel, P.; Wang, L.; Cawley, H.; Ohara, Y.; Valenzuela, P.; Zhang, L.; et al. Dysregulation of HNF1B/Clusterin axis enhances disease progression in a highly aggressive subset of pancreatic cancer patients. *Carcinogenesis* **2022**, *43*, 1198–1210. [CrossRef]

72. Hilmi, M.; Delecourt, F.; Raffenne, J.; Bourega, T.; Dusetti, N.; Iovanna, J.; Blum, Y.; Richard, M.; Neuzillet, C.; Couvelard, A.; et al. Redefining phenotypic intratumor heterogeneity of pancreatic ductal adenocarcinoma: A bottom-up approach. *J. Pathol.* **2025**, *265*, 448–461. [CrossRef] [PubMed]

73. Racle, J.; De Jonge, K.; Baumgaertner, P.; Speiser, D.E.; Gfeller, D. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *eLife* **2017**, *6*, e26476. [CrossRef]

74. Aran, D. Cell-Type Enrichment Analysis of Bulk Transcriptomes Using xCell. In *Bioinformatics for Cancer Immunotherapy: Methods and Protocols*; Boegel, S., Ed.; Springer: New York, NY, USA, 2020; pp. 263–276.

75. Becht, E.; Giraldo, N.A.; Lacroix, L.; Buttard, B.; Elarouci, N.; Petitprez, F.; Selves, J.; Laurent-Puig, P.; Sautes-Fridman, C.; Fridman, W.H.; et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* **2016**, *17*, 218.

76. Jiang, P.; Gu, S.; Pan, D.; Fu, J.; Sahu, A.; Hu, X.; Li, Z.; Traugh, N.; Bu, X.; Li, B.; et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* **2018**, *24*, 1550–1558. [CrossRef]

77. Li, X.; Xiang, Y.; Li, F.; Yin, C.; Li, B.; Ke, X. WNT/β-Catenin Signaling Pathway Regulating T Cell-Inflammation in the Tumor Microenvironment. *Front. Immunol.* **2019**, *10*, 2293. [CrossRef]

78. Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **2015**, *43*, e47. [CrossRef]

79. Yoshihara, K.; Shahmoradgoli, M.; Martínez, E.; Vegesna, R.; Kim, H.; Torres-Garcia, W.; Treviño, V.; Shen, H.; Laird, P.W.; Levine, D.A.; et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* **2013**, *4*, 2612. [CrossRef]

80. Xu, S.; Hu, E.; Cai, Y.; Xie, Z.; Luo, X.; Zhan, L.; Tang, W.; Wang, Q.; Liu, B.; Wang, R.; et al. Using clusterProfiler to characterize multiomics data. *Nat. Protoc.* **2024**, *19*, 3292–3320. [CrossRef] [PubMed]

81. Yu, G. Thirteen years of clusterProfiler. *Innovation* **2024**, *5*, 100722. [CrossRef]

82. Simon, N.; Friedman, J.H.; Hastie, T.; Tibshirani, R. Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent. *J. Stat. Softw.* **2011**, *39*, 1–13. [CrossRef] [PubMed]

83. Yang, W.; Soares, J.; Greninger, P.; Edelman, E.J.; Lightfoot, H.; Forbes, S.; Bindal, N.; Beare, D.; Smith, J.A.; Thompson, I.R.; et al. Genomics of Drug Sensitivity in Cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* **2013**, *41*, D955–D961. [CrossRef] [PubMed]

84. Rigden, D.J.; Fernández, X.M. The 2025 Nucleic Acids Research database issue and the online molecular biology database collection. *Nucleic Acids Res.* **2025**, *53*, D1–D9. [CrossRef]

85. Liu, Y.; Yang, X.; Gan, J.; Chen, S.; Xiao, Z.X.; Cao, Y. CB-Dock2: Improved protein-ligand blind docking by integrating cavity detection, docking and homologous template fitting. *Nucleic Acids Res.* **2022**, *50*, W159–W164. [CrossRef]

86. Mayakonda, A.; Lin, D.-C.; Assenov, Y.; Plass, C.; Koeffler, H.P. Maftools: Efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* **2018**, *28*, 1747–1756. [CrossRef]

87. Yarchoan, M.; Hopkins, A.; Jaffee, E.M. Tumor Mutational Burden and Response Rate to PD-1 Inhibition. *N. Engl. J. Med.* **2017**, *377*, 2500–2501. [CrossRef]

88. Liberzon, A.; Birger, C.; Thorvaldsdóttir, H.; Ghandi, M.; Mesirov, J.P.; Tamayo, P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **2015**, *1*, 417–425. [CrossRef]

89. Hänzelmann, S.; Castelo, R.; Guinney, J. GSVA: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform.* **2013**, *14*, 7. [CrossRef]

90. Ghandi, M.; Huang, F.W.; Jané-Valbuena, J.; Kryukov, G.V.; Lo, C.C.; McDonald, E.R., III; Barretina, J.; Gelfand, E.T.; Bielski, C.M.; Li, H.; et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **2019**, *569*, 503–508. [CrossRef]

91. Uhlén, M.; Fagerberg, L.; Hallström, B.M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, Å.; Kampf, C.; Sjöstedt, E.; Asplund, A.; et al. Proteomics. Tissue-based map of the human proteome. *Science* **2015**, *347*, 1260419.