ORIGINAL ARTICLE

The cellular trajectories of tumor-associated macrophages decipher the heterogeneity of pancreatic cancer

Jiajing Lin[1] · Fengchun Lu[1] · Yuwei Wu[1] · Heguang Huang[1] · Yu Pan[1]

Abstract

Emerging evidence indicates that the interactions and dynamic changes among tumor-associated macrophages (TAMs) are pivotal in molding the tumor microenvironment (TME), thereby influencing diverse clinical outcomes. However, the potential clinical ramifications of these evolutionary shifts in tumor-associated macrophages within pancreatic adenocarcinoma (PAAD) remain largely unexamined. Single-cell RNA sequencing (scRNA-seq) data were retrieved from the Tumor Immune Single-cell Hub. The Seurat and Monocle algorithms were employed to elucidate the progression of TAMs, using non-negative matrix factorization (NMF) to determine molecular classifications. Subsequently, the prognosis, biological characteristics, genomic modifications, and immune landscape across various clusters were interpreted. Furthermore, the sensitivity of potential therapeutic drugs between subtypes was predicted. Cellular experiments were conducted to explore the function of the NR1H3 gene in pancreatic cancer. These experiments encompassed gene knockdown, proliferation assessment, clone formation evaluation, transwell examination, and apoptosis analysis. Trajectory gene expression analysis of tumor-associated macrophages identified three disparate clusters, each associated with different clinical outcomes Compared to clusters C1 and C2, cluster C3 is seemingly at a less advanced pathological stage and associates with a relatively favorable prognosis. Further investigation revealed pronounced genetic instability in cluster C2, whereas cluster C3 demonstrated notable genetic stability. Cluster C1, characterized as "immune-hot," exhibits an abundance of immune cells and elevated immune checkpoint expression, suggesting its suitability for immunotherapy. Furthermore, several potential therapeutic agents have been pinpointed, potentially facilitating the clinical application of these insights. Cell assays indicated that NR1H3 knockdown markedly induced apoptosis and suppressed clonogenesis, migration, and proliferation of pancreatic cancer cells in the PTAU-8988 and PANC-1 cell lines. Overall, our study discerned three clusters with unique characteristics, defined by the evolution of TAMs. We propose customized therapeutic strategies for patients within these specific clusters to improve clinical outcomes and optimize clinical management.

Keywords Pancreatic cancer · Single-cell RNA-seq · Immunotherapy · Heterogeneity · Prognosis

✉ Heguang Huang
heguanghuang22@163.com

✉ Yu Pan
yupan199002@163.com

Jiajing Lin
627919928@qq.com

Fengchun Lu
fengchun160@163.com

Yuwei Wu
wuyuweiv@163.com

1 Department of General Surgery, Fujian Medical University Union Hospital, No. 29 Xinquan Road, Fuzhou 350001, People's Republic of China

Introduction

Pancreatic ductal adenocarcinoma (PDAC) is a highly aggressive gastrointestinal tumor, characterized by a 5-year overall survival (OS) rate of only 9% (Siegel et al. 2022). The treatment options for PDAC patients are limited (Park et al. 2021; Shi et al. 2019). While numerous novel drugs targeting the tumor microenvironment are under clinical trials, such as immune checkpoint inhibitors (ICIs), only a handful has yielded significant extensions in PDAC patients' life expectancy, largely due to an incomplete understanding of this intricate microenvironment (Emens et al. 2017; Ganesh et al. 2019; Bailey et al. 2016).

Moreover, prior research has indicated that patients with PDAC, even at identical clinical stages, can display divergent therapeutic responses and considerable prognostic disparities, a phenomenon attributed to molecular heterogeneity (Bailey et al. 2016). This heterogeneity is frequently neglected by extant clinical classification systems that emphasize clinicopathological features, resulting in suboptimal patient stratification and molecular feature assessment. Consequently, an urgent need exists to deepen our grasp of PDAC's genomic heterogeneity. Establishing an innovative molecular classification system is essential for precise patient stratification and informed clinical decision-making, ultimately enhancing patient outcomes. Single-cell sequencing offers advanced methods for decoding intratumoral heterogeneity by exposing the molecular attributes of diverse cellular constituents. In pancreatic adenocarcinoma (PAAD), significant heterogeneity has been documented within the immune microenvironment and in conditions of hypoxia (Peng et al. 2019; Pan et al. 2019). Comprising a varied mix of immune cells, stromal cells, and tumor cells, the tumor microenvironment (TME) is instrumental in the heterogeneous constitution of tumors and their malignant evolution (Han et al. 2021). Tumor-associated macrophages (TAMs), the predominant immune constituents in the TME, emit a range of mediators that not only dampen antitumor immunity but also foster angiogenesis, thereby intensifying cancer cell proliferation, invasion, intravasation, and eventual metastatic spread (Noy and Pollard 2014). Interactions and dynamic variations among TAMs play a crucial role in shaping the TME, ultimately determining heterogeneous clinical outcomes (Ruffell et al. 2012). Consequently, employing single-cell RNA sequencing is imperative to investigate the dynamic processes of TAM subpopulations and to develop new molecular classifications that accurately reflect the heterogeneous genomic characteristics and clinical outcomes. Incorporating single-cell sequencing data may afford a more comprehensive understanding of the complex cellular interactions within the TME, providing a foundation for developing targeted therapies based on individual patients' TME profiles.

In this context, we utilized single-cell RNA sequencing data to characterize the immunological cellular landscape and identify the trajectory genes fundamental to the kinetic evolution of neoplastic macrophages in PAAD. This study pinpointed trajectory genes instrumental in the evolution of these cells and established three heterogeneous clusters based on The Cancer Genome Atlas (TCGA)-PAAD cohort data, subsequently validated through five independent external cohorts. The three clusters displayed significant variations in prognosis, biological characteristics, genomic changes, and immune microenvironment infiltration, indicating that PAAD is a heterogeneous disease comprising different subtypes necessitating varied treatment approaches.

Furthermore, the study identified potential therapeutic agents within these clusters, suggesting the feasibility of personalized treatment strategies for PAAD. Overall, this study offers valuable insights into PAAD biology and underscores the potential for personalized treatment approaches predicated on the molecular and genetic characteristics of individual tumors.

Method

Dataset sources

We examined single-cell sequencing data (CRA 001160) from the Tumor Immune Single-cell Hub (TISCH) (Sun et al. 2021), encompassing 25 PDAC samples and ten adjacent normal samples. Bulk sequencing data were obtained from TCGA-PAAD, International Cancer Genome Consortium (ICGC)-CA, and Gene Expression Omnibus (GEO) (GSE21501, GSE62452, GSE71729, and GSE57495). Transcriptome and clinical data were concurrently retrieved from these platforms. The transcriptome data underwent logarithmic transformation, with transcripts per million (TPM) values increased by 1. Additionally, protein expression profiles of the published dataset by Cao et al. (2021) were downloaded from Genomic Data Commons.

Processing of single-cell RNA sequencing data

The single-cell sequencing analysis was performed using the "Seurat" package, as described in the Tumor Immune Single-cell Hub (http://tisch.comp-genomics.org/documentation/). Further analysis focused on myeloid cells, starting with the calculation of each gene's standard deviation across various cells. This step involved the use of the top 15 principal components and the 2000 most variable genes. The ScaleData function (Satija et al. 2015) was then applied in a regression analysis to mitigate the influence of mitochondrion-derived unique molecular identifier (UMI) counts. Subsequently, the RunHarmony function from the Harmony R package was employed, integrating multiple samples into a unified analysis using default parameters and the top 15 principal components (Korsunsky et al. 2019).

Subsequently, the FindClusters function (resolution = 0.4) in Seurat was utilized to discern primary cell clusters, later visualized via 2D uniform manifold approximation and projection (UMAP), a prevalent algorithm for data dimensionality reduction. Given that SingleR and CellMarker (Zhang et al. 2019) could only classify clusters into fundamental types, we consulted prior single-cell RNA sequencing (scRNA-seq) analyses to assign more precise subtypes to each cluster (Aran et al. 2019). The Monocle 2 package facilitated the pseudo-temporal analysis of individual

cells (Qiu et al. 2017), with the resulting cellular trajectories arranged in pseudo-time, displaying multiple branches. Genes discerned along these trajectories were then incorporated into subsequent analyses.

Cell–cell communication analysis

To analyze cell–cell communication among various cell types within the tumor microenvironment, we utilized Cell-PhoneDB (http://www.cellphonedb.org/) (Efremova et al. 2020). The normalized expression matrix derived from Seurat facilitated the computation of the mean expression for each receptor–ligand pair, entailing the calculation of the mean of average expressions of the receptor and ligand across respective clusters. Furthermore, the p value was employed to discern the cell type specificity of the interaction.

Function analysis for differently expressed genes

Differentially expressed genes (DEGs) were identified using Seurat's FindMarkers function, with a cutoff threshold set at an adjusted p value of less than 0.05. To elucidate the characteristics of these DEGs, we engaged the fgsea (v1.11.1, R package), a swift preranked gene set enrichment analysis (GSEA) method to perform functional enrichment analysis on the DEGs (Zyla et al. 2019). Consequently, hallmark gene sets from the Molecular Signatures Database were applied to assess each set of differentially expressed genes. GSEA was undertaken to identify gene sets markedly enriched within each distinct cell cluster, focusing on those with an adjusted p value of less than 0.05, denoting their significance.

Cluster identification via non-negative matrix factorization

The NMF package facilitated the execution of a non-negative matrix factorization (NMF) algorithm aimed at discerning molecular clusters through matrix factorization and iterative computations (Brunet et al. 2004).

In the pursuit of clinically contextualizing these findings, genes related to the trajectory underwent scrutiny through univariate Cox regression, leading to the isolation of prognostically pertinent genes regarded as potential candidate genes. Subsequent consensus clustering was achieved using the non-negative matrix derived from these genes, considering potential factorization ranks from 2 to 9, 100 iterations, and employing the "lee" method. To ascertain the optimal rank, the cophenetic coefficient was applied, and the robustness of clustering patterns was evaluated using the silhouette statistic. The ideal factorization rank was identified at the point where the cophenetic correlation coefficient commenced its decline (Brunet et al. 2004). The extent of

similarity between a sample and its corresponding cluster positively correlates with the silhouette coefficient's magnitude, a higher value of which indicates a more precise alignment with its designated cluster (Lovmar et al. 2005).

## Weighted gene co-expression network analysis

The weighted gene co-expression network analysis (WGCNA) approach aims to uncover the connections between gene modules and phenotypes through the analysis of gene co-expression networks (Langfelder and Horvath 2008; Guo et al. 2022). To pinpoint signature genes within distinct clusters, the WGCNA package was employed. After removing outlier samples, a gene co-expression network was established using the top 5000 significant genes. This network was subsequently converted into a scale-free network through the selection of an appropriate soft-thresholding power, "b." The ensuing weighted adjacency matrix was harnessed to generate a topological overlap matrix (TOM), reflecting the interconnectedness among the network's constituents. To quantify the gene dissimilarity, the value of 1-TOM was utilized. Moreover, the dynamic tree algorithm facilitated the identification of diversely colored gene modules. Three modules demonstrating the most potent correlation with phenotypic traits were isolated by examining the nexus between their module eigengenes and the phenotypes in question. In each module, genes with a correlation coefficient greater than 0.4 with that module are identified as characteristic genes. Consequent to this, the genes instrumental in dictating these modules' distinctive co-expression configurations were extracted for in-depth analysis.

Nearest template prediction validation

The nearest template prediction (NTP) method offers a robust approach for assessing the confidence level of class predictions for individual patients (Hoshida 2010). To strengthen the consistency and integrity of the identified clusters, we utilized the NTP algorithm, which was implemented using the CMScaller software package, to corroborate results across multiple cohorts, employing various platforms. The unique genes utilized in the NTP approach were derived from signature genes identified within the respective modules.

Somatic mutation and copy number variation analysis

Incorporating mutation and copy number variation (CNV) data contributes to a deeper understanding of genomic variability. The maftools package is instrumental in visualizing somatic variations, including single nucleotide polymorphisms (SNPs), tumor mutational burden (TMB), and mutation frequency
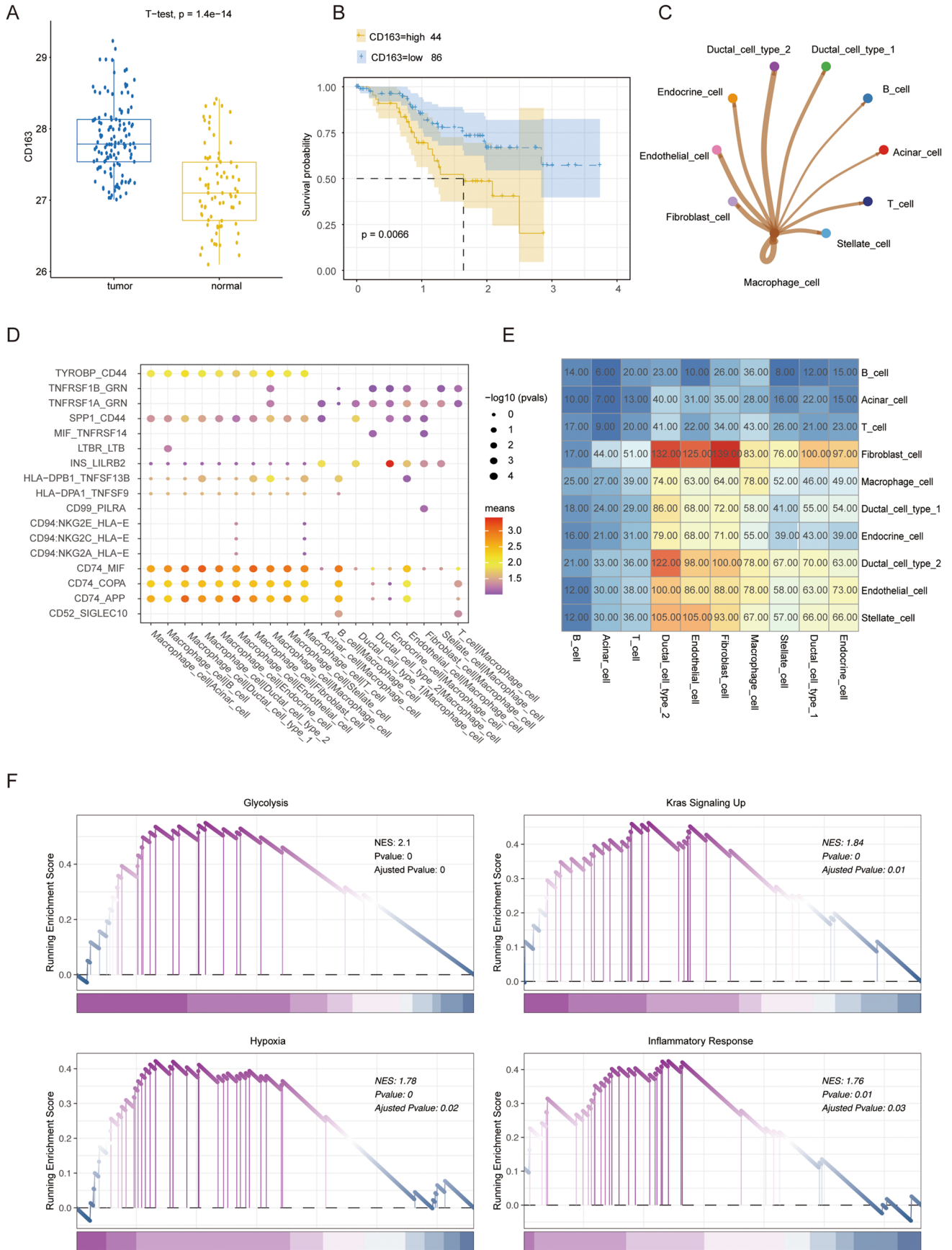
◄**Fig. 1** Cell–cell communication analysis revealed intimate crosstalk between macrophages and other cells in tumor microenvironment. **A** Box plots displaying the protein expression of CD163 in pancreatic tumor and pancreatic tissue. **B** Kaplan–Meier survival curve presenting the overall survival of patients with PDAC. The patients were divided into two groups according to the protein expression of CD163 in PDAC tissues. **C** The Scallop plot showing the macrophage interactions among ten major cell types as in pancreatic tissues predicted by CellPhoneDB. The width of the edge represents the number of significant ligand–receptor pairs between the two cell types. **D** The dot plot generated by CellPhoneDB showing potential ligand–receptor pairs associated with all detected cellular types in pancreatic tissue. Dots colored by mean expression of ligand–receptor pair between two clusters and dot size proportional to the value of $-\log10$ (*p* value). **E** Heatmap displaying the number of potential ligand–receptor pairs in each cell. **F** Gene set enrichment analysis of differentially expressed genes from PDAC versus control group

(Mayakonda et al. 2018). Each patient's TMB was calculated by evaluating the total number of non-synonymous mutations per megabase. Fisher's exact test was applied to discern variations in mutated genes between active immune and active stromal compartments. For significant amplifications or deletions in copy number analysis, GISTIC_2.0 was employed (Mermel et al. 2011). The extent of copy number alterations, at both focal and arm levels, defined the overall burden of copy number diminution or amplification. Notably, Submap, MutSigCV_1.41, and GISTIC_2.0 are accessible via GenePattern at https://cloud.genepattern.org.

The assessment of immune cell infiltration and immunotherapy

CIBERSORT is an established method for the analysis and estimation of immune cell infiltration, evaluating the proportions of diverse cell subtypes in mixed cell samples through the examination of RNA-seq expression profiles. It discerns 22 unique immune cell subtypes, with LM22 showcasing gene expression signatures spanning naive and memory B cells, seven types of T cell types, NK cells, plasma cells, and myeloid subsets. Utilizing CIBERSORT (Newman et al. 2015), we assessed these immune cell subtypes across various clusters and considered the hypothesis regarding the types of immune cells to be accurate and statistically significant for further analysis with a significance level of *p* value < 0.05. Furthermore, we utilized the ESTIMATE (Yoshihara et al. 2013) algorithm to generate stromal and immune scores, which estimate the levels of infiltrating stromal and immune cells in PAAD tissues as well as tumor purity based on the expression profiles. Discrepancies in tumor purity, stromal, and immune scores among the three clusters were analyzed using the Wilcoxon rank-sum test. Antigen presentation potential was appraised based on the expression profiles of nine human leukocyte antigen (HLA) molecules (Huang and Fu 2019). To assess immunotherapy efficacy within diverse clusters, we employed two primary methods: T cell inflammatory signatures (TIS), composed of 18 inflammatory genes, was assessed using the ssGSEA algorithm, where higher scores indicated a more favorable response to PD-1 blockade (Ayers et al. 2017). Submap was employed to gauge the similarity in expression profiles between PAAD patients and immunotherapy patients, providing insights into clinical response similarity (Hoshida et al. 2007). IMvigor210 cohort was used to further investigate the significance of immunotherapy in different clusters. The accuracy of immunotherapeutic prediction was assessed using a receiver operating characteristic (ROC) curve.

The cluster-associated immunotherapy score (CAIS) establishment

WCGNA identified genes correlating with C1 patients. Concurrently, least absolute shrinkage and selection operator (LASSO) analysis selected dependable predictors (Li et al. 2022). The CAIS for each patient in the IMvigor210 cohort was determined using the following formula: risk score = Σ coefficient mRNAn × expression level mRNA. Based on the Z-score, patients within the IMvigor210 cohort were categorized into low-risk (Z-score < 0) and high-risk (Z-score ≥ 0) groups.
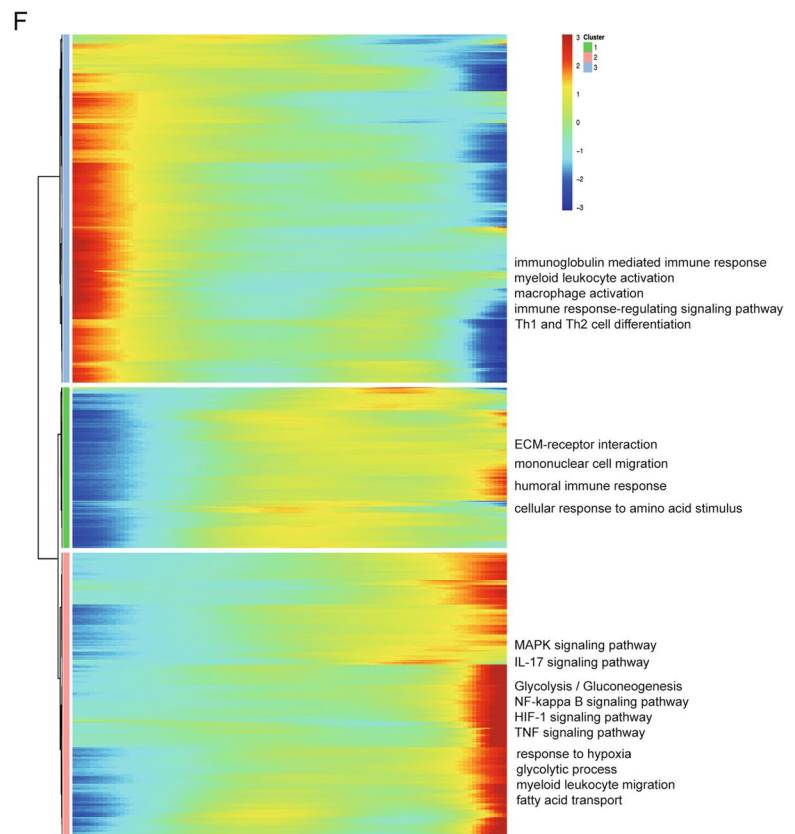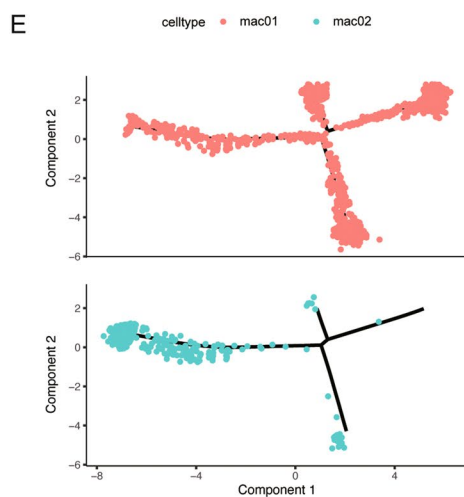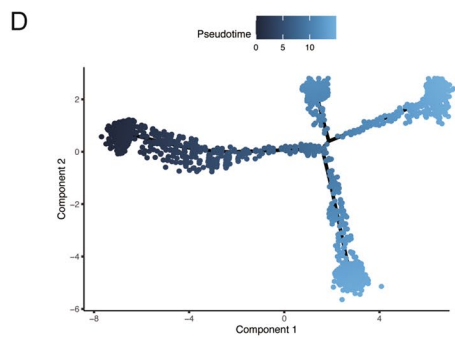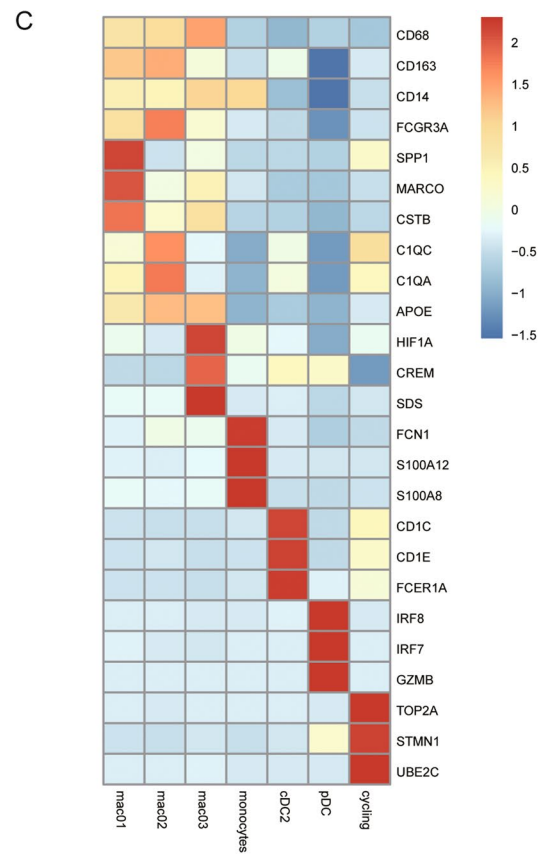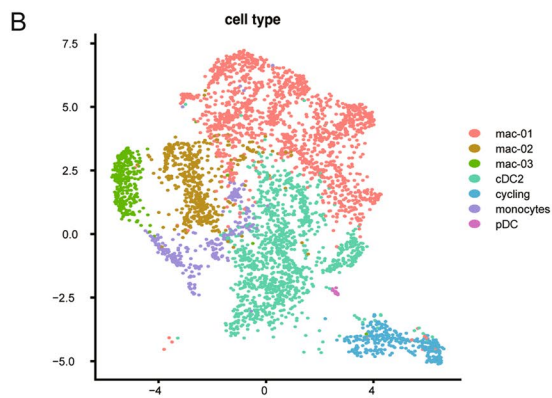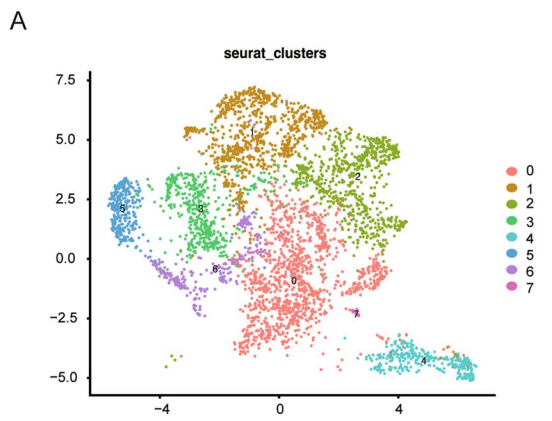
The log-rank test compared survival disparities between groups, while Kaplan–Meier survival curves illustrated the OS for each. The "survival," "survminer," and "ROC" R packages were employed to perform ROC curve analysis.

Clinical treatment evaluation

Drug response predictions, predicated on gene expression data, were conducted using the linear ridge regression model within the pRRophetic package (Geeleher et al. 2014). Through this package, we computed the half-maximal inhibitory concentration (IC50) of clinical tissues, subsequently screening for potential therapeutic agents exhibiting the lowest IC50 values across distinct clusters.

Cell culture and transfection

PATU-8988 and PANC-1 cells were cultivated in a DMEM medium (Gibco, USA), supplemented with 10% fetal bovine serum (FBS) (Gibco, USA), and maintained at 37 °C in a 5% CO_2 atmosphere. Three si-NR1H3s were obtained from Hanheng in Shanghai, China. PATU-8988 and PANC-1 cells were seeded within six-well plates, and small interfering RNA (siRNA) transfection experiments were executed upon reaching an 80% cellular density. Transfection was rigorously performed following the manufacturer's instructions, with protein extraction occurring 72 h post-transfection for ensuing analysis. The siRNA sequences are provided in Table S2.

A

seurat_clusters

- 0
- 1
- 2
- 3
- 4
- 5
- 6
- 7

B

cell type

- mac-01
- mac-02
- mac-03
- cDC2
- cycling
- monocytes
- pDC

C

CD68
CD163
CD14
FCGR3A
SPP1
MARCO
CSTB
C1QC
C1QA
APOE
HIF1A
CREM
SDS
FCN1
S100A12
S100A8
CD1C
CD1E
FCER1A
IRF8
IRF7
GZMB
TOP2A
STMN1
UBE2C

mac01 mac02 mac03 monocytes cDC2 pDC cycling

D

Pseudotime

0   5   10

Component 2

Component 1

E

celltype ● mac01 ● mac02

Component 2

Component 1

F

Cluster
1
2
3

immunoglobulin mediated immune response
myeloid leukocyte activation
macrophage activation
immune response-regulating signaling pathway
Th1 and Th2 cell differentiation

ECM-receptor interaction
mononuclear cell migration
humoral immune response
cellular response to amino acid stimulus

MAPK signaling pathway
IL-17 signaling pathway

Glycolysis / Gluconeogenesis
NF-kappa B signaling pathway
HIF-1 signaling pathway
TNF signaling pathway

response to hypoxia
glycolytic process
myeloid leukocyte migration
fatty acid transport

◀**Fig. 2** The dynamics of macrophages during PDAC progression. **A** UMAP plot of macrophages cells presenting eight clusters. **B** UMAP plot of macrophage cells presenting seven cell types. **C** Top three marker genes of seven macrophage cell types identified in heatmap. **D** Differentiation trajectory of TAMs in PDAC, with a color code for clusters. **E** Trajectory of macrophages along pseudotime in a two-dimensional space. Each point corresponds to a single cell. **F** Heatmap showing the dynamic changes of gene expression along pseudotime. The differentially expressed genes were clustered hierarchically into three groups, and the representative enriched pathways of each group are shown

Construction of stable transfer cell lines

Utilizing a lentiviral vector, we generated stable cell lines exhibiting diminished NR1H3 expression levels. Post-verification of siRNA knockdown efficiency, siRNA-2 oligonucleotides were integrated into the short hairpin RNA (shRNA) construct. Scramble shRNA served as a negative control. After 48 h of virus infection, the cells were supplemented with conventional growth medium and cultured under normal conditions. The cells were screened under maintenance with 2 μg/mL puromycin for 14 days to establish stable cell lines.

Western blot assay

Cells were harvested and thoroughly lysed with RIPA lysis buffer with an inclusion of 1% phenylmethanesulfonyl fluoride. Protein samples were combined with a loading buffer and subjected to a 5-min boil in a 100 °C water bath. Following SDS-PAGE, proteins were transferred onto a polyvinylidene fluoride (PVDF) membrane. Then, PVDF was incubated in 5% skimmed milk powder for 2 h, and the primary antibody was incubated at 4 °C overnight after washing. After the PVDF membrane was washed three times again, the secondary antibody (1:5000 dilution) was incubated for 1 h. Subsequent to the application of enhanced chemiluminescence (ECL) reagent, band detection was performed using a Bio-Rad imaging device. The NR1H3 antibody (14351–1-AP) was purchased from Proteintech. The GAPDH antibody (10494–1-AP) was purchased from Proteintech.

Cell proliferation, apoptosis, and migration assays

Cell proliferation was examined using the EdU Cell Proliferation Assay Kit (C0075S; Beyotime, China). Cells were labeled with 50 mM EdU for 1 h. Post-staining, cells in suspension were analyzed and detected via flow cytometry. Cells were seeded into six-well plates, subsequently stained using an apoptosis kit from Multi Sciences, China (AP107-100). Post-staining, cells in suspension were analyzed and detected via flow cytometry. The cells were subsequently suspended in a DMEM medium without FBS. A total of 2×10^4 cells were seeded in the upper chamber of transwell units, devoid of Matrigel, within 24-well plates (Millipore, Billerica, MA, USA). The lower compartment was supplied with a DMEM medium containing 10% FBS. Following a 48-h co-culture, cells were fixed with methanol. Cells in the upper chamber were eliminated, and the residual cells were treated with a 2% crystal violet stain. Subsequently, microscopic images were captured, and the migrated cells were quantified. Three biological replicates were performed, and the result was subjected to statistical analysis.

Colony formation assay

Transfected PATU-8988 and PANC-1 cells were seeded at a density of 500 cells/6-cm dish (Millipore, Billerica, MA, USA). After a 14-day incubation, cells were stained with 1% crystal violet. Colonies exceeding 50 mm in diameter were quantified using ImageJ. Three biological replicates were performed, and the result was subjected to statistical analysis.

Statistical analysis

The "survival" package facilitated the Kaplan–Meier and Cox regression analyses. Survival statistics for categorical variables were contrasted using the log-rank test. We utilized multivariate Cox regression analysis to determine the hazard ratio and confirm the independent significance of various characteristics. Pearson's correlation analysis was conducted to explore the relationship between two continuous variables. The Kruskal–Wallis test was employed to assess differences among three groups. For predicting binary categorical variables, we utilized the pROC package to generate the ROC curves. All data preprocessing, statistical analysis, and visualizations were performed using R v4.1 software. p values less than 0.05*, less than 0.01**, less than 0.001***, or less than 0.0001**** were considered statistically significant. All experiments were conducted three times. Data analysis was performed using Statistical Products and Services Solutions (SPSS) version 22.0 (IBM, Armonk, NY, USA), with experimental data expressed as mean ± standard deviation. Differences between groups were assessed using a t test.

Result

M2 macrophages are associated with a poor prognosis in pancreatic cancer

To discern the relationship between M2 macrophages and survival outcomes in pancreatic cancer, we harnessed pancreatic proteome data. Our findings indicated elevated CD163 protein expression in cancerous tissue relative to
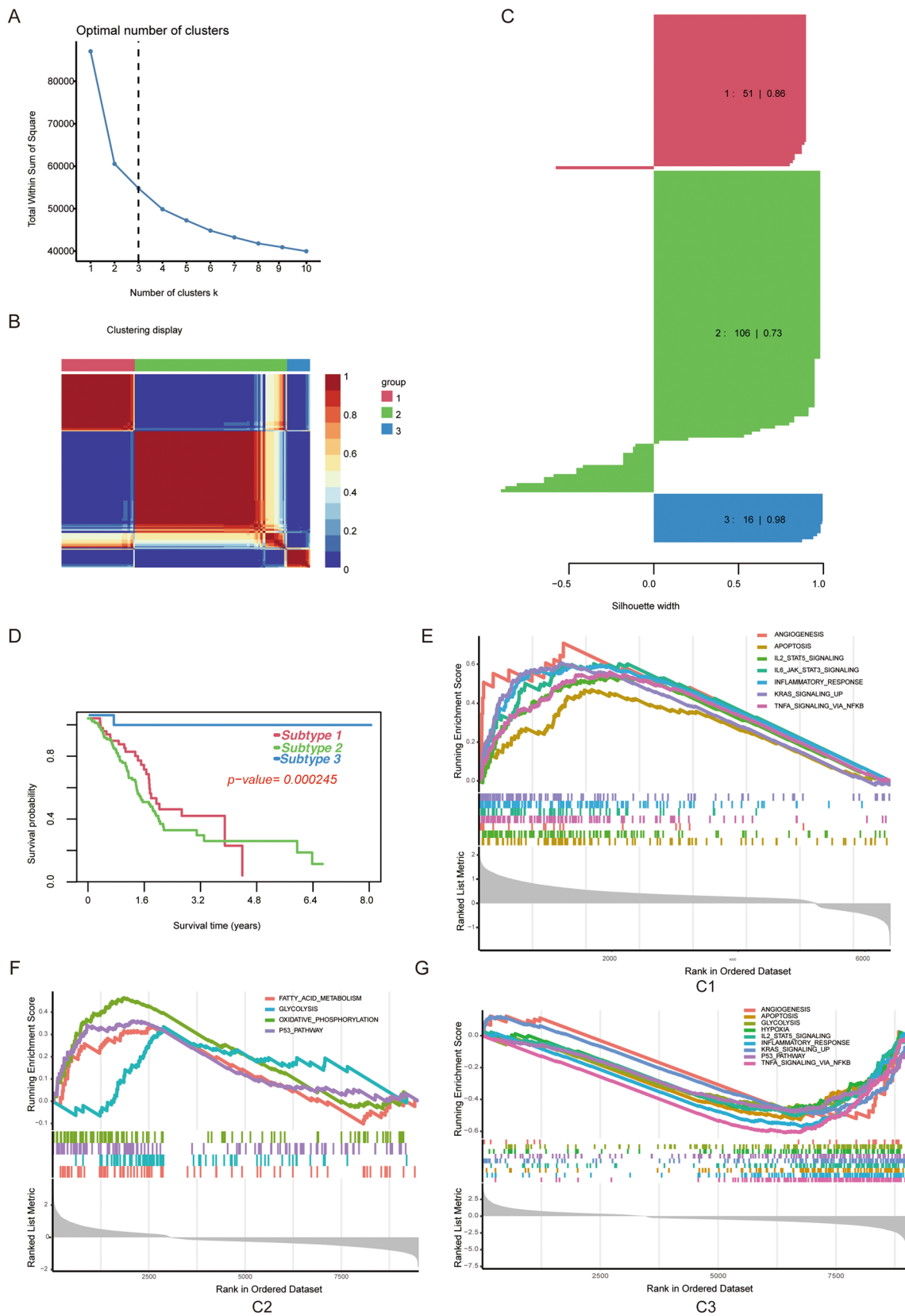
◀**Fig. 3** Development of three molecular clusters with heterogeneous clinical outcomes by non-negative matrix factorization (NMF) analysis. **A** The optimal rank was 3 as the cophenetic coefficient started firstly decreasing. **B** Consensus map of NMF clustering results in The Cancer Genome Atlas-Pancreatic Adenocarcinoma (TCGA-PAAD) cohort. **C** Silhouette statistic of three heterogeneous clusters. **D** Kaplan–Meier curves of overall survival according to the three clusters in the TCGA-PAAD cohort. **E–G** Comparison of pathway between the different molecular clusters. The pathways were associated with hallmark gene sets derived from the Molecular Signatures Database (MSigDB)

normal pancreatic tissue (Fig. 1A). Furthermore, tissues demonstrating high CD163 protein expression correlated with unfavorable prognoses (Fig. 1B). Recognizing that macrophages infiltrating pancreatic tumors bear functional resemblance to M2 macrophages, we also conducted a single-cell transcriptome analysis to explore the heterogeneity of macrophages within the pancreatic cancer tumor microenvironment.

scRNA profiling of pancreatic tissue

We analyzed the scRNA sequencing dataset from TISCH (http://tisch.comp-genomics.org), encompassing 57,443 pancreatic cells, of which 15,467 were from normal pancreatic tissue and 41,976 from pancreatic tumors. This original data was sourced from the Genome Sequence Archive project PRJCA001063. Expression characteristics for each sample are detailed in Fig. S1A. The nCount RNA, indicative of the number of UMIs, displayed a positive correlation with nFeature RNA, representative of the number of genes, exhibiting a correlation coefficient of 0.82 (Supplementary Fig. S1E). Additional quality control results are shown in Supplementary Fig. S1B, C, D, and F. Clinical data pertaining to the pancreatic tissues are tabulated in Table S1.

A UMAP analysis facilitated the visualization of ten distinct cell types, as illustrated in Fig. S2A and B. Utilizing marker genes, these clusters were aligned with known cell lineages from prior research, as shown in Supplementary Fig. S2C and D. The distribution of various cell types across patients is depicted in Supplementary Fig. S2D and E.

Cell–cell communication analysis revealed intimate crosstalk between macrophages and other cells in tumor microenvironment

CellPhoneDB, an interactive web application that discerns cellular interactions from ligand–receptor signaling databases, was employed to investigate the interplay between macrophages and other immune cells in the tumor microenvironment. The Scallop plot identified broadcast ligands and revealed extensive communication through cognate receptors (Fig. 1C). We observed interactions between macrophages

and other cell types involved in chemochemical, immuno-costimulatory, immunosuppressive, and immune response signaling, such as Th1, Th2, and Th17 (Fig. 1D). Further analysis revealed that macrophages preferentially promoted tumor cell proliferation and chronic inflammation in pancreatic cancer by increasing CD74-MIF ligand–receptor pairs. This finding suggests a potential role for macrophages in immunosuppression and tumor cell proliferation promotion via CD74 secretion. Furthermore, in comparison to other cells, macrophages showed a higher number of receptor–ligand pairs and were more evenly distributed (Fig. 1E).

To elucidate the role of infiltrating macrophages in pancreatic cancer, we executed a GSEA pathway enrichment analysis. The results indicated that macrophages infiltrating cancerous tissue were significantly enriched in pathways related to glucose metabolism, hypoxia, Kras, and inflammatory response compared to normal pancreatic tissue (Fig. 1F). These pathways are historically linked with tumor cell proliferation and invasion. Our findings underscore the necessity for more comprehensive studies on infiltrating macrophages to identify TAMs that contribute to tumor growth and immunosuppression.

The dynamics of macrophages during PDAC progression

To probe further into macrophage heterogeneity, we conducted dimensionality reduction and clustering analyses on macrophages sourced from PDAC patients. This approach distinguished eight unique macrophage subpopulations (Fig. 2A). Subsequently, we meticulously analyzed the gene signatures of all 13 myeloid subsets, which were identified in a prior study (Zhang et al. 2020). An in-depth analysis yielded seven distinct subclusters (Fig. 2B), including SPP1 (Mac-01) and C1QC (Mac-02) tumor-associated macrophages, HIF1A-expressing macrophages, monocyte-like cells, DC2 and pDC dendritic cells, and cycling cells. The expression of distinct marker genes for each myeloid cluster was manually annotated (Fig. 2C). Mac-01, referred to as tumor-associated macrophage–SPP1 (TAM–SPP1), is defined by elevated expression of macrophage receptor with collagenous structure (MARCO), cystatin B (CSTB), and SPP1, known promoters of oncogenic activity in pancreatic cancer (Yoshihara et al. 2013). Conversely, Mac-02, indicative of an inflammatory subset, exhibited robust expression of inflammation-associated genes C1QC and C1QA, hence its designation as tumor-associated macrophage–C1QC (TAM–C1QC). TAMs, forming the bulk of myeloid cells in PAAD tissue, aggregated into subgroups marked by high SPP1 and C1QC gene expression. To trace the temporal evolution of TAMs from normal tissue to the tumor milieu, we employed monocle2, constructing a pseudotime trajectory of TAM states (Fig. 2E). This analysis revealed an ascent in TAM–SPP1 cells with tumor progression, suggesting a potential adverse prognosis
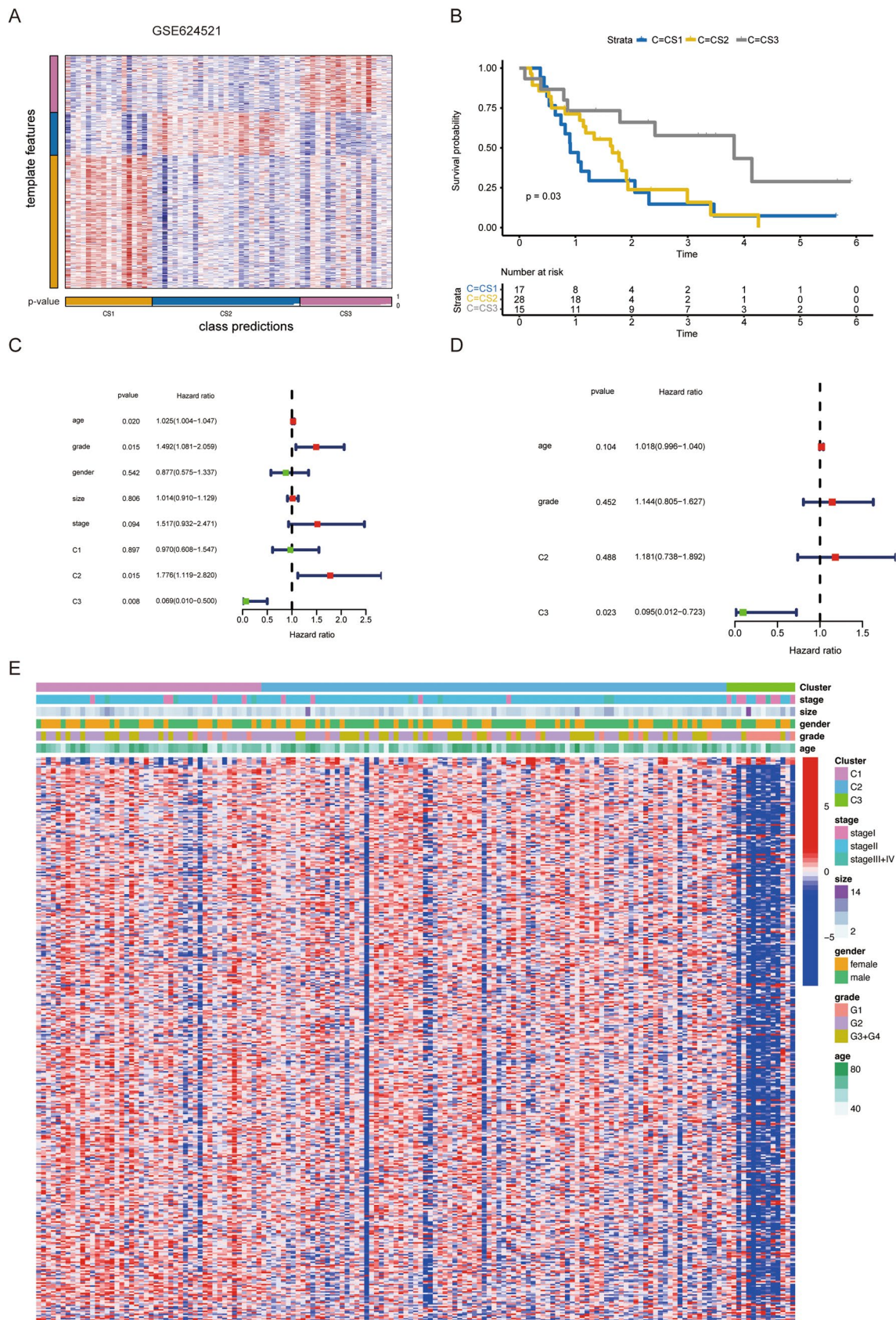
◀**Fig. 4** Validation and clinical features of three heterogeneous clusters. **A** Heatmap of the expression level of the template feature between the three clusters in the GSE62452 cohort. **B** Kaplan–Meier curves of overall survival (OS) according to the three clusters in the GSE62452 cohort. **C** Univariate Cox regression of OS in TCGA-PAAD. **D** Multivariate Cox regression of OS in TCGA-PAAD. **E** Heatmap depicted the correlation between the patterns and different clinicopathological features

(Fig. 2D), while TAM-C1QB cells diminished as the tumor advanced (Fig. 2D). Although traditionally classified into M1 (pro-inflammatory) or M2 (pro-tumorigenic) categories, our examination of M1 and M2 signature gene expressions within TAMs revealed that, despite C1QC + TAMs showing elevated M1 signature expression relative to SPP1 + TAMs, both subsets exhibited prominent M2 signature expression, challenging their distinction based on the standard M1/M2 classification.

Upon analyzing the single-cell transcriptomes of TAMs across their trajectory, we identified 1000 genes demonstrating notable expression shifts. These genes fell into three distinct expression patterns: the first category showed diminishing expression levels through the trajectory, correlating with immune response modulation, and Th1 and Th2 cell differentiation pathways. The second category, elevated during the initial phase of tumor infiltration, was implicated in mononuclear cell migration. The third set, activated in the tumor microenvironment's advanced stage, was involved in diverse pathways, encompassing hypoxia, MAPK signaling pathway, IL-17 signaling pathway, glycolysis, NF-kappa B signaling pathway, HIF-1 signaling pathway, and TNF signaling pathway (Fig. 2F).

The identification of three molecular clusters

We utilized genes associated with the trajectory to identify potential prognostic candidate genes. We applied the NMF method to discern heterogeneous molecular clusters based on these genes' expression. Figure 3A illustrates that the optimal number of clusters was 3, determined by the significant rapid decrease in the cophenetic coefficient. The consensus matrix, presented in Fig. 3B, further substantiates that the three clusters facilitate ideal stratification. To assess the robustness of these molecular clusters, we utilized profile statistics and further classified the samples by their profile width (Isella et al. 2017). Consequently, samples with a positive profile width were segregated into three distinct stable clusters, as illustrated in Fig. 3C. To advance clinical applicability, we delved into the prognostic implications of these clusters. The OS of C2 and C1 is bleak, while that of C3 is favorable ($P < 0.05$), as revealed in Fig. 3D. We undertook GSEA analysis with hallmark gene sets to elucidate the distinct biological attributes of each cluster. C1 was significantly enriched in proliferative pathways, encompassing angiogenesis and apoptosis. C2 was predominantly associated

with metabolic pathways, including fatty acid and glycolysis. It is worth noting that there exist significant connections between the tumor immune microenvironment and metabolic activities, including fatty metabolism and bile glycolysis metabolism in the pancreas. In contrast, C3 was markedly characterized by suppressed proliferative pathways (Fig. 3E–G).

The characteristic genes of three clusters and the nearest template prediction verify three heterogeneous clusters

The characteristic genes of the three clusters were identified using the WGCNA package. Initially, outlier samples were removed, and then the remaining samples were clustered (Supplementary Fig. S3A). Employing a soft-thresholding power of 6, the scale-free R_2 value was 0.9, indicating the development of a scale-free network (Supplementary Fig. S3A). The association between each module and the three clusters was subsequently examined to ascertain their correlation. Significantly, the black, turquoise, and yellow modules were most strongly correlated with C1, C2, and C3, respectively, as depicted in Supplementary Fig. S3B. These modules were then utilized to identify the characteristic genes of each cluster. The validity and integrity of the gene module assembly were affirmed by the correlation coefficients between gene significance and module membership, showcased in Supplementary Fig. S3C–E. Genes within each module were designated as characteristic genes and are detailed in Supplementary Table S3. To verify the reliability and stability of the three clusters, NTP analysis, predicated on signature gene expression from the characteristic genes, was performed. The NTP method was deployed to assess and validate the prediction confidence for each patient, drawing upon five distinct cohorts from various platforms, namely GSE28735, GSE21501, GSE71729, GSE62452 (Fig. 4A), and ICGC-AU. In line with prior research, patients with a false discovery rate (FDR) of less than 0.05 were chosen for further analysis. Within the GSE62452 cohorts, the prognostic significance of the three clusters was elaborated through Kaplan–Meier analyses, the outcomes of which reaffirmed that C1 and C2 had detrimental OS rates, whereas C3 manifested the most propitious OS ($P < 0.05$), aligning with antecedent observations, as displayed in Fig. 4B. Univariate and multivariate Cox regression analyses revealed that C3 was independent prognostic indicators in the TCGA-PAAD cohorts, as shown in Fig. 4C and D. Moreover, the expression levels of the prognostic candidate genes were relatively diminished in the C3 cluster. Additionally, patients within the C3 cluster were diagnosed at earlier stages compared to other subtypes, as indicated in Fig. 4E. Collectively, the three clusters exhibited heterogeneous clinical prognoses in PAAD.
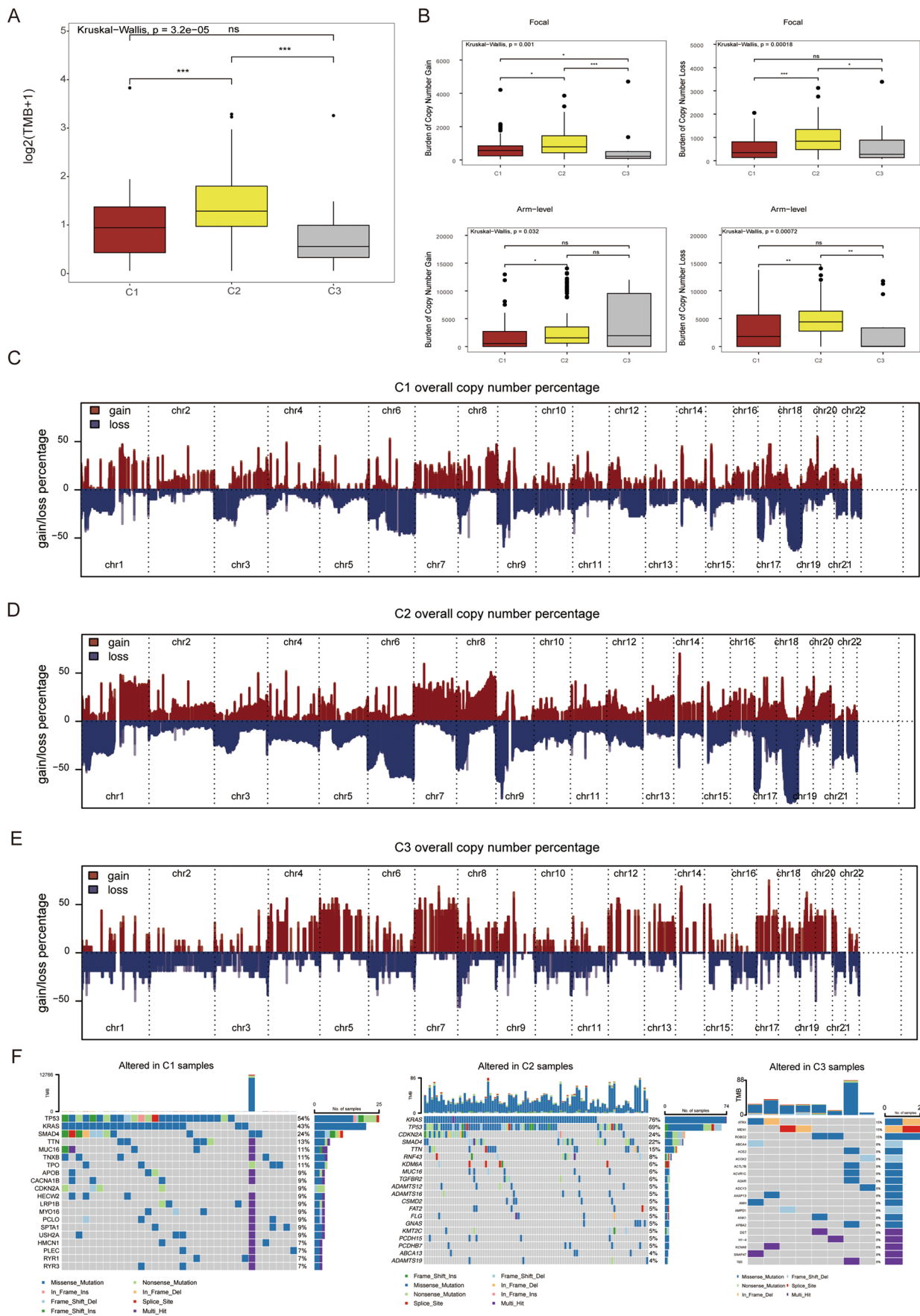
◀**Fig. 5** Genetic and copy number alterations across immune subtypes. **A** Distribution of TMB among the three heterogeneous clusters. **B** Focal and broad copy number alterations among the three heterogeneous clusters. The statistical significance of pairwise comparisons: ns, not significant; *$P > 0.05$; **$P > 0.01$; ***$P \leq 0.001$. **C–E** Copy number profiles for C1, C2, and C3 subtypes, with gains in dark red and losses in midnight blue. Gene segments are placed according to their location on chromosomes, ranging from chromosome 1 to chromosome 22. **F** The waterfall plot depicted the differences in frequently mutated genes (FMGs) of hepatocellular carcinoma among the three clusters. The right panel shows the mutation rate, and genes were ordered by their mutation frequencies

The landscape of genomic variations

Recent studies suggest a correlation between a tumor's genetic landscape and its cytolytic activity, with specific somatic mutations and high TMB linked to antitumor immunity. Utilizing genomic data from the TCGA database, we explored these underlying genomic mechanisms. Our analysis revealed that patients in the C2 cluster had significantly higher TMB values than those in the C3 group (Fig. 5A; C2 vs. C3: Wilcoxon rank-sum test, $P < 0.001$; C2 vs. C1: Wilcoxon rank-sum test, $P < 0.001$).

To delineate the genomic disparities among the three clusters, we assessed copy number alterations. We observed a higher burden of both gains and losses at the focal level, as well as gains at the arm level, in the C2 group compared to the others (Fig. 5B). The G-score distribution across all chromosomes in these subtypes is depicted in Fig. 5C–E. Notably, C1 and C2 shared similar G-scores on chromosomes 6, 8, 17, and 19 (Fig. 5C, D), potentially indicative of an adverse prognosis. In contrast, C3's G-score diverged from those of C1 and C2 (Fig. 5E), underscoring the substantial role focal copy number alterations may play in the survival discrepancies observed in pancreatic cancer.

Furthermore, we characterized the genomic variation landscape across the three heterogeneous clusters. Figure 5F illustrates the mutation frequency of the top 20 frequently mutated genes (FMGs). Among these, TP53 and KARS mutations were prominent in C1, while KARS mutations predominated in C2. Prior research has linked the progressive accumulation of gene mutations to tumorigenesis (Sanchez-Vega et al. 2018). In comparison, C3 exhibited the lowest frequency of gene mutations, implying a more favorable clinical outcome. Collectively, patients in C2 manifested notable genomic variations, suggestive of elevated genomic instability.

The assessment of immune infiltration and immunotherapy

In this study, we delved deeper into the relationship between the three heterogeneous clusters and the characteristics of the TME. Our findings showed that C1 had a higher concentration of infiltrating immune cells relative to the other clusters (Fig. 6A). Specifically, patients in the C1 group tended to have an "immune-hot" subtype, characterized by a high number of immune cells in the TME, including CD4+ T cells, CD8+ T cells, activated dendritic cells, and natural killer cells, among others ($P < 0.05$) (Fig. 6A). Moreover, patients within the C1 cluster exhibited significantly higher immune scores compared to those in other clusters (Fig. 6B). In contrast, a diminished level of tumor purity was identified in the C1 cluster relative to others (Fig. 6C).

Given the association of C1 with adverse prognoses, we postulated that immune evasion could be a contributing factor to the unfavorable outcomes observed in this group. Consequently, we initiated a detailed examination of immune checkpoints. C1 presented the most substantial expression of immune checkpoints, including CD80, CD86, PDCD1, and CD274 (PD-L1), suggesting a potential increased susceptibility to ICI therapy (Fig. 6D). For a broader perspective on immunotherapy, we applied the TIS and Submap methods to evaluate clinical efficacy across the distinct clusters. Additionally, we noted a marked escalation in the expression of HLA molecules within the C1 subset, indicating a pronounced antigen presentation capability in these patients (Fig. 6E). Predictably, the C1 cluster recorded the highest TIS score, denoting intensified immune activation and a higher propensity for a positive response to ICIs ($P < 0.01$) (Fig. 6F). Through the Submap algorithm, we discovered that the C1 subset showed greater congruence in expression profiles with individuals who had demonstrated a positive response to PD-L1 inhibitors (Bonferroni-corrected $P < 0.01$), indicating that they may derive greater therapeutic benefit from anti-PD-L1 therapy (Fig. 6G).

In conclusion, precision immunotherapy presents itself as a potentially effective strategy for patients classified within the C1 cluster. Subsequent to this, we isolated genes (ENTPD1, WIPF1, STAB1, FERMT2, RGS18, HGF, IL16, PDCD1LG2, NR1H3) indicative of C1 to formulate a metric indicative of the cluster's affinity for immunotherapy, termed the cluster-associated immunotherapy score. We normalized the CAIS using individual Z-scores, categorizing patients into high- or low-risk sectors based on these scores. Among the IMvigor210 cohort undergoing anti-PD-L1 immunotherapy, patients were stratified into either high or low CAIS strata. It was observed that individuals in the high CAIS category faced significantly briefer survival periods compared to those in the low CAIS division (Fig. 6H). The predictive reliability of immunotherapeutic effectiveness, as represented by the area under the CAIS curve, was established at 0.89 within the IMvigor210 assembly (Fig. 6I). Furthermore, the rate of positive responses to anti-PD-L1 treatment
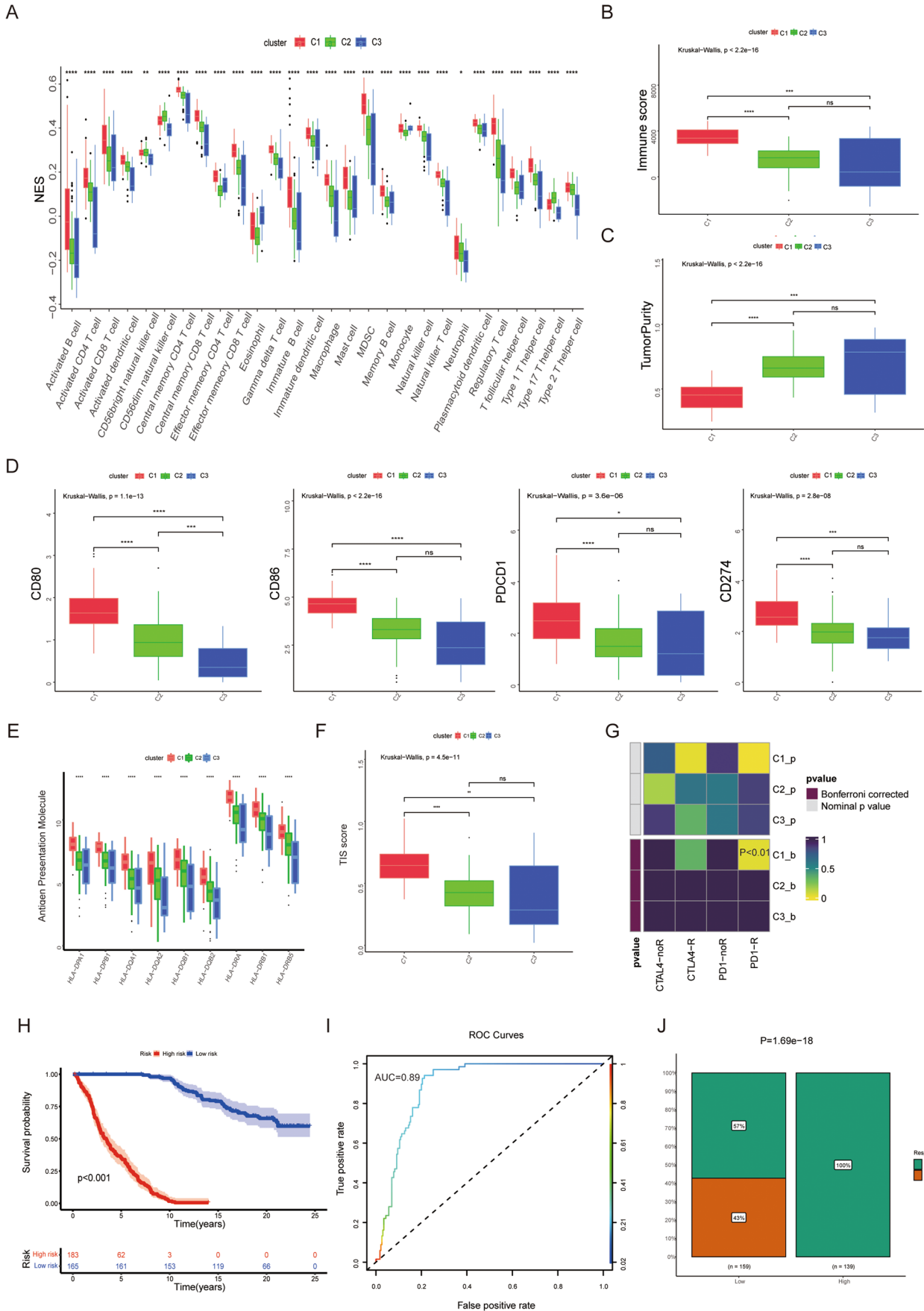
◀**Fig. 6** Immune landscape and immunotherapy responses. **A** Comparison of Infiltration abundance of immune cells evaluated by CIBERSORT analysis algorithm among the three clusters. **B** Comparison of immune score calculated through ESTIMATE among the three clusters. **C** Comparison of tumor purity score calculated through ESTIMATE among the three clusters. **D** Immune checkpoint profiles of three clusters. **E** Distribution of nine human leukocyte antigen molecular expressions among the three clusters. **F** Distribution difference of T cell inflammatory signature prediction scores among the three clusters. **G** Submap analysis manifesting that C1 could be more sensitive to anti-PD-1 therapy (Bonferroni, $P < 0.01$). **H** Kaplan–Meier curves to anti-PD-1 therapy for patients with high and low cluster-associated immunotherapy score (CAIS) in IMvigor210 cohort. CR, complete response; PD, progressive disease; PR, partial response, SD, stable disease. **I** Receiver operating characteristic curves of CAIS to predict the benefits of immunotherapy in IMvigor210 cohort. **J** Immunotherapy response ratio of CAIS in IMvigor210 cohort. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ,**** $P < 0.0001$. ns, not significant

was markedly superior in the low CAIS faction compared to the high CAIS assembly (Fig. 6J). Collectively, these insights propose CAIS as a potent predictor for immunotherapy, advocating for the inclusion of patients in the C1 classification as prime candidates for immunotherapeutic regimens.

Identification of potential therapeutic drugs

The pRRophetic package was utilized to pinpoint potential therapeutic drugs through the evaluation of various agents' sensitivity, quantified by their IC50. Patients in cluster C1 of the TCGA-PAAD cohort demonstrated a markedly superior response to docetaxel, ibrutinib, osimertinib, and entospletinib relative to other clusters, as evidenced by Fig. 7A. Moreover, the reduced IC50 values for sapitinib, lapatinib, acetalax, and ulixertinib suggested an enhanced receptiveness in patients within cluster C2 to these medications, depicted in Fig. 7B. Figure 7C inferred that vorinostat, doramapimod, sorafenib, and linsitinib might serve as viable therapeutic agents for patients in cluster C3. Consequently, the implementation of tailored therapeutic approaches contingent on patient clusters could culminate in improved clinical outcomes.

The validation of the role played by gene NR1H3 in pancreatic cancer cell lines was conducted *in vitro*

Our review of the existing literature on various genes (ENTPD1 (Novak et al. 2020), WIPF1 (Pan et al. 2018), STAB1 (Guo et al. 2017), FERMT2 (Yoshida et al. 2017), RGS18 (Liu et al. 2023), HGF (Pothula et al. 2020), IL16 (Liao et al. 2017), PDCD1LG2 (Davidson et al. 2022)) revealed that the preponderance has been confirmed to be implicated in the progression of pancreatic cancer. However, the role of NR1H3 remained unverified. Consequently, we elected to undertake functional assays focusing on NR1H3.

Transfection of three si-NR1H3 molecules into PATU-8988 and PANC-1 cell lines revealed that si-RNA2 manifested an inhibitory impact on NR1H3 expression (Fig. 8A, B). Following the suppression of NR1H3 in PATU-8988 and PANC-1 cell lines, a notable decline in the proliferation rate of pancreatic cancer cells was observed (Fig. 8C, D). Additionally, clonogenic assays indicated a significant diminution in the colony-forming ability of both the PATU-8988 and PANC-1 cell lines post-NR1H3 knockdown (Fig. 8E, F). In subsequent transwell assays, the abrogation of NR1H3 considerably curtailed the migratory capabilities of pancreatic cancer cells in both tested cell lines (Fig. 8G, H). Ultimately, the knockdown of NR1H3 markedly catalyzed apoptosis in both the PATU-8988 and PANC-1 cell lines (Fig. 8I, J).

Discussion

In this study, we harnessed advanced technology to delve into tumor heterogeneity, employing single-cell RNA sequencing analysis to discern distinct macrophage subpopulations and elucidate their dynamic intricacies. Subsequently, the NMF algorithm was applied to identify diverse molecular subtypes, and the WGCNA algorithm was used to isolate unique genes. Multiple assessment metrics facilitated the recognition of three stable subpopulations within the TCGA-PAAD cohort. Comprehensive analyses explored prognostic implications, uncovering that C3 possesses a favorable prognosis and potential as an independent prognostic indicator. An extensive examination of genomic variation, immune infiltration, and immunotherapy discrepancies across the subpopulations pinpointed prospective therapeutic agents beneficial for each subtype. Integrating these insights, we determined that patients in cluster C2 experienced the most significant genomic modifications, whereas those in cluster C1 might demonstrate an improved response to immunotherapy.

Our initial observation linked heightened CD163 protein expression in pancreatic cancer to an adverse prognosis. This finding prompted further investigation into macrophage diversity within the tumor microenvironment, utilizing single-cell sequencing. Cluster analysis pinpointed tumor-associated macrophages, particularly TAM–SPP1 and TAM–C1QC. We subsequently confirmed these macrophage subtypes in pancreatic cancer samples via immunohistochemistry and immunofluorescence techniques. Employing trajectory analysis, we decoded the temporal dynamics of tumor-associated macrophages. Moreover, we conducted an in-depth investigation into the functional role of one of the identified hallmark genes, specifically NR1H3, within the context of PAAD tumorigenesis. Our research findings indicate that NR1H3 plays a pivotal role in the pathogenic
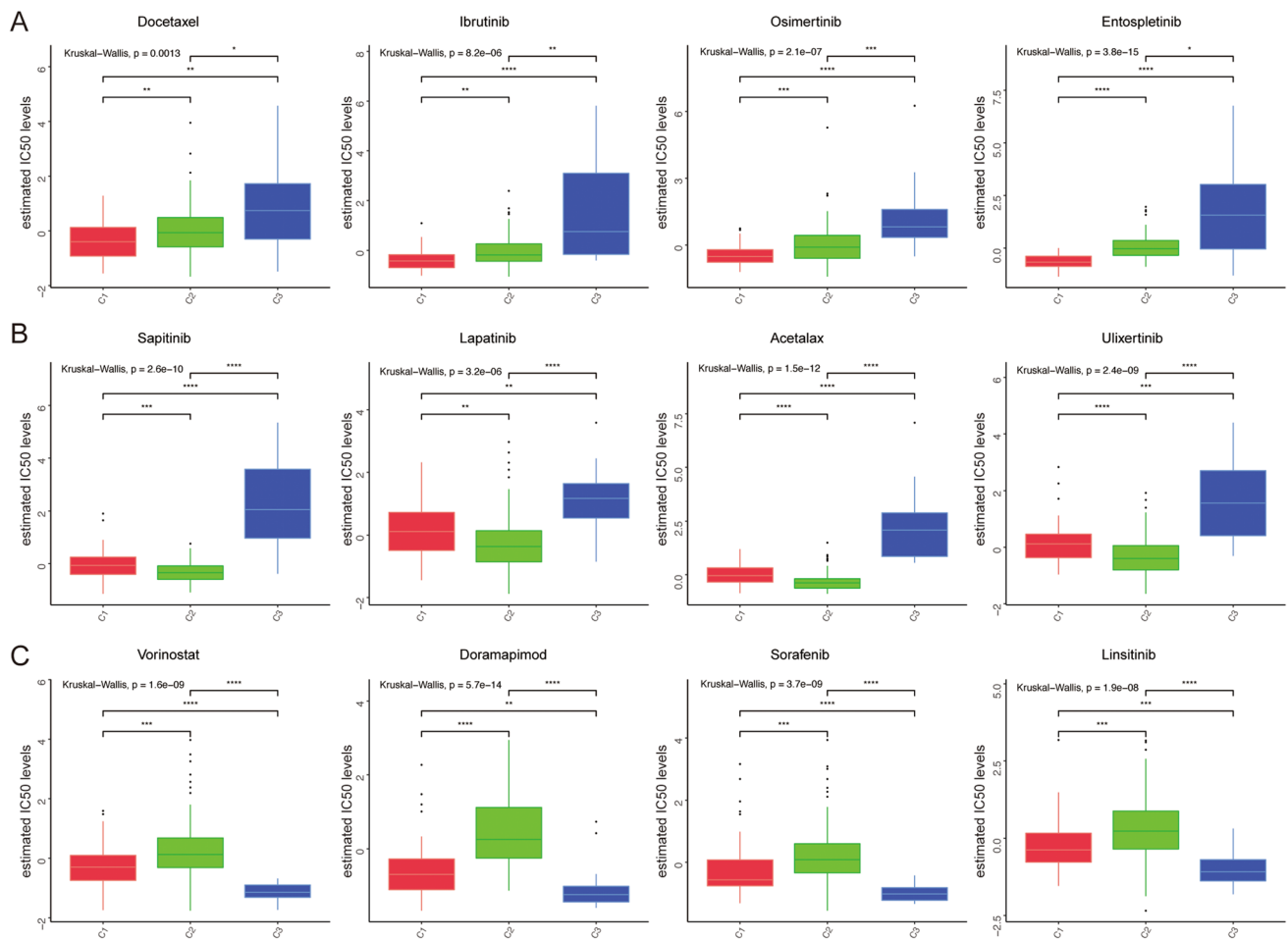
Fig. 7 Box plots depicted the differences in the estimated IC50 levels. **A** Docetaxel, ibrutinib, osimertinib, and entospletinib. **B** Sapitinib, lapatinib, acetalax, and ulixertinib. **C** Vorinostat, doramapimod, sorafenib, and linsitinib among the three heterogeneous clusters. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$

mechanisms of PAAD, exerting control over the proliferation and migratory behavior of malignant cells.

Cancer progression is intricately linked to the functions of TAMs, key constituents of the tumor microenvironment. TAMs are bifurcated into two categories based on their polarization states: pro-inflammatory (M1) or anti-inflammatory (M2) TAMs. In colon cancer single-cell studies, TAMs were classified as C1QC TAMs and SPP1 TAMs. SPP1 TAMs showed enrichment in pathways associated with cell migration, ECM–receptor interaction, and tumor angiogenesis, whereas C1QC TAMs were abundant in pathways related to complement pathway activation and antigen processing and presentation (Zhang et al. 2020). The study's findings revealed that SPP1 + TAMs could engage in paracrine signaling with other macrophage cells through SPP1 secretion and binding to the CD44 cell surface receptor. This finding aligns with a previous research identifying CD44 as the receptor for SPP1-mediated cancer metastasis (Wei et al. 2021).

In comparison to Liu et al.'s (2022) research, we constructed a molecular subtyping based on tumor-associated macrophages in pancreatic cancer. Notably, anti-PD-1 and PD-L1 therapies often exhibit limited efficacy in pancreatic cancer patients. This may be attributed to the role of tumor-associated macrophages in facilitating immune evasion reactions in T cells. Consequently, we conducted a comprehensive analysis of these tumor-associated macrophages and selected NR1H3 among the characteristic genes within the C1 subtype for functional experimental validation. Our findings indicate that NR1H3 holds promise as a relevant therapeutic target for pancreatic cancer patients. Moreover, in comparison to Yang et al.'s (2023) study, we utilized congruent single-cell data for pancreatic cancer, all of which were subjected to analysis focusing on tumor-associated macrophages. In our research, we employed a time-based dynamic trajectory to categorize patients into three distinct subtypes and subsequently identified novel treatment modalities tailored to patients within each subtype. This
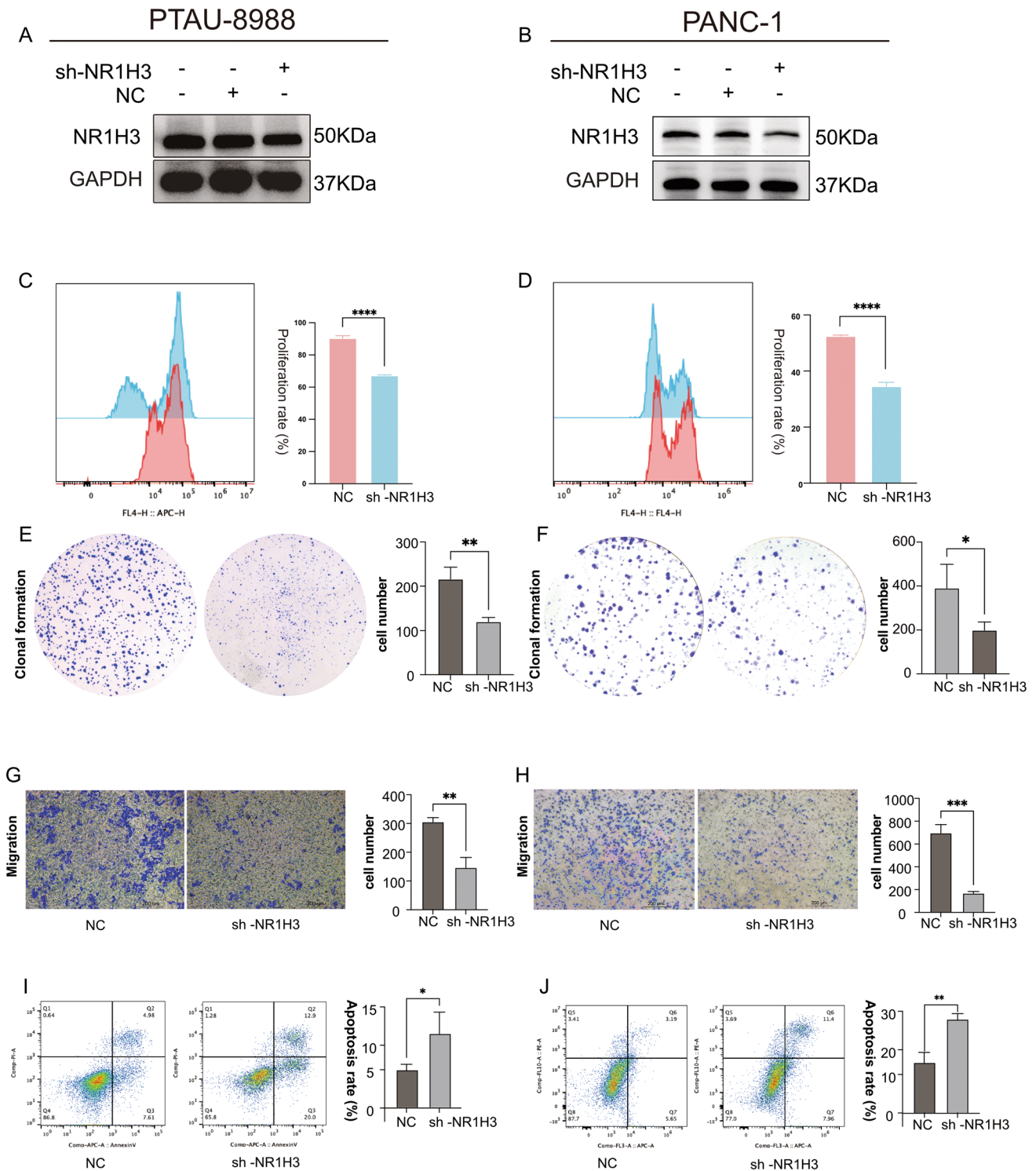
Fig. 8 The role of gene NR1H3 in pancreatic cancer cell lines was verified in vitro. **A** Western blot verified the knockdown efficiency of NR1H3 in PTAU-8988 cells. **B** Western blot verified the knockdown efficiency of NR1H3 in PANC-1 cells. **C** The EdU assay was used to detect the effect of NR1H3 knockdown on cell proliferation in PTAU-8988 cells. **D** The EdU assay was used to detect the effect of NR1H3 knockdown on cell proliferation in PANC-1 cells. **E**, **F** Clonal formation assays revealed that the ability of the PTAU-8988 cell line and the PANC-1 cell line to produce colonies was considerably reduced following NR1H3 knockdown. **G**, **H** Subsequent Transwell assays revealed that the abrogation of NR1H3 considerably curtailed the migratory capabilities of pancreatic cancer cells in both tested cell lines. **I**, **J** The effect of NR1H3 knockdown on cell apoptosis was detected by apoptosis assay in the PTAU-8988 cell line and the PANC-1 cell line. $*P < 0.05$; $**P < 0.01$; $***P < 0.001$; $****P < 0.0001$

precision subtyping approach holds the potential to significantly advance the prospects of precise therapeutic strategies for pancreatic cancer patients.

We observed that TAM undergo a dynamic shift, marked by a systematic decrease in immune inflammatory pathways, including those controlling regulate immune responses and signaling pathways modulating Th1 and Th2 cell differentiation. In contrast, pathways overseeing tumor proliferation, metastasis, and angiogenesis witnessed a progressive amplification. We believe that genes along TAM's evolutionary trajectory are crucial for prognosis and heterogeneous clinical outcomes, thereby serving as a strong basis for developing a molecular classification. PAAD is renowned for its profound heterogeneity and dire prognosis. Even among patients at equivalent clinical stages, standardized therapies exhibit notable diversity, primarily due to the neglect of individual molecular disparities (Yoshida et al. 2017). These varied clinical results emphasize the urgency of probing the genomic specifics of PAAD patients and devising a novel molecular classification. Concerning survival prognosis, our data suggest that patients of different molecular subtypes indeed experience disparate outcomes, especially those within the C3 subtype, which presents a favorable prognosis. This insight corresponds with previous assertions and paves the way for more precise therapeutic approaches for patients of varied subtypes, ultimately aiming to improve pancreatic cancer patient survival rates.

Additionally, the three clusters demonstrate unique genomic characteristics. The genomic instability of C2 is particularly striking, evidenced by both somatic mutations and CNV analyses. Historical data suggest that patients harboring TP53 and KRAS mutations exhibit a higher tendency for immune evasion and face grim prognoses (Zhang et al. 2021). Aligning with these insights, our research also shows that C2, containing the highest frequency of KRAS mutations, correlates with a bleak prognosis. In contrast, C1 and C3 maintain more stable genomic configurations and might possess a higher probability of responding positively to immunotherapy.

The effectiveness of various clinical treatments was evaluated and contrasted across three different clusters. C1, characterized by its "immunofever" signature, boasted a proliferation of immune cells and an amplified expression of immune checkpoints, including CD4[+] T cells, CD8[+] T cells, activated dendritic cells, and CD274 (PD-L1), among others (Mariathasan et al. 2018; Lee and Radford 2019). Activated CD8 T cells contribute to anti-tumor activity by targeting and destroying tumor cells, whereas dendritic cells spearhead the immune response. Furthermore, C1 has manifested an elevated propensity for antigen presentation. Per the TIS and Submap analytical methodologies, C1 could exhibit increased receptiveness to immunotherapy. Consequently, patients classified under C1 are encouraged to contemplate immunotherapy as a viable treatment alternative.

As previously noted, the C2 phenotype is marked by pronounced genomic instability, an escalated malignant profile, and a grim prognostic outlook. It becomes crucial to explore additional strategies to enhance both patient and treatment outcomes. Interestingly, certain patients display an inclination for specific drug therapies, while others encounter adverse reactions to pharmacological interventions. Despite the limited efficacy of existing targeted therapies and immunotherapy in addressing pancreatic cancer, strides in fundamental research and clinical trials have catalyzed a significant shift towards integrating multimodal targeted therapy with immunotherapy, representing a prominent therapeutic approach (Chao et al. 2018). To facilitate the delivery of accurate treatment, the pRRophetic algorithm was utilized to pinpoint potential therapeutic drugs for C2, including lapatinib and ulixertinib. Lapatinib kinases exhibit a significant promise as therapeutic targets for PAAD (Wu et al. 2015), while ulixertinib, acting as an ERK1/2 kinase inhibitor, has shown preclinical anticancer activity in pancreatic cancer cell lines harboring BRAF and RAS mutations (Jiang et al. 2018).

Liver X receptor alpha (LXRα), encoded by the NR1H3 gene, is a nuclear receptor that forms a subfamily within the NR superfamily, alongside liver X receptor beta (Vogeler et al. 2016). Previously considered a key regulatory factor in macrophage function and lipid homeostasis (Zelcer et al. 2009), recent research has revealed associations between NR1H3 and various cancer types. Vedin et al. found that activation of NR1H3 can inhibit the proliferation of colon cancer cells and reduce glutathione levels, consistent with increased cellular oxidative stress (Vedin et al. 2013). In breast cancer, some researchers posit NR1H3 as a potential tumor suppressor gene (Vedin et al. 2009). In bladder cancer, authors have linked NR1H3 expression with poor prognosis (Wu et al. 2017), aligning with our research findings. These discoveries underscore the critical pathophysiological roles of NR1H3 in various cancer processes and highlight its potential as a therapeutic target in PAAD.

Although enhancing prognosis is compelling, acknowledging certain constraints is critical. First, the specimens utilized in this research were drawn from retrospective data; therefore, conducting prospective studies is crucial to validate these findings. Second, a comprehensive multicenter dataset, inclusive of eligible immunotherapy patients, is indispensable for evaluating clinical effectiveness. Third, emerging therapeutic candidates require thorough investigation and validation via clinical trial studies. Lastly, it is important to note that gene expression levels measured using scRNA-seq technology may not necessarily reflect the actual protein expression levels.

In a word, we have identified the heterogeneity of tumors in PAAD, delineating them into three distinct clusters. Each cluster correlates with unique molecular signatures, clinical outcomes, biological characteristics, genomic variations, immune landscapes, and treatment responses. Specifically, patients in C1 are advised to pursue immunotherapy, whereas those in C2 are steered towards targeted therapy. Patients in C3 exhibit notable genetic stability strongly associated with a positive survival prognosis. Our research introduces a robust classification framework designed to improve clinical outcomes and streamline clinical management. Furthermore, our study identified the gene of NR1H3 which the characteristic gene associated with C1 may represent a potential therapeutic target for pancreatic cancer.

Abbreviations *95% CI*: 95% Confidence interval; *AUC*: Area under the curve; *CAIS*: Cluster-associated immunotherapy score; *CNV*: Copy number variation; *DEG*: Differential expression gene; *GEO*: Gene Expression Omnibus; *HR*: Hazard ratio; *IC50*: Half-maximal inhibitory concentration; *ICIs*: Immune checkpoint inhibitors; *KM*: Kaplan–Meier; *LASSO*: Least absolute shrinkage and selection operator; *NMF*: Non-negative matrix factorization; *NTP*: Nearest template prediction; *OS*: Overall survival; *PAAD*: Pancreatic adenocarcinoma; *PDAC*: Pancreatic ductal adenocarcinoma; *ROC*: Receiver operating characteristic; *SNP*: Single nucleotide polymorphism; *ssGSEA*: Single-sample gene set enrichment analysis; *Submap*: Unsupervised subclass mapping; *TAM*: Tumor-associated macrophage; *TAM–C1QC*: Tumor-associated macrophage–C1QC; *TAM–SPP1*: Tumor-associated macrophage–SPP1; *TCGA*: The Cancer Genome Atlas; *TIS*: T cell inflammatory signature; *TISCH*: Tumor Immune Single-cell Hub; *TMB*: Tumor mutational burden; *TME*: Tumor microenvironment; *TPM*: Transcripts per million; *UMAP*: Uniform manifold approximation and projection

Data availability Public data used in this work can be acquired from The Cancer Genome Atlas (TCGA) portal (https://portal.gdc.cancer.gov/), International Cancer Genome Consortium (ICGC) portal (https://dcc.icgc.org/), and Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/). The 50 hallmark gene sets could be obtained from Hallmark Molecular Signatures Database (MSigDB; https://www.gsea-msigdb.org/gsea/msigdb/). The Connectivity Map analysis could be generated from CMap database (https://clue.io).

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

Aran D, Looney AP, Liu L et al (2019) Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat Immunol 20(2):163–172. https://doi.org/10.1038/s41590-018-0276-y

Ayers M, Lunceford J, Nebozhyn M et al (2017) IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. J Clin Invest 127(8):2930–2940. https://doi.org/10.1172/jci91190

Bailey P, Chang DK, Nones K et al (2016) Genomic analyses identify molecular subtypes of pancreatic cancer. Nature 531(7592):47–52. https://doi.org/10.1038/nature16965

Brunet JP, Tamayo P, Golub TR, Mesirov JP (2004) Metagenes and molecular pattern discovery using matrix factorization. Proc Natl Acad Sci U S A 101(12):4164–4169. https://doi.org/10.1073/pnas.0308531101

Cao L, Huang C, Cui ZD et al (2021) Clinical proteomic tumor analysis consortium. Proteogenomic characterization of pancreatic ductal adenocarcinoma. Cell 184(19):5031–5052.e26. https://doi.org/10.1016/j.cell.2021.08.023

Chao Y, Xu L, Liang C et al (2018) Combined local immunostimulatory radioisotope therapy and systemic immune checkpoint blockade imparts potent antitumour responses. Nat Biomed Eng 2(8):611–621. https://doi.org/10.1038/s41551-018-0262-6

Davidson C, Taggart D, Sims AH, Lonergan DW, Canel M, Serrels A (2022) FAK promotes stromal PD-L2 expression associated with poor survival in pancreatic cancer. Br J Cancer 127(10):1893–1905. https://doi.org/10.1038/s41416-022-01966-5

Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R (2020) Cell PhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat Protoc 15(4):1484–1506. https://doi.org/10.1038/s41596-020-0292-x

Emens LA, Ascierto PA, Darcy PK et al (2017) Cancer immunotherapy: opportunities and challenges in the rapidly evolving clinical landscape. Eur J Cancer 81:116–129. https://doi.org/10.1016/j.ejca.2017.01.035

Ganesh K, Stadler ZK, Cercek A et al (2019) Immunotherapy in colorectal cancer: rationale, challenges and potential. Nat Rev Gastroenterol Hepatol 16(6):361–375. https://doi.org/10.1038/s41575-019-0126-x

Geeleher P, Cox N, Huang RS (2014) pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. PloS One 9(9):e107468. https://doi.org/10.1371/journal.pone.0107468

Guo L, Zheng J, Yu T, Liu Y, Duo L (2017) Elevated expression of SATB1 is involved in pancreatic tumorigenesis and is associated with poor patient survival. Mol Med Rep 16(6):8842–8848. https://doi.org/10.3892/mmr.2017.7683

Guo C, Xie B, Liu Q (2022) Weighted gene co-expression network analysis combined with machine learning validation to identify key hub biomarkers in colorectal cancer. Funct Integr Genomics 23(1):24. https://doi.org/10.1007/s10142-022-00949-2

Han J, DePinho RA, Maitra A (2021) Single-cell RNA sequencing in pancreatic cancer. Nat Rev Gastroenterol Hepatol 18(7):451–452. https://doi.org/10.1038/s41575-021-00471-z

Hoshida Y (2010) Nearest template prediction: a single-sample-based flexible class prediction with confidence assessment. PLoS One. 5(11):e15543. https://doi.org/10.1371/journal.pone.0015543

Hoshida Y, Brunet JP, Tamayo P, Golub TR, Mesirov JP (2007) Subclass mapping: identifying common subtypes in independent disease data sets. PLoS One 2(11):e1195. https://doi.org/10.1371/journal.pone.0001195

Huang TX, Fu L (2019) The immune landscape of esophageal cancer. Cancer Commun (Lond) 39(1):79. https://doi.org/10.1186/s40880-019-0427-z

Isella C, Brundu F, Bellomo SE et al (2017) Selective analysis of cancer-cell intrinsic transcriptional traits defines novel clinically relevant subtypes of colorectal cancer. Nat Commun 8:15107. https://doi.org/10.1038/ncomms15107

Jiang H, Xu M, Li L et al (2018) Concurrent HER or PI3K inhibition potentiates the antitumor effect of the ERK inhibitor ulixertinib in preclinical pancreatic cancer models. Mol Cancer Ther 17(10):2144–2155. https://doi.org/10.1158/1535-7163.Mct-17-1142

Korsunsky I, Millard N, Fan J et al (2019) Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 16(12):1289–1296. https://doi.org/10.1038/s41592-019-0619-0

Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9:559. https://doi.org/10.1186/1471-2105-9-559

Lee YS, Radford KJ (2019) The role of dendritic cells in cancer. Int Rev Cell Mol Biol 348:123–178. https://doi.org/10.1016/bs.ircmb.2019.07.006

Li R, Zhang S, Liu G (2022) Identification and validation of a pyroptosis-related prognostic model for colorectal cancer. Funct Integr Genomics 23(1):21. https://doi.org/10.1007/s10142-022-00935-8

Liao X, Huang K, Huang R et al (2017) Genome-scale analysis to identify prognostic markers in patients with early-stage pancreatic ductal adenocarcinoma after pancreaticoduodenectomy. Onco Targets Ther 10:4493–4506. https://doi.org/10.2147/ott.S142557

Liu L, Liu Z, Gao J et al (2022) CD8+ T cell trajectory subtypes decode tumor heterogeneity and provide treatment recommendations for hepatocellular carcinoma. Front Immunol 13:964190. https://doi.org/10.3389/fimmu.2022.964190

Liu X, Song J, Zhang H et al (2023) Immune checkpoint HLA-E:CD94-NKG2A mediates evasion of circulating tumor cells from NK cell surveillance. Cancer Cell 41(2):272-287.e9. https://doi.org/10.1016/j.ccell.2023.01.001

Lovmar L, Ahlford A, Jonsson M, Syvänen AC (2005) Silhouette scores for assessment of SNP genotype clusters. BMC Genomics 6:35. https://doi.org/10.1186/1471-2164-6-35

Mariathasan S, Turley SJ, Nickles D et al (2018) TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. Nature 554(7693):544–548. https://doi.org/10.1038/nature25501

Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP (2018) Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res 28(11):1747–1756. https://doi.org/10.1101/gr.239244.118

Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G (2011) GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol 12(4):R41. https://doi.org/10.1186/gb-2011-12-4-r41

Newman AM, Liu CL, Green MR et al (2015) Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 12(5):453–457. https://doi.org/10.1038/nmeth.3337

Novak I, Yu H, Magni L, Deshar G (2020) Purinergic signaling in pancreas-from physiology to therapeutic strategies in pancreatic cancer. Int J Mol Sci 21(22):8781. https://doi.org/10.3390/ijms21228781

Noy R, Pollard JW (2014) Tumor-associated macrophages: from mechanisms to therapy. Immunity 41(1):49–61. https://doi.org/10.1016/j.immuni.2014.06.010

Pan Y, Lu F, Xiong P et al (2018) WIPF1 antagonizes the tumor suppressive effect of miR-141/200c and is associated with poor survival in patients with PDAC. J Exp Clin Cancer Res 37(1):167. https://doi.org/10.1186/s13046-018-0848-6

Pan Y, Lu F, Fei Q et al (2019) Single-cell RNA sequencing reveals compartmental remodeling of tumor-infiltrating immune cells induced by anti-CD47 targeting in pancreatic cancer. J Hematol Oncol 12(1):124. https://doi.org/10.1186/s13045-019-0822-6

Park W, Chawla A, O'Reilly EM (2021) Pancreatic cancer: a review. JAMA 326(9):851–862. https://doi.org/10.1001/jama.2021.13027

Peng J, Sun BF, Chen CY et al (2019) Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. Cell Res 29(9):725–738. https://doi.org/10.1038/s41422-019-0195-y

Pothula SP, Xu Z, Goldstein D, Pirola RC, Wilson JS, Apte MV (2020) Targeting HGF/c-MET axis in pancreatic cancer. Int J Mol Sci 21(23):9170. https://doi.org/10.3390/ijms21239170

Qiu X, Mao Q, Tang Y et al (2017) Reversed graph embedding resolves complex single-cell trajectories. Nat Methods 14(10):979–982. https://doi.org/10.1038/nmeth.4402

Ruffell B, Affara NI, Coussens LM (2012) Differential macrophage programming in the tumor microenvironment. Trends Immunol 33(3):119–126. https://doi.org/10.1016/j.it.2011.12.001

Sanchez-Vega F, Mina M, Armenia J et al (2018) Oncogenic signaling pathways in The Cancer Genome Atlas. Cell 173(2):321-337.e10. https://doi.org/10.1016/j.cell.2018.03.035

Satija R, Farrell JA, Gennert D, Schier AF, Regev A (2015) Spatial reconstruction of single-cell gene expression data. Nat Biotechnol 33(5):495–502. https://doi.org/10.1038/nbt.3192

Shi S, Hua J, Liang C et al (2019) Proposed modification of the 8th edition of the AJCC staging system for pancreatic ductal adenocarcinoma. Ann Surg 269(5):944–950. https://doi.org/10.1097/sla.0000000000002668

Siegel RL, Miller KD, Fuchs HE, Jemal A (2022) Cancer statistics, 2022. CA Cancer J Clin 72(1):7–33. https://doi.org/10.3322/caac.21708

Sun D, Wang J, Han Y et al (2021) TISCH: a comprehensive web resource enabling interactive single-cell transcriptome visualization of tumor microenvironment. Nucleic Acids Res 49(D1):D1420-d1430. https://doi.org/10.1093/nar/gkaa1020

Vedin L-L, Lewandowski SA, Parini P, Gustafsson J-Å, Steffensen KR (2009) The oxysterol receptor LXR inhibits proliferation of human breast cancer cells. Carcinogenesis 30(4):575–579

Vedin LL, Gustafsson JÅ, Steffensen KR (2013) The oxysterol receptors LXRα and LXRβ suppress proliferation in the colon. Mol Carcinog 52(11):835–844

Vogeler S, Bean TP, Lyons BP, Galloway TS (2016) Dynamics of nuclear receptor gene expression during Pacific oyster development. BMC Dev Biol 16(1):33. https://doi.org/10.1186/s12861-016-0129-6

Wei J, Chen Z, Hu M et al (2021) Characterizing intercellular communication of pan-cancer reveals SPP1+ tumor-associated macrophage expanded in hypoxia and promoting cancer malignancy through single-cell RNA-seq data. Front Cell Dev Biol 9:749210

Wu Z, Gabrielson A, Hwang JJ et al (2015) Phase II study of lapatinib and capecitabine in second-line treatment for metastatic pancreatic cancer. Cancer Chemother Pharmacol 76(6):1309–1314. https://doi.org/10.1007/s00280-015-2855-z

Wu J, Wan F, Sheng H et al (2017) NR1H3 expression is a prognostic factor of overall survival for patients with muscle-invasive bladder cancer. J Cancer 8(5):852–860. https://doi.org/10.7150/jca.17845

Yang K, Yang T, Yu J, Li F, Zhao X (2023) Integrated transcriptional analysis reveals macrophage heterogeneity and macrophage-tumor cell interactions in the progression of pancreatic ductal adenocarcinoma. BMC Cancer 23(1):199. https://doi.org/10.1186/s12885-023-10675-y

Yoshida N, Masamune A, Hamada S et al (2017) Kindlin-2 in pancreatic stellate cells promotes the progression of pancreatic cancer. Cancer Lett 390:103–114. https://doi.org/10.1016/j.canlet.2017.01.008

Yoshihara K, Shahmoradgoli M, Martínez E et al (2013) Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun 4:2612. https://doi.org/10.1038/ncomms3612

Zelcer N, Hong C, Boyadjian R, Tontonoz P (2009) LXR regulates cholesterol uptake through Idol-dependent ubiquitination of the LDL receptor. Science 325(5936):100–104. https://doi.org/10.1126/science.1168974

Zhang X, Lan Y, Xu J et al (2019) Cell Marker: a manually curated resource of cell markers in human and mouse. Nucleic Acids Res 47(D1):D721-d728. https://doi.org/10.1093/nar/gky900

Zhang L, Li Z, Skrzypczynska KM et al (2020) Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer. Cell 181(2):442-459.e29. https://doi.org/10.1016/j.cell.2020.03.048

Zhang F, Zhong W, Li H, Huang K, Yu M, Liu Y (2021) TP53 Mutational status-based genomic signature for prognosis and predicting therapeutic response in pancreatic cancer. Front Cell Dev Biol 9:665265. https://doi.org/10.3389/fcell.2021.665265

Zyla J, Marczyk M, Domaszewska T, Kaufmann SHE, Polanska J, Weiner J (2019) Gene set enrichment for reproducible science: comparison of CERNO and eight other algorithms. Bioinformatics 35(24):5146–5154. https://doi.org/10.1093/bioinformatics/btz447