OPEN ACCESS

Original research

Single-cell RNA-seq analysis reveals BHLHE40-driven pro-tumour neutrophils with hyperactivated glycolysis in pancreatic tumour microenvironment

Liwen Wang,[1,2,3] Yihao Liu,[1,2,3] Yuting Dai,[4] Xiaomei Tang,[1,2,3] Tong Yin,[4] Chaofu Wang,[5] Ting Wang,[5] Lei Dong,[5] Minmin Shi,[1,2,3] Jiejie Qin,[1,2,3] Meilin Xue,[1,2,3] Yizhi Cao,[1,2,3] Jia Liu,[1,2,3] Pengyi Liu,[1,2,3] Jinyan Huang,[4,6] Chenlei Wen,[1,2,3] Jun Zhang,[1,2,3] Zhiwei Xu,[1,2,3] Fan Bai,[7] Xiaxing Deng,[1,2,3] Chenghong Peng,[1,2,3] Hao Chen,[1,2,3] Lingxi Jiang,[1,2,3] Saijuan Chen,[4,8] Baiyong Shen [1,2,3]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/gutjnl-2021-326070).

For numbered affiliations see end of article.

Correspondence to
Dr Baiyong Shen, Department of General Surgery, Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital, Shanghai 200025, China; shenby@shsmu.edu.cn
Dr Lingxi Jiang; jlx12120@rjh.com.cn
Saijuan Chen; sjchen@stn.sh.cn

LW, YL, YD and XT contributed equally.

Received 8 September 2021
Accepted 27 May 2022

To cite: Wang L, Liu Y, Dai Y, *et al*. *Gut* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/gutjnl-2021-326070

ABSTRACT

Objective Innate immunity plays important roles in pancreatic ductal adenocarcinoma (PDAC), as non-T-cell-enriched tumour. Neutrophils are major players in innate immune system. Here, we aimed to explore the heterogeneity and pro-tumour mechanisms of neutrophils in PDAC.

Design We analysed single-cell transcriptomes of peripheral blood polymorphonuclear leucocytes (PMNs) and tumour-infiltrating immune cells from five patients with PDAC, and performed immunofluorescence/immunohistochemistry staining, multi-omics analysis and *in vitro* experiments to validate the discoveries of bioinformatics analysis.

Results Exploration of the heterogeneity of tumour-associated neutrophils (TANs) revealed a terminally differentiated pro-tumour subpopulation (TAN-1) associated with poor prognosis, an inflammatory subpopulation (TAN-2), a population of transitional stage that have just migrated to tumour microenvironment (TAN-3) and a subpopulation preferentially expressing interferon-stimulated genes (TAN-4). Glycolysis signature was upregulated along neutrophil transition trajectory, and TAN-1 was featured with hyperactivated glycolytic activity. The glycolytic switch of TANs was validated by integrative multi-omics approach of transcriptomics, proteomics and metabolomics analysis. Activation of glycolytic activity by LDHA overexpression induced immunosuppression and pro-tumour functions in neutrophil-like differentiated HL-60 (dHL-60) cells. Mechanistic studies revealed BHLHE40, downstream to hypoxia and endoplasmic reticulum stress, was a key regulator in polarisation of neutrophils towards TAN-1 phenotype, and direct transcriptional regulation of BHLHE40 on TAN-1 marker genes was demonstrated by chromatin immunoprecipitation assay. Pro-tumour and immunosuppression functions were observed in dHL-60 cells overexpressing BHLHE40. Importantly, immunohistochemistry analysis of PDAC tissues revealed the unfavourable prognostic value of BHLHE40^{+} neutrophils.

Conclusion The dynamic properties of TANs revealed by this study will be helpful in advancing PDAC therapy targeting innate immunity.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ As non-T-cell enriched tumour, pancreatic ductal adenocarcinoma (PDAC) exhibited limited response to current immunotherapeutic strategies with immune checkpoint blockade or chimeric antigen receptor T.
⇒ Abundant infiltration of tumour-associated neutrophils (TANs) was found in PDAC tumour microenvironment, and was recognised as an important unfavourable prognostic factor in most solid tumours.
⇒ As an important player of innate immunity, neutrophils promote tumour progression via promoting cell proliferation, angiogenesis, tissue remodelling, immunosuppression and metastasis.
⇒ Neutrophils are delicate cells, and were not detected in most of previous single cell RNA-sequencing studies showing the comprehensive gene expression atlas of main cell types in human PDAC tumour microenvironment.

WHAT THIS STUDY ADDS

⇒ Compared with neutrophils from peripheral blood, TANs are composed of a heterogeneous population.
⇒ We observed four neutrophil subpopulations with distinctive features in PDAC tumour microenvironment, unveiling the neutrophil heterogeneity at single-cell level for the first time.
⇒ A terminally differentiated subpopulation of neutrophils with hyperactivated glycolytic activity and pro-tumour functions in PDAC milieu is associated with worse prognosis of patients.
⇒ Glycolytic activity is upregulated along neutrophil transition trajectory in PDAC.
⇒ Glycolytic switch promotes immunosuppression and pro-tumour functions of neutrophils.
⇒ BHLHE40, activated by hypoxia and endoplasmic reticulum stress, is a key regulator driving neutrophils towards pro-tumour and immunosuppressive subtype.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE AND/OR POLICY

⇒ Our work provides a comprehensive atlas of neutrophils from PDAC tumour microenvironment.

⇒ The molecular mechanisms of driving neutrophils towards pro-tumour subtype demonstrated in our study will facilitate the development of novel immunotherapies targeting the pro-tumour subcluster or immunometabolism of TANs for patients with PDAC, providing alterative choice to current immunotherapy which is mainly based on the role of cytotoxic T cells.

INTRODUCTION

Pancreatic cancer, comprised mostly of pancreatic ductal adenocarcinoma (PDAC), is one of the most lethal malignancies and one of the leading causes of cancer-related mortality, with a dismal 5-year survival rate of 9%.[1] Immunotherapeutic strategies with immune checkpoint blockade or chimeric antigen receptor T yield limited efficacy in PDAC as high stromal density creates a physical barrier to T-cell infiltration.[2 3] Minimal antitumour T-cell infiltration was observed in the immune contexture of PDAC.[4] In non-T-cell-infiltrated tumours, it is required to trigger innate immune activation which facilitates signals for effector T-cell trafficking and bridges towards adaptive immunity.[5 6] Therefore, therapeutic interventions targeting innate immune system may provide alternative choice to current immunotherapies in PDAC.

Similar to other solid tumours, PDAC is featured by extensive infiltration of immunosuppressive cells in tumour microenvironment.[3 7 8] The fibro-inflammatory filtrates are educated by cancer cells to provide a favourable microenvironment, supporting the immune escape, malignant transformation and progression of neoplastic cells.[3 7 8] Among various types of tumour-infiltrating immune cells, the important role of neutrophils in tumour progression has attracted extensive research interest during the past decade.[9] Neutrophils are a highly plastic population that present with heterogeneous phenotypes in response to various environmental cues.[9 10] N2-polarised neutrophils represent the vast majority of neutrophils in tumour microenvironment, supporting tumour progression via promoting tumour cell proliferation, angiogenesis, tissue remodelling, immunosuppression and metastasis.[9 10] Infiltration of tumour-associated neutrophils (TANs) has been recognised as an unfavourable prognostic factor in most solid malignancies.[11] However, there is still evidence of neutrophils that oppose tumour progression as they exhibit direct cytotoxicity against tumour cells[12 13] or antibody-dependent cellular cytotoxicity.[14] Therefore, the characterisation of TANs in tumour remains obscure.

In a previous study, bioinformatics analysis of bulk RNA-sequencing (RNA-seq) data from TCGA cohort revealed that TAN infiltration is relatively abundant in PDAC compared with other cancer types, implying the importance of TANs in PDAC microenvironment.[15] However, most of prior studies using single cell sequencing approach failed to detect human neutrophils from PDAC tissues.[16–21] We hypothesise that this is due to the systematic bias resulting from the nature of neutrophils as delicate cells with a short lifespan.[22] High sensitivity and various enzymes in neutrophil-derived granules contribute to their vulnerability *in vitro*.[22] Therefore, neutrophils might not be able to survive the process of tissue digestion or single-cell capture. Microwell-based single-cell capture method using BD Rhapsody provides an alternative option for characterising human neutrophils.[23 24]

In this study, we succeeded in acquiring single-cell transcriptomes of 21 972 neutrophils from five patients with PDAC using BD Rhapsody. We explored the heterogeneity of TANs in PDAC microenvironment, and further validated the existence and clinical relevance of newly discovered neutrophil subtypes with immunofluorescence (IF) and immunohistochemistry (IHC) staining on PDAC tissues. In addition, we uncovered the underlying mechanisms driving TANs towards pro-tumour phenotypes, and identified potential therapeutic targets associated with neutrophils.

MATERIALS AND METHODS

Processing of clinical samples

For peripheral blood, red blood cells were depleted by gravity sedimentation, and neutrophils were labelled with CD66b-phycoerythrin (PE) monoclonal antibody (BioLegend, 305106), and isolated by anti-PE beads and MACS column (Miltenyi), according to manufacturer's protocol. PDAC tumour tissues were digested into single-cell suspension with Tumor Dissociation Kit (Miltenyi, 130-095-929). Infiltrating immune cells were isolated using CD45 microbeads and MACS column (Miltenyi), while infiltrating neutrophils were isolated using the same protocol as that for peripheral blood. Detailed information of patient enrolment, sample processing and purity assessment is provided in online supplemental methods.

Single-cell RNA-sequencing and data analysis

Single-cell capture was achieved by BD Rhapsody system. Whole transcriptome libraries were prepared according to the BD Rhapsody single-cell whole-transcriptome amplification workflow, and sequenced using HiSeq Xten (Illumina, San Diego, California, USA) on 150 bp paired-end run. Raw data were processed using fastp to filter adaptor sequences and remove low-quality reads.[25] Cell barcode whitelist was identified by UMI-tools.[26] The UMI-based clean data were mapped to human reference genome (Ensemble V.91) using STAR algorithm.[27] UMI count matrices were generated for each sample, and imported into Seurat R toolkit (V.3.2.3).[28] Cluster analysis was performed by Seurat,[28] pseudotime trajectory analysis was performed with Monocle2 package (V.2.18.0),[29] cell-cell communication was analysed with CellPhoneDB[30] and gene regulatory networks were constructed with Single-Cell Regulatory Network Inference and Clustering (SCENIC) analysis.[31] Detailed description of single-cell sequencing and data analysis are provided in online supplemental methods.

RESULTS

Single-cell landscape of neutrophils from peripheral blood and PDAC tumours

Based on PDAC cohort from our centre, we observed substantial infiltration of TANs in areas within or adjacent to malignant cells in PDAC tissues (online supplemental figure S1A-B), and the unfavourable prognosis in patients with PDAC with higher neutrophil infiltration (online supplemental figure S1C). To comprehensively catalogue the populations of neutrophils and their crosstalk with other immune filtrates in tumour microenvironment, we generated single-cell RNA-seq data of $CD66b^+$ peripheral blood polymorphonuclear leucocytes (PMNs) and $CD45^+$ tumour-infiltrating immune cells from five treatment-naïve patients with PDAC. PMNs from healthy controls and patients with chronic pancreatitis were also transcriptionally

characterised as controls (figure 1A, left panel and online supplemental table S1). To validate the findings based on single-cell RNA-seq, we collected paired CD66b$^+$ cells from peripheral blood and tumour tissue (PMNs and TANs, respectively) from 24 patients with PDAC for multi-omics analysis and quantitative PCR (qPCR) analysis (validation cohort 1), and recruited additional 114 patients for IHC, IF, and spatial transcriptomics study (validation cohort 2) (figure 1A, right panel, online supplemental figure S1D and online supplemental table S2).

After initial quality control, we obtained 96 692 883 unique transcripts from a total of 51 823 cells, including 33 891 cells originated from peripheral blood and 17 932 cells from PDAC tumour tissues (online supplemental table S3). Following principal component analysis and graph-based cluster analysis, we identified nine cell clusters that were annotated as neutrophils (two clusters), macrophages, mast cells, B cells, plasma cells, T cells, ductal cells and fibroblasts, according to well-known cell type marker genes (figure 1B, online supplemental figure S2A-D, online supplemental table S4). We identified two different types of neutrophils, among which neutrophil 1 was strongly enriched in peripheral blood (predominantly composed of PMNs), while neutrophil 2 was present exclusively in tumour tissues (composed of TANs) (figure 1B, online supplemental figure S2E).

To compare the expression profiles of neutrophils from different origins, we visualised the transcriptomes of neutrophils from each sample by t-SNE analysis, which mapped the high-dimensional data to a two-dimensional space, with pairwise similarities of input objects preserved.[32] According to the t-SNE plot, TANs have a unique RNA profile, while PMNs from patients with PDAC are very similar to PMNs from healthy controls or patients with chronic pancreatitis (online supplemental figure S3A-B), indicating significant transcriptional reprogramming of neutrophils after recruitment into tumour tissues. Pathway enrichment analysis revealed inhibition of neutrophil differentiation in PMNs from patients with PDAC compared with that from healthy controls/chronic pancreatitis, whereas inhibited innate immunity, activated multiple cancer-associated signalling pathways and profound metabolic changes were observed in TANs (online supplemental figure S3C).

Cell-cell communication analysis revealed intimate crosstalk between neutrophils and macrophages in tumour microenvironment

To explore the crosstalk between neutrophils and other immune cells in tumour microenvironment, we applied CellPhoneDB, an interactive web application that infers cellular interactions according to ligand-receptor signalling database.[30] Notably, compared with other types of immune cells, macrophages expressed significantly higher number of receptors corresponding to ligands from neutrophils, and conversely, macrophages also expressed significantly higher number of ligands corresponding to receptors expressed by neutrophils (figure 1C,D), indicating the close interaction between neutrophils and macrophages in tumour microenvironment. Our data suggested that tumour-associated macrophages attracted neutrophils via CCL13-CCR1, CCL3-CCR1, CCL3L3-CCR1, CXCL2-CXCR1, CXCL2-CXCR2 and CXCL8-CXCR2 axes, while macrophages were recruited by neutrophils through CCL3L3-CCR1 axis (figure 1C). Consistent with the reciprocal recruitment implied by the expression of chemokine and chemokine receptors, macrophage and neutrophil signatures were positively correlated in treatment-naïve PDAC from The Cancer Genome Atlas-Pancreatic adenocarcinoma (TCGA-PAAD) cohort (online supplemental figure S4A). In addition, IF staining confirmed the physical proximity between neutrophils and macrophages in PDAC tissues (figure 1E). Of note, macrophages stimulated neutrophils with pro-inflammatory cytokines interleukin 1 (IL-1) and tumour necrosis factor (TNF) (figure 1C), and significant activation of TNFα/nuclear factor kappa B (NF-κB) and IL-1 signalling pathways were observed in TANs in comparison with PMNs (figure 1F). Interestingly, IF staining of PDAC tissues validated the activation of NF-κB and MAPK signalling pathways (revealed by the presence of p65 and c-JUN in nuclei) downstream to TNFα and IL-1 in TANs (online supplemental figure S4B), and qPCR analysis also demonstrated the upregulated expression of multiple TNFα and IL-1 target genes (online supplemental figure S4C),[33–35] confirming the TNFα- and IL-1-induced activation of TANs.

TANs from PDAC tumours are composed of a heterogeneous population

To further investigate the heterogeneity of neutrophils, we performed dimensionality reduction and clustering of PMNs and TANs from patients with PDAC. The data without batch effect correction are shown in online supplemental figure S5A, in which both PMNs and TANs clustered according to their sample of origin. After correction of batch effect, six PMN subclusters and six TAN subclusters were identified (figure 2A,B, online supplemental figure S5B,C, online supplemental tables S5,S6), and each neutrophil subcluster was present in all the PDAC patient samples (online supplemental figure S5D). PMN-5 was excluded due to preferential expression of eosinophil marker Charcot-Leyden crystal galectin (online supplemental figure S5C). TAN-5 consisted of low-quality cells and was also disregarded for further analysis (online supplemental figure S5E). Correlation analysis revealed that all of the PMN subclusters had similar expression profiles, while TAN subclusters were more distinct from each other (figure 2C), demonstrating that tumour-infiltrating neutrophils is a heterogeneous population, which could be attributed to varying phenotypes in response to diverse environmental stimuli in tumour microenvironment.

In consideration of the heterogeneity of TANs, we further explored the characteristics of each TAN subcluster. Of note, no cluster-specific distinctive features were identified in TAN-0, as the marker genes of TAN-0 were also highly expressed in other TAN subclusters (online supplemental figure S5C). TAN-1 could be recognised as the 'pro-tumour subpopulation', which highly expressed pro-angiogenic factor VEGFA,[36] pro-metastatic factor PLAU[37 38] and LGALS3, a molecule with multiple functions beneficial for tumour progression, including enhancing proliferation and stemness of malignant cells, promoting angiogenesis, suppressing immune surveillance, promoting M2 macrophage differentiation and contributing to drug resistance[39 40] (figure 2D (i)). TAN-2 consisted of an inflammatory subpopulation, which strongly expressed inflammation-associated genes NLRP3 and PDE4B,[41 42] and neutrophil activation marker CD69[43 44] (figure 2D (ii)). TAN-2 also strongly expressed pro-tumour molecules IL1RN and adrenomedullin (ADM) (figure 2D (ii)), which has been found to support tumour progression in other cancer types.[45 46] Among all the TAN subclusters, the gene expression profile of TAN-3 was most similar to PMNs (figure 2C), and TAN-3 highly expressed genes associated with transendothelial migration of neutrophils (VNN2 and SELL, figure 2D (iii)),[47 48] suggesting that TAN-3 is a population of transitional stage neutrophils that have just migrated to tumour microenvironment, converting from PMNs to TANs. TAN-4 exhibited

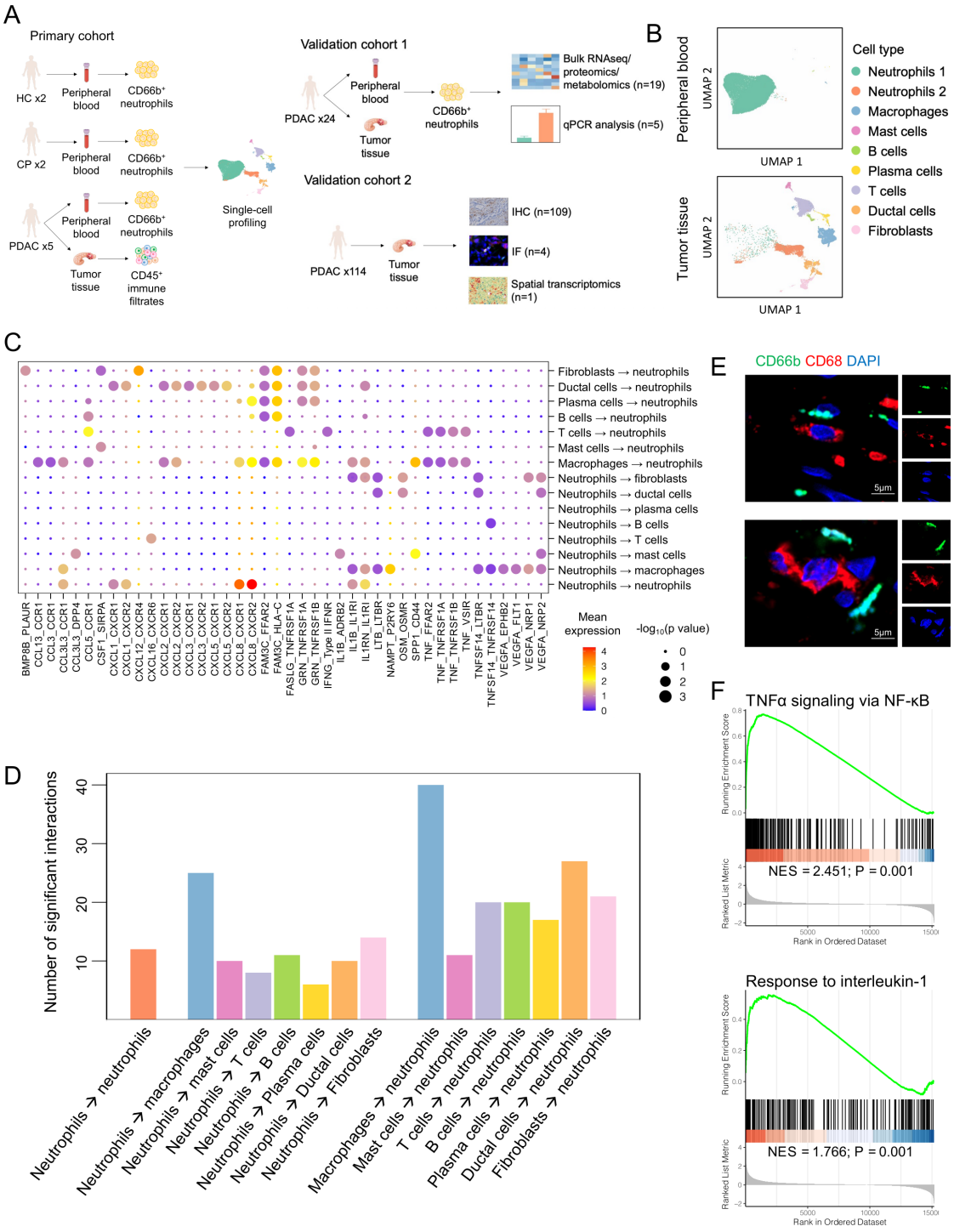

Figure 1 Single-cell atlas from peripheral blood and PDAC tumours, and cell-cell communication between neutrophils and other cells in PDAC microenvironment. (A) Graphical scheme describing the experimental workflow, including single-cell sequencing study of primary cohort, multi-omics analysis and quantitative PCR analysis of validation cohort 1, and IHC, IF and ST study of validation cohort 2. The number of patients/healthy donors in each subgroup are shown in the figure. (B) UMAP plot depicting the major cell types identified by single-cell sequencing, including CD66b$^+$ PMNs from peripheral blood of healthy controls, patients with chronic pancreatitis and patients with PDAC, and CD45$^+$ immune filtrates from PDAC tumour tissues. (C) Dot plot depicting the selected ligand-receptor interactions between neutrophils and other cells in PDAC tumour microenvironment. Mean expressions and p values were calculated by CellPhoneDB (http:// www.cellphonedb.org/), and were indicated by circle colour and size, respectively. (D) Bar plot depicting the numbers of significant ligand-receptor interactions between neutrophils and other cells in PDAC tumour microenvironment. Cell type labels were written as (the cell type expressing the ligand) → (the cell type expressing the corresponding receptor). (E) IF staining of CD66b and CD68 on PDAC tissue. (F) Gene set enrichment analysis (GSEA) plot showing the enrichment scores for TNFα and IL-1 pathways in TANs. Pathway enrichment analysis was performed based on single-cell RNA sequencing (RNA-seq) data of PMNs and TANs from patients with PDAC. CP, chronic pancreatitis; IF, immunofluorescence; IHC, immunohistochemistry; IL, interleukin; HC, healthy controls; NES, normalised enrichment score; NF-κB, nuclear factor kappa B; PDAC, pancreatic ductal adenocarcinoma; PMNs, polymorphonuclear leucocytes; ST, spatial transcriptomics; TNF, tumour necrosis factor; UMAP, uniform manifold approximation and projection.

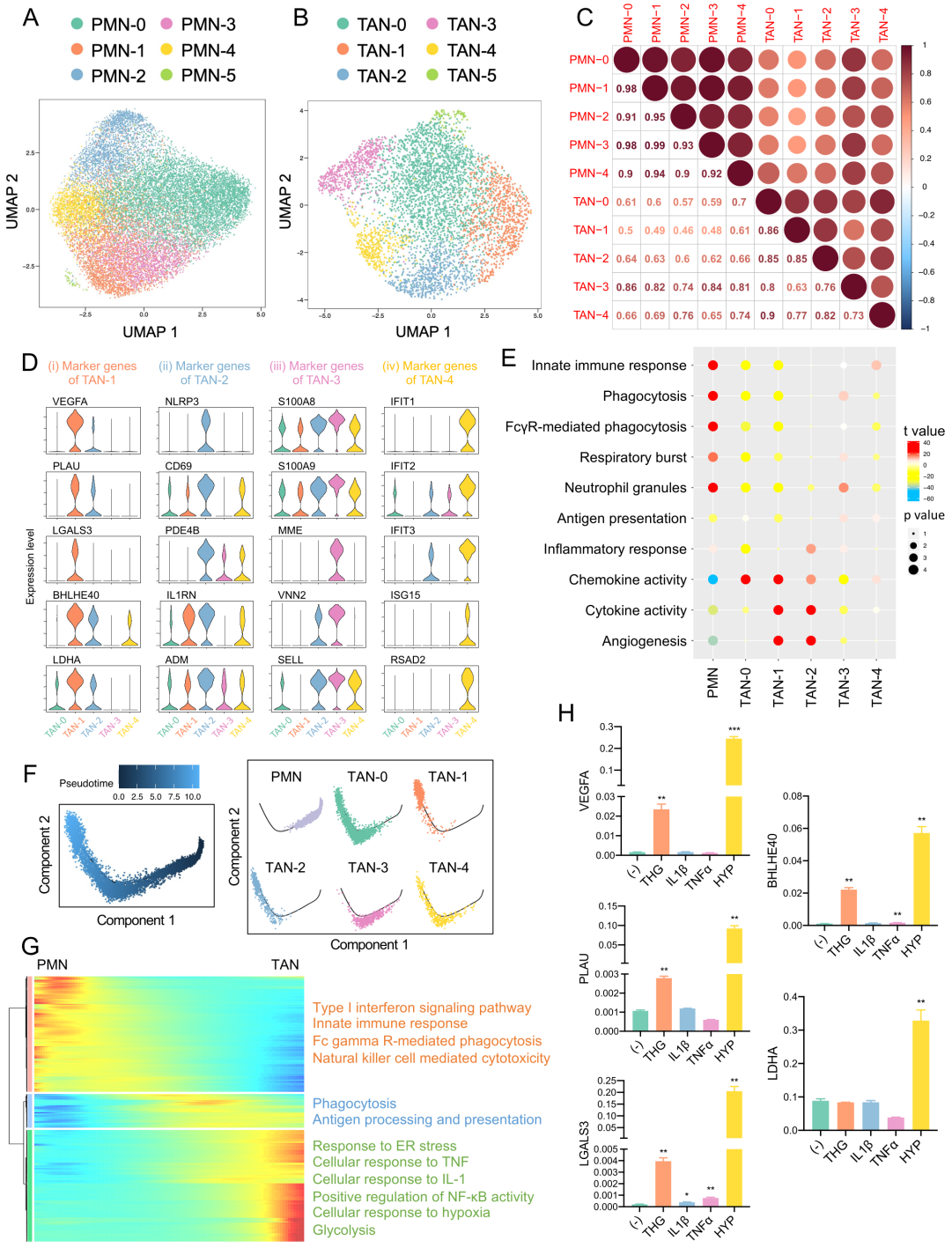

Figure 2 TANs from PDAC tumours are composed of a heterogeneous population. (A) UMAP plot showing subclusters of PMNs from patients with PDAC. (B) UMAP plot showing subclusters of TANs from patients with PDAC. All the cells from PDAC tumour tissues annotated as neutrophils were analysed as TANs, including neutrophil clusters 1 and 2 in figure 1B. (C) Pearson's correlation coefficients between expression profiles of neutrophils from each subcluster. The average normalised expression of top 2000 highly variable genes in each subcluster were calculated, and Pearson's correlation analysis was performed. (D) Violin plots showing marker genes of (i) TAN-1 subcluster, (ii) TAN-2 subcluster, (iii) TAN-3 subcluster and (iv) TAN-4 subcluster. (E) Comparison of pathway activities between the different neutrophil subclusters. The pathways were associated with neutrophil functions. The pathway activities were scored per cell by gene set variation analysis. T values and p values were calculated based on linear models analysing difference between neutrophils from one cluster and neutrophils from all other clusters, and were indicated by circle colour and size, respectively. (F) Trajectory of neutrophils along pseudotime in a two-dimensional space. Each point corresponds to a single cell. (G) 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 were shown. (H) Expression of TAN-1 marker genes in neutrophil-like dHL-60 cells stimulated with THG, IL-1β, TNFα and hypoxia for 24 hours, analysed by qPCR. qPCR data were normalised to fold over β-actin (housekeeping gene), and represented as mean with SD. *P<0.05; **p<0.01; ***p<0.001. dHL-60, differentiated HL-60; ER, endoplasmic reticulum; HYP, hypoxia; IL, interleukin; NF-κB, nuclear factor kappa B; PDAC, pancreatic ductal adenocarcinoma; PMN, polymorphonuclear leucocytes; qPCR, quantitative PCR; TAN, tumour-associated neutrophils; THG, thapsigargin; TNF, tumour necrosis factor; UMAP, uniform manifold approximation and projection.

a unique transcriptional signature, expressing interferon (IFN)-stimulated genes, including IFIT1, IFIT2, IFIT3, ISG15 and RSAD2 (figure 2D (iv)).[49]

To characterise the functions of each neutrophil subcluster, we analysed pathway activities with gene set variation analysis[50] (figure 2E). PMNs displayed high expression of genes involved in normal immune functions of neutrophils, including innate immune response, phagocytosis, respiratory burst and synthesis of neutrophil granules. Similarly, TAN-3 also highly expressed genes associated with phagocytosis, respiratory burst and neutrophil granules, and TAN-4 was associated with innate immune response. However, TAN-3 and TAN-4 preferentially displayed high activities of antigen processing and presentation, in contrast to PMNs. TAN-1 and TAN-2 similarly expressed high levels of chemokines, cytokines and angiogenic factors. TAN-2 also showed strong inflammation signature expression, consistent with the expression of inflammatory genes described above.

Trajectory analysis revealed that neutrophils terminally differentiated into pro-tumour TAN-1 state

To further study the dynamic transitional process of neutrophils from peripheral blood into tumour microenvironment, we applied monocle2[29] to construct a pseudotime map of neutrophil state trajectory. The trajectory was determined to initiate with PMNs as beginning, through TAN-3 as the transitional state between PMNs and TANs, followed by an intermediate tumour-infiltrating state characterised by TAN-0 and TAN-4, and finally reached a terminally differentiate state of TAN-2 and TAN-1 (figure 2F). Next, we analysed the single-cell transcriptomes along trajectory, and identified 1757 genes with significant expression changes, which could be clustered into three expression patterns: group 1 included genes that showed decreased expression levels along trajectory. Pathway enrichment analysis revealed that these genes were associated with IFN signalling pathway and innate immune functions. The genes in group 2 were upregulated at the early stage of tumour infiltration, and these genes participated in phagocytosis and antigen presentation. The genes in group 3 were activated at the late stage in tumour microenvironment, and were enriched in hypoxia, endoplasmic reticulum (ER) stress, IL-1 and TNF signalling pathway and glycolysis, etc (figure 2G). To further explore how the activation of those signalling pathways influence neutrophil phenotype, neutrophil-like differentiated HL-60 cells (dHL-60) were treated with ER stress inducer thapsigargin (THG), pro-inflammatory cytokines IL-1β and TNFα, and hypoxia *in vitro*. Interestingly, we observed that both THG and hypoxia caused marked upregulation of TAN-1 marker expression (figure 2H), indicating that ER stress and hypoxia are potent stimulators of TAN-1 polarisation.

TAN-1 is associated with worse prognosis in patients with PDAC

To validate these newly identified TAN subclusters in PDAC tissues, we performed IF staining, and confirmed the existence of neutrophils expressing marker genes of each TAN subcluster (VEGFA, NLRP3, MME and IFIT2) (figure 3A). To further investigate the spatial distribution of these TAN subtypes, we stained serial tumour sections with CD66b and marker genes of TAN subclusters using IHC (figure 3B, online supplemental figure S6A). The median percentage of $VEGFA^+$ TANs, $NLRP3^+$ TANs, MME^+ TANs and $IFIT2^+$ TANs among total cells in PDAC tissues were 2.6%, 1.4%, 4.7% and 0.9%, respectively. Notably, most of $VEGFA^+$ TANs were in proximity to malignant cells, spatially enabling them to conduct pro-tumour functions, while $IFIT2^+$ TANs were frequently present in fibrotic stromal tissues (figure 3B). Further analysis on clinical data revealed that $VEGFA^+$ TANs were significantly associated with later American Joint Committee on Cancer (AJCC) stage (online supplemental figure S6B) and worse prognosis (figure 3C, online supplemental figure S6C) in patients with PDAC.

To further confirm the clinical relevance of TAN-1 subcluster based on public database, we retrieved RNA-seq and clinical data from TCGA-PAAD cohort,[51] as well as the published dataset by Cao *et al*.[52] Consistent with results from our cohort, TAN-1 signature was an unfavourable prognostic factor in treatment-naïve patients with standard PDAC histology from those two cohorts (figure 3D–E).

Metabolism analysis revealed glycolytic switch in TANs, mostly upregulated in TAN-1

Considering the profound metabolic reprogramming of neutrophils in tumour microenvironment as described above (online supplemental figure S3C), and the significant impact of metabolic programmes on immune cell functions,[53 54] we analysed the hallmark metabolic features of each TAN subcluster, and found that the activities of glycolysis and hypoxia were significantly higher in TAN-1 compared with other neutrophil subclusters (figure 4A, online supplemental figure S7A,B), and were upregulated along neutrophil transition trajectory (figure 4B). Consistently, flow cytometry analysis revealed upregulated expression of glucose transporter GLUT1 and glycolytic enzymes HK2, PFKFB3 and LDHA in neutrophils expressing TAN-1 marker LGALS3 isolated from PDAC tissues (figure 4C, online supplemental figure S7C,D). To further explore whether the metabolic features of TANs is associated with their spatial distribution, we generated spatial transcriptomes from 2498 spots on PDAC tissue section (online supplemental figure S7E,F), with an average of approximately 13 095 UMIs and 3969 unique genes detected per spot. Interestingly, compared with the neutrophil-enriched spots in stromal area, the glycolytic activity of neutrophil-enriched spots within or adjacent to tumour area was significantly upregulated (figure 4D,E), implying the glycolytic switch of neutrophils induced by malignant cells in PDAC tumour microenvironment.

To further validate the glycolytic switch of TANs, we isolated paired PMNs and TANs from 19 patients with PDAC (figure 1A), and applied integrative multi-omics approach to compare their glycolysis activity on transcriptome, proteome and metabolome levels. Analysis of bulk RNA-seq data revealed 3920 genes differentially expressed between PMNs and TANs (online supplemental table S7), and pathway enrichment analysis validated the upregulation of glycolysis in TANs (figure 4F and online supplemental table S8). Similarly, according to protein expression profiles identified using data-independent acquisition mass spectrometry (DIA-MS)-based quantitative proteomics analysis (online supplemental table S9), glycolysis pathway was significantly enriched in TANs (figure 4G). Metabolomic analysis on neutrophil lysates also revealed the significant upregulation of several glycolytic intermediates in TANs (figure 4H,I).

Glycolytic switch enhances pro-tumour functions in TANs

Next, we investigated the association between glycolytic switch and pro-tumour functions of neutrophils. We focused on LDHA, which is a critical gene in glycolysis pathway (figure 4A), and also one of the marker genes of TAN-1 (figure 2D (i)). Metabolic assays revealed that LDHA overexpression in dHL-60 cells (online supplemental figure S8A,B) resulted in enhanced

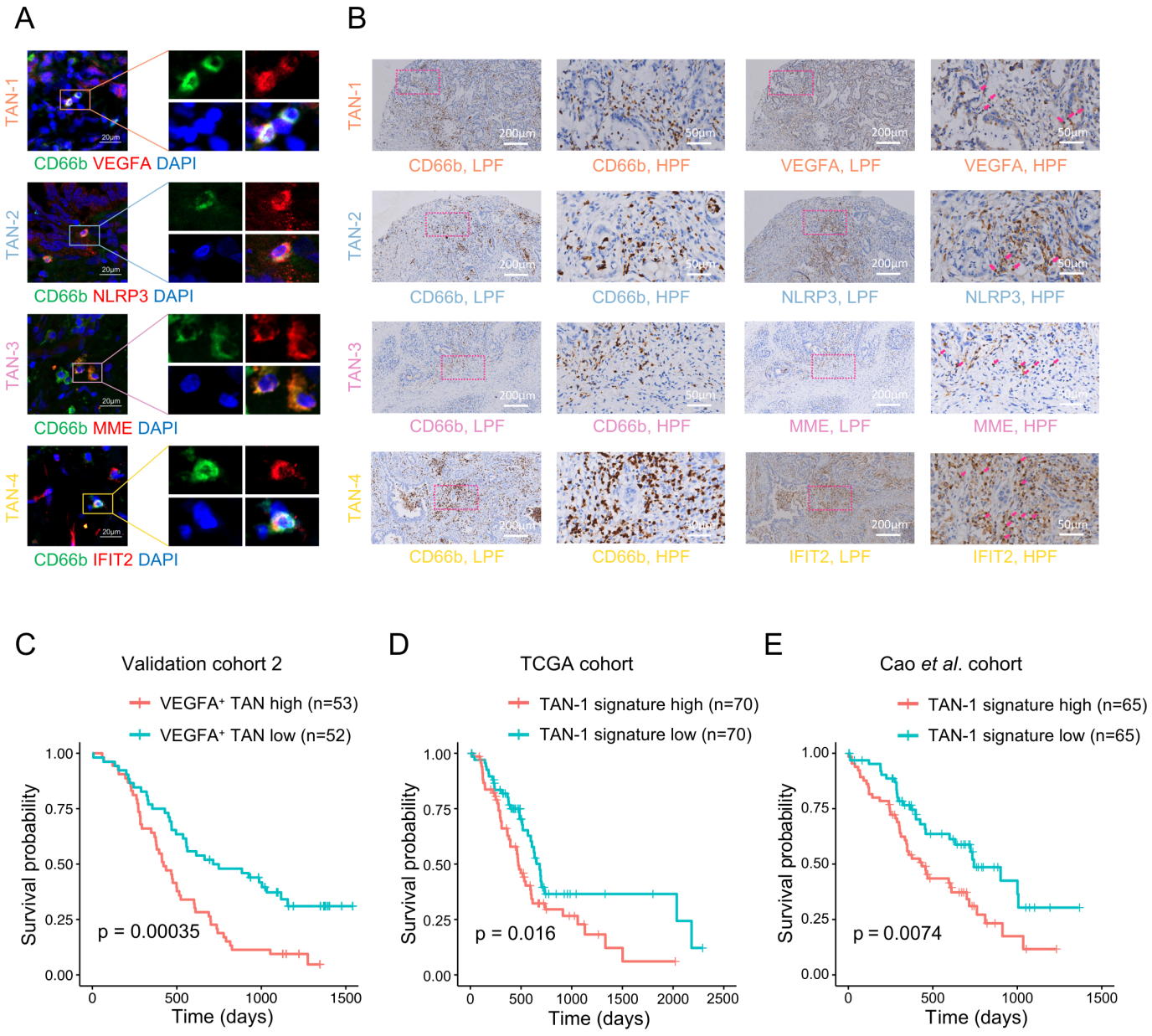

Figure 3 Spatial distribution and clinical relevance of TAN subclusters. (A) IF staining of marker genes of neutrophils (CD66b) and TAN subclusters (VEGFA for TAN-1, NLRP3 for TAN-2, MME for TAN-3 and IFIT2 for TAN-4) on PDAC tissue. (B) IHC images of representative PDAC tissues stained for marker genes of neutrophils (CD66b) and TAN subclusters (VEGFA for TAN-1, NLRP3 for TAN-2, MME for TAN-3 and IFIT2 for TAN-4) on serial slides. Neutrophils were identified according to CD66b staining on serial slides, and the polynucleated morphology. Pink arrows highlight the neutrophils expressing TAN subcluster markers. (C) Kaplan-Meier survival curve presenting the overall survival of patients with PDAC in IHC analysis. The patients were divided equally into two groups according to the percentage of VEGFA$^+$ TANs among total cells in PDAC tissues. (D) Kaplan-Meier survival curve presenting the overall survival of treatment-naïve patients with standard PDAC histology from TCGA-PAAD cohort. The patients were divided equally into two groups according to the expression of TAN-1 signature, assessed by GSVA of TAN-1 marker gene expression in each sample. (E) Kaplan-Meier survival curve presenting the overall survival of treatment-naïve patients with standard PDAC histology from the published dataset by Cao *et al*. The patients were divided equally into two groups according to the expression of TAN-1 signature, assessed by GSVA of TAN-1 marker gene expression in each sample. GSVA, gene set variation analysis; HPF, high power field; IHC, immunohistochemistry; IF, immunofluorescence; LPF, low power field; PDAC, pancreatic ductal adenocarcinoma; TAN, tumour-associated neutrophils .

glucose consumption (online supplemental figure S8C) and lactate production (online supplemental figure S8D), confirming the upregulation of glycolytic activity in LDHA-overexpressed dHL-60. PDAC tumour cells (PATU-8988 or Aspc-1) co-cultured with LDHA-overexpressed dHL-60 showed enhanced proliferation ability (online supplemental figure S8E-H), unravelling the pro-tumour effects resulted from metabolic reprogramming in neutrophils. Previous studies demonstrated that LDHA-associated lactic acid accumulation in tumour microenvironment inhibited anti-tumour immunosurveillance,[55] which prompted us to study the impact of LDHA overexpression on immunosuppressive functions of neutrophils. Interestingly, we observed that co-culturing with LDHA-overexpressed dHL-60 cells resulted in reduced expression of IFNγ and TNFα in CD8$^+$ T cells stimulated with phorbol 12-myristate 13-acetate (PMA) and ionomycin, as well as slightly suppressed T-cell proliferation

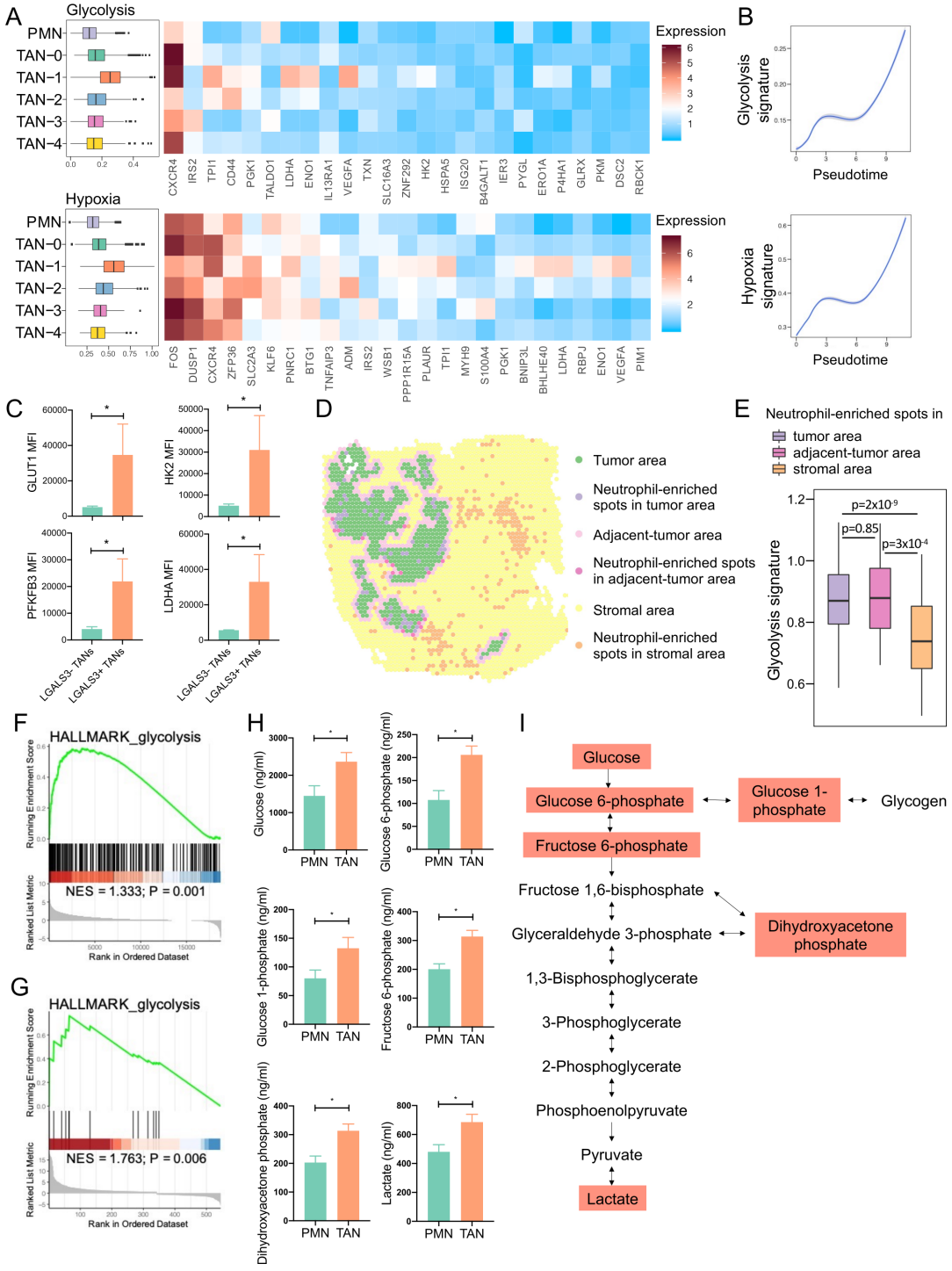

Figure 4 Metabolism analysis revealed glycolytic switch in TANs, mostly upregulated in TAN-1. (A) Box plots of expression of hallmark glycolysis and hypoxia signature in each neutrophil subcluster (left panel), and heatmap of average normalised expression of genes in signature (right panel). (B) Two-dimensional plots showing the dynamic expression of glycolysis signature and hypoxia signature along pseudotime trajectory. (C) Flow cytometry analysis of the expression of GLUT1, HK2, PFKFB3 and LDHA in LGALS3$^+$ and LGALS3$^-$ TANs from PDAC tissues. MFI of GLUT1, HK2, PFKFB3 and LDHA expression in LGALS3$^+$ and LGALS3$^-$ TANs were measured in tumour tissues from four patients with PDAC, and represented as mean with SD. (D) Spatial feature plot of neutrophil-enriched spots in tumour, adjacent-tumour and stromal area. Neutrophil signatures of each spatial spot were calculated with single sample gene set enrichment analysis (ssGSEA) based on the marker genes of neutrophils identified according to single-cell sequencing, and the spots with top 10% neutrophil signatures (ssGSEA score >0.53) were defined as neutrophil-enriched spots. (E) Box plot showing glycolysis signatures of neutrophil-enriched spots in tumour, adjacent-tumour and stromal area. (F–G) GSEA plots showing the NES for HALLMARK glycolysis pathway in TANs. Pathway enrichment analysis was performed based on transcriptomics (F) and proteomics (G) data of paired CD66b$^+$ PMNs and TANs from patients with PDAC. (H) Metabolomic analysis of glycolysis intermediates in PMN and TAN lysates. Data were represented as mean with SD. (I) Diagram of glycolysis pathway. Red frames indicated the metabolites significantly increased ($p<0.05$) in TANs in comparison with PMNs from patients with PDAC. MFI, mean fluorescence intensity; NES, normalised enrichment score; PDAC, pancreatic ductal adenocarcinoma; PMN, polymorphonuclear leucocytes; TAN, tumour-associated neutrophils. *P<0.05.

in response to CD3/CD28 activation (online supplemental figure S8I,J).

Analysis of upstream regulators revealed that BHLHE40 drives neutrophils towards pro-tumour phenotype

SCENIC[31] was applied to assess upstream transcription factors (regulons) driving the heterogeneity of TANs (online supplemental figure S9A). We identified four TAN clusters with distinctive regulon activity, associated with the original partition of TAN subclusters based on RNA profiles (online supplemental figure S9B,C). Of note, BHLHE40, one of the marker genes of TAN-1, was among the transcription factors most significantly upregulated in TAN-1 (figure 5A, online supplemental figure S9A). According to the gene regulatory network constituted by SCENIC, BHLHE40 is a downstream target gene of transcription factors HIF1A and XBP1, and BHLHE40 regulates the expression of TAN-1 markers VEGFA, PLAU, LGALS3, LDHA and BHLHE40 (autoregulation), and TAN-2 markers PDE4B and IL1RN (figure 5B), suggesting that BHLHE40, induced by hypoxia and ER stress, promotes the differentiation of neutrophils towards pro-tumour phenotype in tumour microenvironment.

To further validate the regulatory effects of BHLHE40, we analysed the influence of BHLHE40 overexpression on the expression profiles of neutrophil-like dHL-60 cells. qPCR confirmed that BHLHE40 overexpression resulted in significant upregulation of the expression of VEGFA, PLAU, LGALS3, LDHA and PDE4B (figure 5C). Chromatin immunoprecipitation-qPCR assays also confirmed the binding of BHLHE40 to the promoter regions of those genes (figure 5D), demonstrating the direct transcriptional regulation of pro-tumour genes by BHLHE40 in neutrophils. Interestingly, consistent with the regulatory network inferred by SCENIC, we observed significant induction of TAN-1 marker expression by hypoxia and ER stress, and a strong synergistic effect of these two factors, while the induction of TAN-1 markers was dramatically decreased in BHLHE40 knockdown dHL-60 cells (figure 5E), demonstrating that BHLHE40 plays a critical role in polarisation of neutrophils towards TAN-1 phenotype.

Next, we sought to explore the influence of BHLHE40 on pro-tumour functions of neutrophils. Interestingly, we observed significantly enhanced proliferation and migration capacity in PDAC cells co-cultured with BHLHE40-overexpressed dHL-60 (figure 6A,B, online supplemental figure S9D-F), demonstrating that the upstream regulator BHLHE40 drives neutrophils towards pro-tumour subtype. In addition, immunological assessment revealed that BHLHE40-overexpressed dHL-60 cells exerted suppressive effect on pro-inflammatory cytokine production of $CD8^+$ T cells as well as proliferation capacity of lymphocytes (figure 6C,D), suggesting BHLHE40 as a potential regulator of myeloid-derived suppressor cells (MDSCs). Nuclear BHLHE40-positive TANs were identified according to IF staining (online supplemental figure S9G), and IHC analysis revealed co-localisation of $BHLHE40^+$ neutrophils and neutrophils expressing TAN-1 marker gene VEGFA in PDAC microenvironment (figure 6E). Of note, higher infiltration level of $BHLHE40^+$ neutrophils was associated with worse prognosis (figure 6F), further demonstrating the pro-tumour role of the upstream regulator BHLHE40 in TANs.

DISCUSSION

Cancers develop in complicated microenvironment, in which the tumour-associated stroma is hijacked to create a favourable niche supporting tumour progression.[3] On the other hand, given the functional plasticity of stromal cells, they are also capable of inhibiting tumour growth on re-education,[56] which provides insights for development of novel therapeutic strategies against malignancies. In genetically engineered mouse models of PDAC, inhibiting neutrophil infiltration in tumour microenvironment resulted in tumour regression and prolonged survival,[15 57] demonstrating the roles of TANs in supporting PDAC progression, suggesting that TANs could be a potential therapeutic target.

Previous single cell RNA-seq studies showed the comprehensive gene expression atlas of main cell types in PDAC microenvironment.[16–21] However, the features of neutrophils are poorly studied, despite the fact that substantial TAN infiltration is present in PDAC and associated with unfavourable prognosis. In this study, we aimed to explore the neutrophil heterogeneity in PDAC tumour microenvironment at single-cell level, to identify the particular subpopulation that contributes to tumour progression, and to investigate the underlying mechanisms driving neutrophils towards pro-tumour phenotype. In an effort to characterise the reprogramming of neutrophils in tumour microenvironment, the counterpart in peripheral blood was also profiled as control, based on the fact that the neutrophils in tumour microenvironment are constantly recruited from circulation in response to chemotactic factors,[9] and neutrophil infiltration was uncommon in non-carcinomatous tissues.[58]

As human neutrophils have a short half-life *in vivo* (8 hours) and are extremely delicate and vulnerable *in vitro*,[22] it is difficult to capture them and decipher their reprogramming in tumour microenvironment. Of note, neutrophils were not detected in most of previous single-cell studies of human PDAC using 10x Genomics Chromium platform,[16–21] although there was only one exception, in which Steele *et al* successfully profiled granulocytes expressing typical marker genes from PDAC tissues.[59] Alternatively, we successfully acquired expression profiles of 21 972 neutrophils with BD Rhapsody. Similarly, in a study on prostate cancer, granulocytes were detected in the samples sequenced by BD Rhapsody, whereas not acquired using 10x Genomics.[24] Therefore, it is meaningful to directly compare these two single-cell sequencing platforms in their ability of capturing neutrophils.

The short lifespan and vulnerability of neutrophils also makes it challenging to study their functions *in vitro*. Here, we took advantage of the HL-60 cell line, acute promyelocytic leukaemia cells that could be differentiated into neutrophil-like state when treated with dimethyl sulfoxide or several other stimuli.[60 61] This enabled us to study the pro-tumour and immunosuppressive functions of neutrophils and to investigate how particular factors influence their phenotypes and functional states. Although dHL-60 cells could not completely substitute for neutrophils, we believe that experiments with this model provided insights for neutrophil biology.

Our data provided evidence for the continuum of transitional states of neutrophils in PDAC microenvironment: through transendothelial migration of neutrophils from circulation into PDAC microenvironment, the neutrophils presented the phenotype of transitional state (TAN-3), and then turned to the intermediate state (TAN-0), that represent the majority of neutrophils in tumour microenvironment without distinctive features. A small proportion of neutrophils were activated by IFN signals and acquired N1-like phenotype (TAN-4).[9 62] Through education in tumour microenvironment, neutrophils were gradually converted into inflammatory status (TAN-2), that contributed to cancer-associated inflammation, and terminally

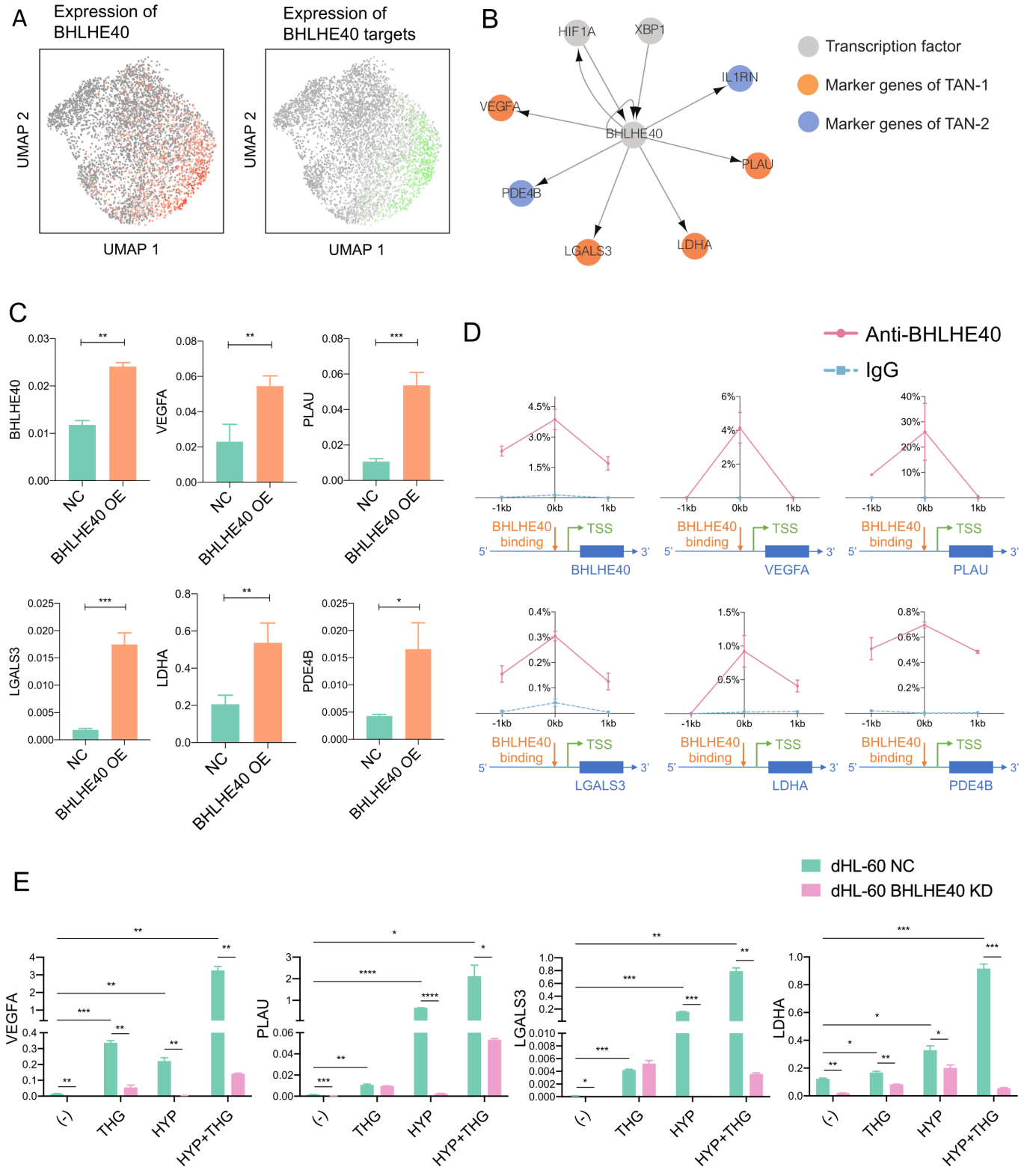

Figure 5 Analysis of upstream regulators revealed BHLHE40 as the key transcription factor of TAN-1. (A) UMAP plots displaying the expression of BHLHE40 in TANs (left panel), and the area under the curve (AUC) score of estimated regulon activity of BHLHE40 in TANs (right panel). (B) Representative regulatory network of BHLHE40 revealed by Single-Cell Regulatory Network Inference and Clustering (SCENIC). (C) Expression of BHLHE40 targets in control and BHLHE40-overexpressed dHL-60 cells, analysed by qPCR. The mRNA expression levels of each gene were normalised to fold over β-actin (housekeeping gene). Data were represented as mean with SD. (D) Binding of BHLHE40 at the promoter regions of target genes in dHL-60 cells, analysed by chromatin immunoprecipitation-qPCR. (E) Expression of BHLHE40 targets in control and BHLHE40 knockdown dHL-60 cells stimulated with THG and/or HYP for 24 hours, analysed by qPCR. The mRNA expression levels of each gene were normalised to fold over β-actin (housekeeping gene). Data were represented as mean with SD . *P<0.05; **p<0.01; ***p<0.001; ****p<0.0001. dHL-60, neutrophil-like differentiated HL-60; HYP, hypoxia; KD, knockdown; mRNA, messenger RNA; NC, negative control; OE, overexpression; THG, thapsigargin; TSS, transcription start site; qPCR, quantitative PCR; TAN, tumour-associated neutrophils; UMAP, uniform manifold approximation and projection.

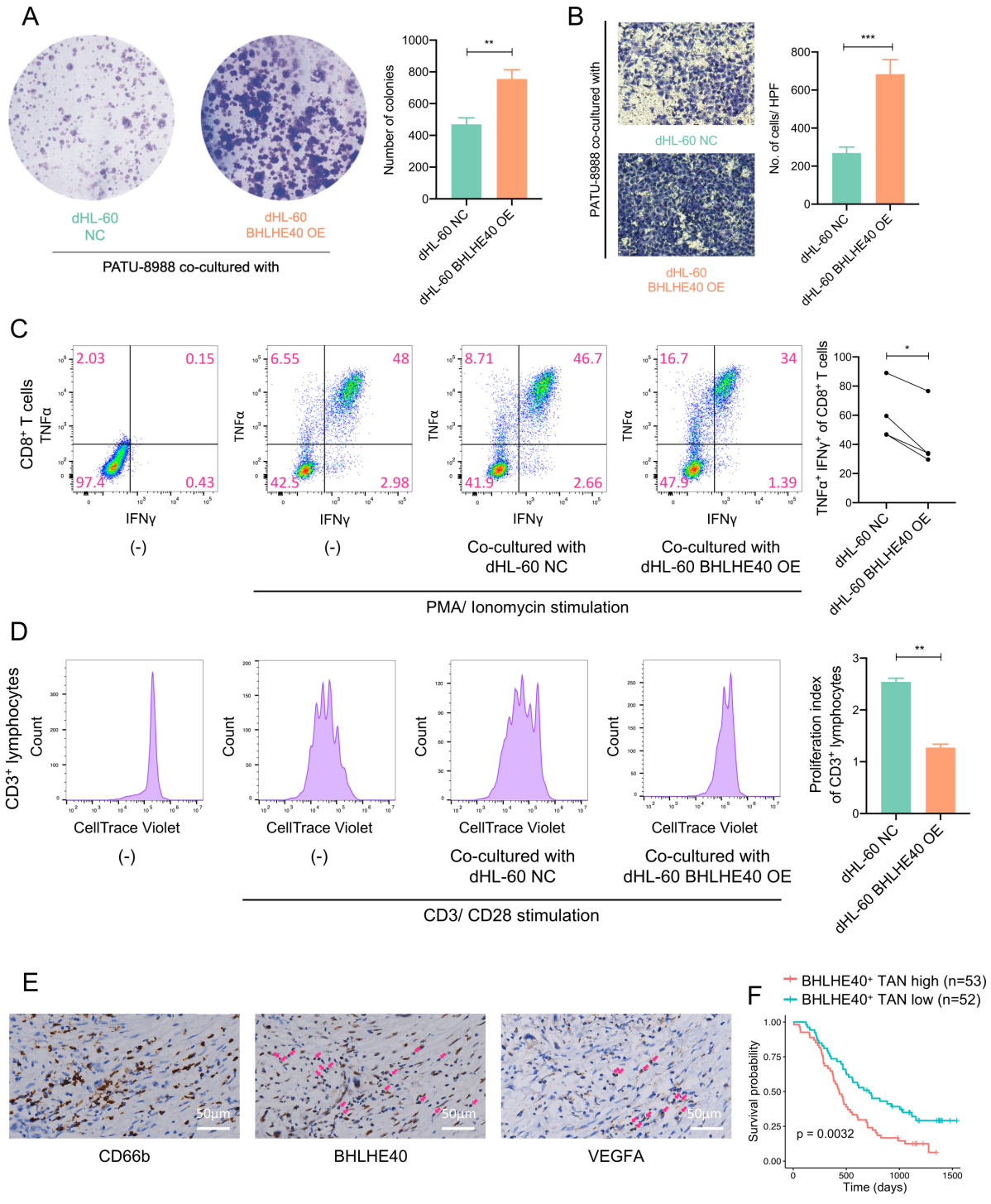

Figure 6 BHLHE40 drives neutrophils towards pro-tumour phenotype. (A) Colony formation assay, in which control and BHLHE40-overexpressed dHL-60 cells were cultured in upper chambers, and PATU-8988 cells were cultured in lower chambers. The numbers of colonies were represented as mean with SD. (B) PATU-8988 cells were co-cultured with control and BHLHE40-overexpressed dHL-60 cells for 3 days, and their migration capacity was assessed by transwell assay. The numbers of migrated cells were represented as mean with SD. (C) PBMCs were co-cultured with control and BHLHE40-overexpressed dHL-60 cells for 3 days, stimulated with PMA/ionomycin, and the percentage of IFNγ$^+$ and TNFα$^+$ cells among stimulated CD8$^+$ T cells was analysed by flow cytometry. Data from four different donors were summarised in the right panel. (D) PBMCs were stained with CellTrace Violet, stimulated with anti-human CD3/CD28 for 4 days in the absence or presence of dHL-60 cells (control or BHLHE40-overexpressed), and the proliferation of CD3$^+$ lymphocytes was analysed by flow cytometry. The proliferation index of triplicate cultures were represented as mean with SD. (E) IHC images of representative PDAC tissues stained for CD66b (left), BHLHE40 (middle) and VEGFA (right) on serial slides. Neutrophils were identified according to CD66b staining, and the polynucleated morphology. Pink arrows highlight the neutrophils expressing BHLHE40 or VEGFA. (F) Kaplan-Meier survival curve presenting the overall survival of patients with PDAC in IHC analysis. The patients were divided equally into two groups according to the percentage of BHLHE40$^+$ TANs among total cells in PDAC tissues . *P<0.05; **p<0.01; ***p<0.001. dHL-60, neutrophil-like differentiated HL-60; HPF, high power field; IFN, interferon; IHC, immunohistochemistry; NC, negative control; OE, overexpression; PBMCs, peripheral blood mononuclear cells; PDAC, pancreatic ductal adenocarcinoma; TAN, tumour-associated neutrophils; TNF, tumour necrosis factor.

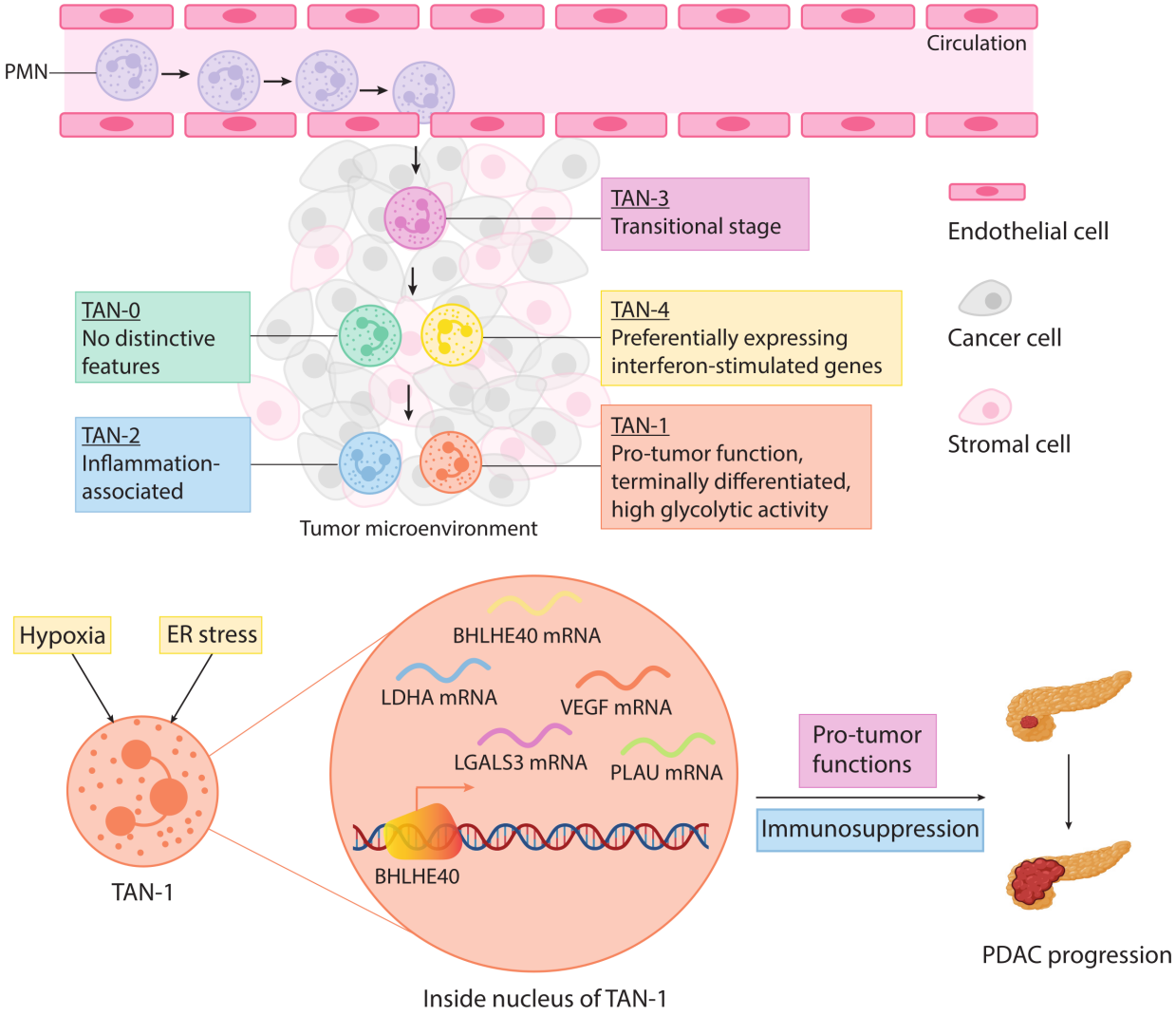

Figure 7 Graphical summary of major findings in this study. ER, endoplasmic reticulum; PDAC, pancreatic ductal adenocarcinoma; PMN, polymorphonuclear leucocytes; TAN, tumour-associated neutrophils.

differentiated to TAN-1, the 'pathogenic subpopulation' preferentially expressing pro-tumour molecules, associated with worse prognosis in patients with PDAC (figure 7). Zilionis *et al* deciphered the diversity of neutrophils in lung cancers, and reported three modules of expression: (i) neutrophils expressing canonical neutrophil markers (S100A8 and S100A9), similar with TAN-3 in our study; (ii) neutrophils expressing inflammatory cytokines and pro-tumour genes, corresponding to TAN-1 and TAN-2; (iii) neutrophils displaying strong expression of type I IFN-response genes, similar with TAN-4.[63] Comparison of these two single-cell datasets suggested the similarity of TAN population structures in different cancer types.

It has been widely accepted that metabolism states influence phenotype and polarisation of immune cells.[53 54] In this study, pro-tumour TAN-1 subpopulation exhibited hyperactivated glycolytic activity (figure 4A), and glycolytic switch was associated with pro-tumour functions in neutrophils (online supplemental figure S8). Our results is agreement with the 'reverse Warburg effect' model, that glycolysis in non-malignant stromal cells result in production of energy-rich metabolites such as lactate and pyruvate, which can be taken up by cancer cells as an alternative TCA cycle substrate, facilitating energy production and tumour growth.[64 65] Besides enhancing the proliferation of malignant cells, it has been discovered that glycolytic enzyme LDHA-associated lactic acid could abrogate tumour immunosurveillance.[55 66 67] Consistently, LDHA overexpression in dHL-60 cells resulted in suppression of T-cell proliferation and activation (online supplemental figure S8). Collectively, glycolysis could be a potential therapeutic target to reverse the pro-tumour and immunosuppressive activity of neutrophils in PDAC.

BHLHE40 has been emerged as a key regulator of immunity. It has been reported that BHLHE40 regulated cytokine production in T cells, and the proliferation of macrophages in both homeostasis and type 2 immunity.[68 69] However, little is known regarding its role in neutrophils. In our study, BHLHE40 has been identified as a downstream target of both hypoxia and ER stress, the two potent stimulators in tumour microenvironment driving neutrophils towards pro-tumour TAN-1 phenotype. Importantly, our data demonstrated that BHLHE40 was a key regulator of pro-tumour neutrophils (figure 6). PMN-MDSCs have been defined as a population of pathologically activated neutrophils with immunosuppressive functions.[70 71] Hypoxia-induced HIF1α activation and ER stress have been recognised as important regulators converting neutrophils to immunosuppressive PMN-MDSCs in tumour microenvironment.[71–73] Interestingly, our data also demonstrated the important role of

BHLHE40 in regulating the immunosuppressive functions of neutrophils (figure 6), suggesting BHLHE40 as a potential regulator of PMN-MDSCs. Taken together, our study shed light on the potential of BHLHE40 as a therapeutic target for TANs in PDAC.

In conclusion, we identified a pro-tumour subcluster of neutrophils in PDAC tumour microenvironment, revealed the association between high glycolytic activity and pro-tumour functions in TANs, and demonstrated hypoxia-induced and ER stress-induced BHLHE40 activation as underlying mechanism driving TANs towards pro-tumour phenotype.

Author affiliations
[1]Department of General Surgery, Pancreatic Disease Center, Ruijin Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, People's Republic of China
[2]Research Institute of Pancreatic Diseases, Shanghai Key Laboratory of Translational Research for Pancreatic Neoplasms, Shanghai Jiaotong University School of Medicine, Shanghai, People's Republic of China
[3]State Key Laboratory of Oncogenes and Related Genes, Institute of Translational Medicine, Shanghai Jiaotong University, Shanghai, People's Republic of China
[4]Shanghai Institute of Hematology, State Key Laboratory of Medical Genomics, National Research Center for Translational Medicine at Shanghai, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, People's Republic of China
[5]Department of Pathology, Ruijin Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, People's Republic of China
[6]Center for Biomedical Big Data, the First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang Province, People's Republic of China
[7]Biomedical Pioneering Innovation Center (BIOPIC), School of Life Sciences, Peking University, Beijing, People's Republic of China
[8]Sino-French Research Center for Life Sciences and Genomics, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, People's Republic of China

Acknowledgements We thank NovelBio for the support of single-cell sequencing. We thank Dr Bing Su and Dr Jing Wang from Shanghai Institute of Immunology for critically reading our manuscript and providing insightful suggestions.

Contributors LW, LJ and BS designed the study and wrote the manuscript. LW, YL, XT, MS, JQ, MX, YC, JL and PL conducted experiments. LW, YD and JH performed bioinformatics analysis. TY participated in proteomics and metabolomics analysis. CWa, TW and LD participated in pathological work. CWe, JZ, ZX and HC participated in patient recruitment and sample collection. FB, XD, CP and LJ helped optimise the research and proofread the paper. SC and BS supervised the study and revised the manuscript. All authors read and approved the final manuscript. BS, SC and LJ are responsible for the overall content as the guarantor.

Funding This work was supported by the National Natural Science Foundation of China (No. 81871906 and 82073326 to BS, 81802316 to LJ, 82002460 to LW), China Postdoctoral Science Foundation (No. 2019M661552 to LW) and Shanghai Pilot Program for Basic Research-Shanghai Jiao Tong University (No. 21TQ1400205 to LJ).

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study was approved by The Ethics Committees of Ruijin Hospital affiliated to School of Medicine, Shanghai Jiaotong University (reference number 2021-161). Participants were given informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Single-cell and bulk RNA sequencing data generated in this study are deposited at the National Omics Data Encyclopedia (NODE) with the accession code OEP003254.

ORCID iD
Baiyong Shen http://orcid.org/0000-0002-3994-248X

REFERENCES

1 Mizrahi JD, Surana R, Valle JW, *et al*. Pancreatic cancer. *Lancet* 2020;395:2008–20.
2 Schizas D, Charalampakis N, Kole C, *et al*. Immunotherapy for pancreatic cancer: a 2020 update. *Cancer Treat Rev* 2020;86:102016.
3 Ho WJ, Jaffee EM, Zheng L. The tumour microenvironment in pancreatic cancer - clinical challenges and opportunities. *Nat Rev Clin Oncol* 2020;17:527–40.
4 Vonderheide RH, Bayne LJ. Inflammatory networks and immune surveillance of pancreatic carcinoma. *Curr Opin Immunol* 2013;25:200–5.
5 Gajewski TF, Schreiber H, Fu Y-X. Innate and adaptive immune cells in the tumor microenvironment. *Nat Immunol* 2013;14:1014–22.
6 Hinshaw DC, Shevde LA. The tumor microenvironment Innately modulates cancer progression. *Cancer Res* 2019;79:4557–66.
7 Looi C-K, Chung FF-L, Leong C-O, *et al*. Therapeutic challenges and current immunomodulatory strategies in targeting the immunosuppressive pancreatic tumor microenvironment. *J Exp Clin Cancer Res* 2019;38:162.
8 Karamitopoulou E. Tumour microenvironment of pancreatic cancer: immune landscape is dictated by molecular and histopathological features. *Br J Cancer* 2019;121:5–14.
9 Jaillon S, Ponzetta A, Di Mitri D, *et al*. Neutrophil diversity and plasticity in tumour progression and therapy. *Nat Rev Cancer* 2020;20:485–503.
10 Mantovani A, Marchesi F, Jaillon S, *et al*. Tumor-associated myeloid cells: diversity and therapeutic targeting. *Cell Mol Immunol* 2021;18:566–78.
11 Gentles AJ, Newman AM, Liu CL, *et al*. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med* 2015;21:938–45.
12 Granot Z, Henke E, Comen EA, *et al*. Tumor entrained neutrophils inhibit seeding in the premetastatic lung. *Cancer Cell* 2011;20:300–14.
13 Cui C, Chakraborty K, Tang XA, *et al*. Neutrophil elastase selectively kills cancer cells and attenuates tumorigenesis. *Cell* 2021;184:3163–77.
14 van Egmond M, Bakema JE. Neutrophils as effector cells for antibody-based immunotherapy of cancer. *Semin Cancer Biol* 2013;23:190–9.
15 Chao T, Furth EE, Vonderheide RH. CXCR2-Dependent accumulation of tumor-associated neutrophils regulates T-cell immunity in pancreatic ductal adenocarcinoma. *Cancer Immunol Res* 2016;4:968–82.
16 Peng J, Sun B-F, Chen C-Y, *et al*. Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. *Cell Res* 2019;29:725–38.
17 Elyada E, Bolisetty M, Laise P, *et al*. Cross-Species single-cell analysis of pancreatic ductal adenocarcinoma reveals antigen-presenting cancer-associated fibroblasts. *Cancer Discov* 2019;9:1102–23.
18 Lin W, Noel P, Borazanci EH, *et al*. Single-cell transcriptome analysis of tumor and stromal compartments of pancreatic ductal adenocarcinoma primary tumors and metastatic lesions. *Genome Med* 2020;12:80.
19 Lee JJ, Bernard V, Semaan A, *et al*. Elucidation of tumor-stromal heterogeneity and the ligand-receptor interactome by single-cell transcriptomics in real-world pancreatic cancer biopsies. *Clin Cancer Res* 2021;27:5912–21.
20 Wang Y, Liang Y, Xu H, *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.
21 Moncada R, Barkley D, Wagner F, *et al*. Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nat Biotechnol* 2020;38:333–42.
22 Cassatella MA, Östberg NK, Tamassia N, *et al*. Biological roles of neutrophil-derived granule proteins and cytokines. *Trends Immunol* 2019;40:648–64.
23 Xie X, Shi Q, Wu P, *et al*. Single-Cell transcriptome profiling reveals neutrophil heterogeneity in homeostasis and infection. *Nat Immunol* 2020;21:1119–33.
24 Peng S, Hu P, Xiao Y-T, *et al*. Single-Cell analysis reveals EP4 as a target for restoring T-cell infiltration and sensitizing prostate cancer to immunotherapy. *Clin Cancer Res* 2022;28:552–67.
25 Chen S, Zhou Y, Chen Y, *et al*. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 2018;34:i884–90.
26 Smith T, Heger A, Sudbery I. UMI-tools: modeling sequencing errors in unique molecular identifiers to improve quantification accuracy. *Genome Res* 2017;27:491–9.
27 Dobin A, Davis CA, Schlesinger F, *et al*. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;29:15–21.

28 Stuart T, Butler A, Hoffman P, *et al*. Comprehensive integration of single-cell data. *Cell* 2019;177:1888–902.
29 Trapnell C, Cacchiarelli D, Grimsby J, *et al*. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* 2014;32:381–6.
30 Efremova M, Vento-Tormo M, Teichmann SA, *et al*. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc* 2020;15:1484–506.
31 Aibar S, González-Blas CB, Moerman T, *et al*. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* 2017;14:1083–6.
32 Krijthe J. Rtsne: T-Distributed stochastic neighbor embedding using Barnes-Hut implementation, 2015. Available: https://github.com/jkrijthe/Rtsne
33 Newton K, Dixit VM. Signaling in innate immunity and inflammation. *Cold Spring Harb Perspect Biol* 2012;4. doi:10.1101/cshperspect.a006049. [Epub ahead of print: 01 Mar 2012].
34 Chu W-M. Tumor necrosis factor. *Cancer Lett* 2013;328:222–5.
35 Weber A, Wasiliew P, Kracht M. Interleukin-1 (IL-1) pathway. *Sci Signal* 2010;3:cm1.
36 Fukumura D, Kloepper J, Amoozgar Z, *et al*. Enhancing cancer immunotherapy using antiangiogenics: opportunities and challenges. *Nat Rev Clin Oncol* 2018;15:325–40.
37 Lin M, Zhang Z, Gao M, *et al*. MicroRNA-193a-3p suppresses the colorectal cancer cell proliferation and progression through downregulating the PLAU expression. *Cancer Manag Res* 2019;11:5353–63.
38 Moquet-Torcy G, Tolza C, Piechaczyk M, *et al*. Transcriptional complexity and roles of Fra-1/AP-1 at the uPA/Plau locus in aggressive breast cancer. *Nucleic Acids Res* 2014;42:11011–24.
39 Vuong L, Kouverianou E, Rooney CM, *et al*. An orally active galectin-3 antagonist inhibits lung adenocarcinoma growth and augments response to PD-L1 blockade. *Cancer Res* 2019;79:1480–92.
40 Ruvolo PP. Galectin 3 as a guardian of the tumor microenvironment. *Biochim Biophys Acta* 2016;1863:427–37.
41 Moossavi M, Parsamanesh N, Bahrami A, *et al*. Role of the NLRP3 inflammasome in cancer. *Mol Cancer* 2018;17:158.
42 Azam MA, Tripuraneni NS. Selective phosphodiesterase 4B inhibitors: a review. *Sci Pharm* 2014;82:453–81.
43 Agraz-Cibrián JM, Delgado-Rizo V, Segura-Ortega JE, *et al*. Impaired neutrophil extracellular traps and inflammatory responses in the peritoneal fluid of patients with liver cirrhosis. *Scand J Immunol* 2018;88:e12714.
44 Tsuyusaki J, Kuroda F, Kasuya Y, *et al*. Cigarette smoke-induced pulmonary inflammation is attenuated in CD69-deficient mice. *J Recept Signal Transduct Res* 2011;31:434–9.
45 Di Mitri D, Toso A, Chen JJ, *et al*. Tumour-infiltrating Gr-1+ myeloid cells antagonize senescence in cancer. *Nature* 2014;515:134–7.
46 Larráyoz IM, Martínez-Herrero S, García-Sanmartín J, *et al*. Adrenomedullin and tumour microenvironment. *J Transl Med* 2014;12:339.
47 Yu Y, Araki Y, Sendo F. Tyrosine phosphorylation of a 34-kDa protein induced by cross-linking a novel glycosylphosphatidylinositol-anchored glycoprotein (GPI-80) on human neutrophils that may regulate their adherence and migration. *IUBMB Life* 2000;49:43–7.
48 Rahman I, Collado Sánchez A, Davies J, *et al*. L-selectin regulates human neutrophil transendothelial migration. *J Cell Sci* 2021;134. doi:10.1242/jcs.250340. [Epub ahead of print: 08 02 2021].
49 Schneider WM, Chevillotte MD, Rice CM. Interferon-stimulated genes: a complex web of host defenses. *Annu Rev Immunol* 2014;32:513–45.
50 Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics* 2013;14:7.
51 Cancer Genome Atlas Research Network. Data from: integrated genomic characterization of pancreatic ductal adenocarcinoma. *The Cancer Genome Atlas* 2020.
52 Cao L, Huang C, Cui Zhou D. Data from: Proteogenomic characterization of pancreatic ductal adenocarcinoma. *LinkedOmics* 2020.
53 Artyomov MN, Van den Bossche J. Immunometabolism in the single-cell era. *Cell Metab* 2020;32:710–25.
54 Makowski L, Chaib M, Rathmell JC. Immunometabolism: from basic mechanisms to translation. *Immunol Rev* 2020;295:5–14.
55 Brand A, Singer K, Koehl GE, *et al*. LDHA-Associated lactic acid production blunts tumor immunosurveillance by T and NK cells. *Cell Metab* 2016;24:657–71.
56 Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. *Nat Med* 2013;19:1423–37.
57 Steele CW, Karim SA, Leach JDG, *et al*. CXCR2 inhibition profoundly suppresses metastases and augments immunotherapy in pancreatic ductal adenocarcinoma. *Cancer Cell* 2016;29:832–45.
58 Reid MD, Basturk O, Thirabanjasak D, *et al*. Tumor-infiltrating neutrophils in pancreatic neoplasia. *Mod Pathol* 2011;24:1612–9.
59 Steele NG, Carpenter ES, Kemp SB, *et al*. Multimodal mapping of the tumor and peripheral blood immune landscape in human pancreatic cancer. *Nat Cancer* 2020;1:1097–112.
60 Blanter M, Gouwy M, Struyf S. Studying neutrophil function in vitro: cell models and environmental factors. *J Inflamm Res* 2021;14:141–62.
61 Rincón E, Rocha-Gregg BL, Collins SR. A map of gene expression in neutrophil-like cell lines. *BMC Genomics* 2018;19:573.
62 Coffelt SB, Wellenstein MD, de Visser KE. Neutrophils in cancer: neutral no more. *Nat Rev Cancer* 2016;16:431–46.
63 Zilionis R, Engblom C, Pfirschke C, *et al*. Single-Cell transcriptomics of human and mouse lung cancers reveals conserved myeloid populations across individuals and species. *Immunity* 2019;50:1317–34.
64 Pavlides S, Whitaker-Menezes D, Castello-Cros R, *et al*. The reverse Warburg effect: aerobic glycolysis in cancer associated fibroblasts and the tumor stroma. *Cell Cycle* 2009;8:3984–4001.
65 Yoshida GJ. Metabolic reprogramming: the emerging concept and associated therapeutic strategies. *J Exp Clin Cancer Res* 2015;34:111.
66 Colegio OR, Chu N-Q, Szabo AL, *et al*. Functional polarization of tumour-associated macrophages by tumour-derived lactic acid. *Nature* 2014;513:559–63.
67 Haas R, Smith J, Rocher-Ros V, *et al*. Lactate regulates metabolic and pro-inflammatory circuits in control of T cell migration and effector functions. *PLoS Biol* 2015;13:e1002202.
68 Cook ME, Jarjour NN, Lin C-C, *et al*. Transcription factor Bhlhe40 in immunity and autoimmunity. *Trends Immunol* 2020;41:1023–36.
69 Jarjour NN, Schwarzkopf EA, Bradstreet TR, *et al*. Bhlhe40 mediates tissue-specific control of macrophage proliferation in homeostasis and type 2 immunity. *Nat Immunol* 2019;20:687–700.
70 Zhou J, Nefedova Y, Lei A, *et al*. Neutrophils and PMN-MDSC: their biological role and interaction with stromal cells. *Semin Immunol* 2018;35:19–28.
71 Veglia F, Sanseviero E, Gabrilovich DI. Myeloid-derived suppressor cells in the era of increasing myeloid cell diversity. *Nat Rev Immunol* 2021;21:485–98.
72 Mortezaee K, Majidpoor J. The impact of hypoxia on immune state in cancer. *Life Sci* 2021;286:120057.
73 Condamine T, Dominguez GA, Youn J-I, *et al*. Lectin-type oxidized LDL receptor-1 distinguishes population of human polymorphonuclear myeloid-derived suppressor cells in cancer patients. *Sci Immunol* 2016;1. doi:10.1126/sciimmunol.aaf8943. [Epub ahead of print: 05 08 2016].

Supplementary figures

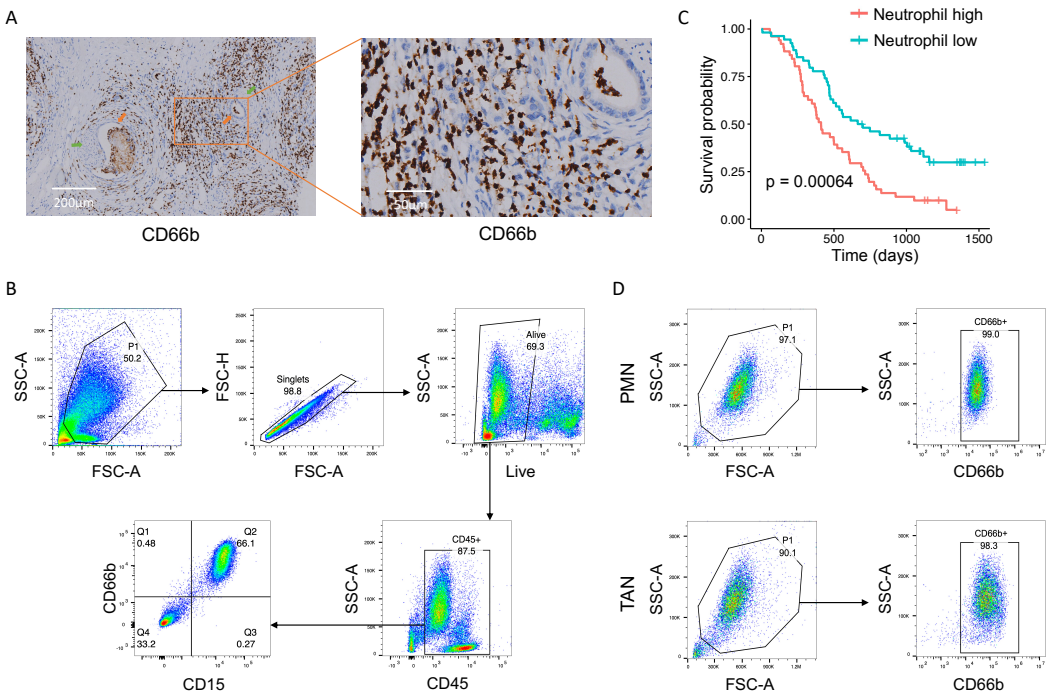

Figure S1. Neutrophils are important infiltrating immune cells in PDAC tumor microenvironment

(A) IHC images of representative PDAC tissues stained for marker genes of neutrophils (CD66b). On the left figure, two duct-like structures formed by malignant cells were highlighted by orange arrows, and peripheral nerves were highlighted by green arrows. Neural invasion of those ductal cells could be observed. (B) Flow cytometry analysis of percentage of $CD15^{+}$ $CD66b^{+}$ neutrophils among $CD45^{+}$ infiltrating immune cells in PDAC tissues. (C) Kaplan-Meier survival curve presenting the overall survival of PDAC patients, divided equally into two groups according to the percentage of infiltrating neutrophils. (D) Typical phenotype of PMNs and TANs isolated from peripheral blood and tumor tissue of PDAC patients using CD66b positive selection. Abbreviations: FSC, forward scatter; SSC, side scatter; PMN, polymorphonuclear leukocytes; TAN, tumor-associated neutrophils.

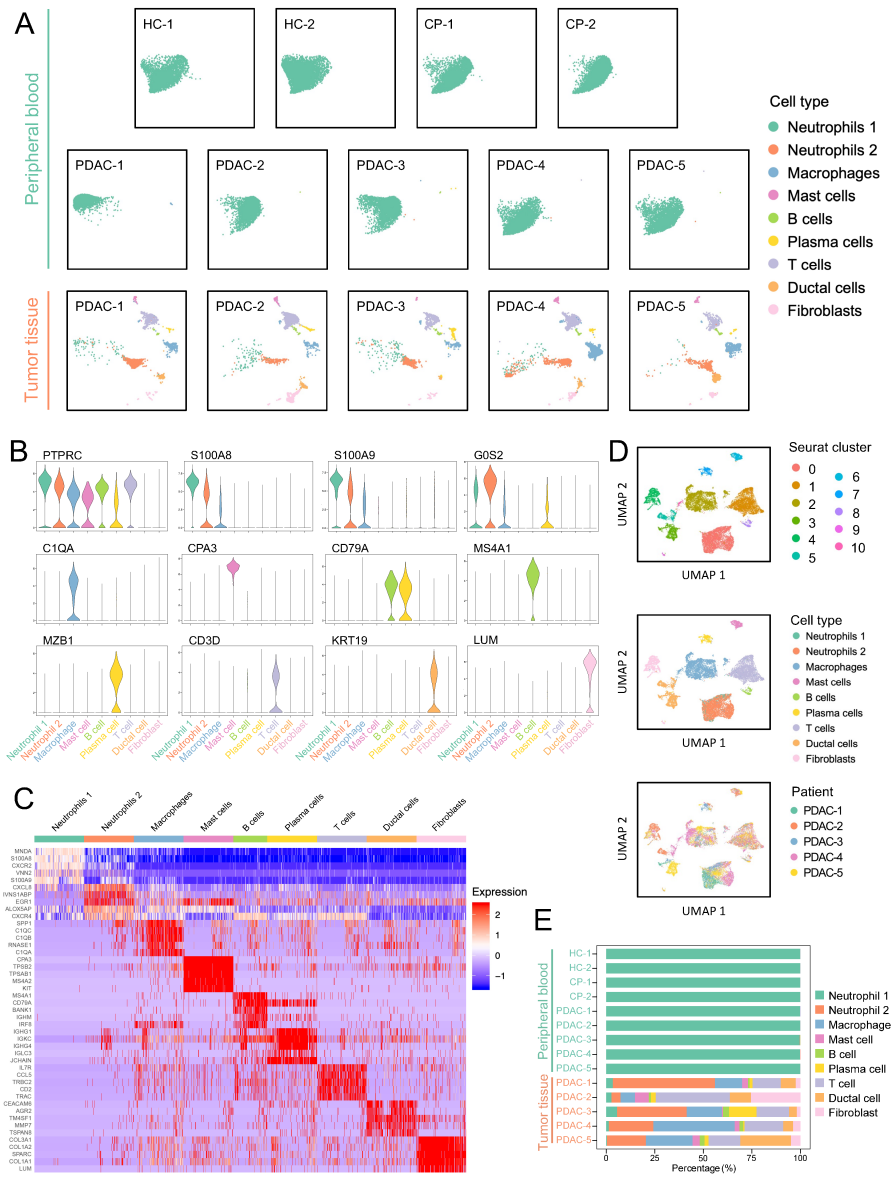

Figure S2. Single-cell atlas from peripheral blood and PDAC tumor tissues, related to Figure 1
(A) UMAP plots of the single cells identified from each sample, with each cell color-coded for cell type. (B) Violin plots showing the expression of representative classical cell type markers. (C) Heatmap showing the expression of top five marker genes in each cell type. (D) Unsupervised clustering of cells from PDAC tissue separately, visualized by UMAP plots, with each cell color-coded for clusters (upper panel), cell type (as annotated in Figure 1B) (middle panel), and the corresponding donor (lower panel). (E) Bar plot depicting the proportions of cell types identified in each sample. Abbreviations: HC, healthy controls; CP, chronic pancreatitis; PDAC, pancreatic ductal adenocarcinoma; UMAP, uniform manifold approximation and projection.

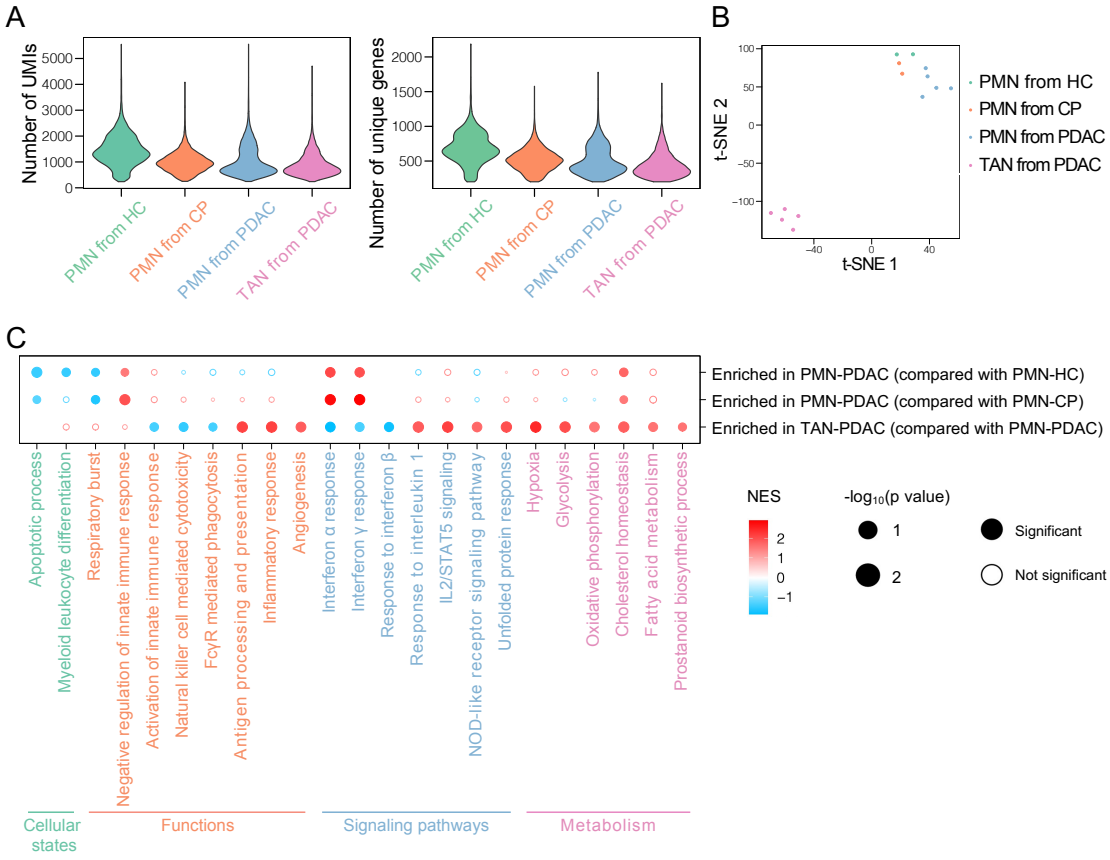

Figure S3. Comparison of the expression profiles of neutrophils from different sources
(A) Violin plots summarizing number of UMIs and number of unique genes detected in neutrophils from different sources. (B) Visualization of the expression profiles of neutrophils from each sample by t-distributed stochastic neighbor embedding (t-SNE) dimensionality reduction analysis. (C) Dot plot showing representative enriched HALLMARK/ KEGG/ GO pathways in TANs and PMNs from PDAC patients. Pathway enrichment analysis was performed with gene set enrichment analysis (GSEA). NES and p values were indicated by circle color and size, respectively. Abbreviations: UMI, unique molecular identifier; HC, healthy controls; CP, chronic pancreatitis; PDAC, pancreatic ductal adenocarcinoma; PMN, polymorphonuclear leukocytes; TAN, tumor-associated neutrophils; t-SNE, t-distributed stochastic neighbor embedding; NES, normalized enrichment score.

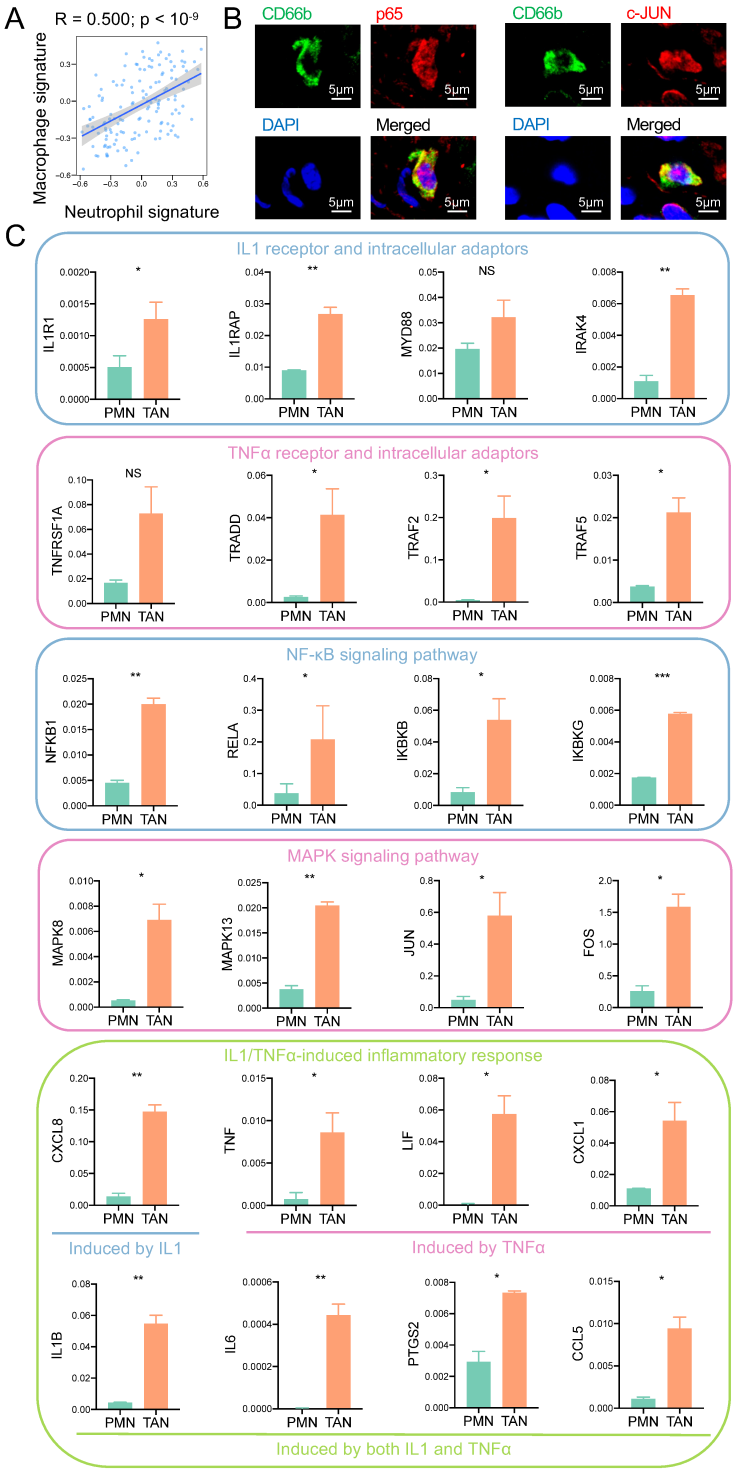

Figure S4. Activation of TNFα and IL1 signaling pathways in TANs, related to Figure 1

(A) Scatterplot showing the expression of neutrophil signatures and macrophage signatures in treatment-naïve PDAC samples from TCGA-PAAD cohort. The association between these variables was assessed by Pearson correlation analysis. (B) Immunofluorescence (IF) staining of CD66b, p65 and c-JUN on PDAC tissue. (C) Comparison of the expression of genes associated with IL1/TNFα signaling between PMNs and TANs isolated from PDAC patients, with qPCR analysis. qPCR data were normalized to fold over β-actin (housekeeping gene), and represented as mean with standard deviation. Abbreviations: PMN, polymorphonuclear leukocytes; TAN, tumor-associated neutrophils.

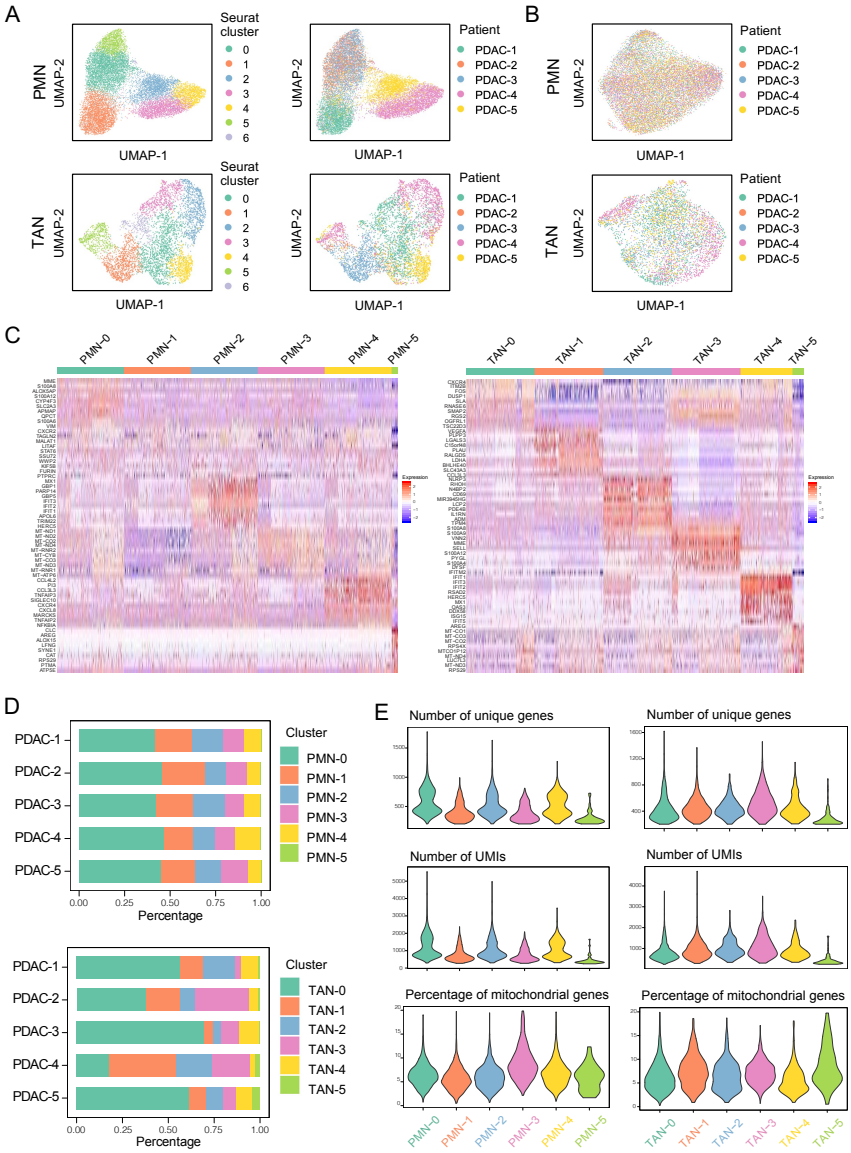

Figure S5. Subclusters of PMNs and TANs from PDAC patients, related to Figure 2

(A) Unsupervised clustering of PMNs and TANs from PDAC patients without batch effect correction, visualized with UMAP plots, with each cell color-coded for clusters (left panel), and the corresponding donor (right panel). (B) UMAP plots of PMNs and TANs from PDAC patients, with each cell color-coded for the corresponding donor, related to Figures 2A-B. (C) Heatmap showing expression of top ten marker genes in each neutrophil cluster. (D) Bar plot depicting the proportions of PMN subclusters and TAN subclusters identified in each sample. (E) Violin plots summarizing number of unique genes, number of UMIs, and percentage of mitochondrial genes in each cell across neutrophil subcluster. Abbreviations: PMN, polymorphonuclear leukocytes; TAN, tumor-associated neutrophils; PDAC, pancreatic ductal adenocarcinoma; UMAP, uniform manifold approximation and projection; UMI, unique molecular identifier.

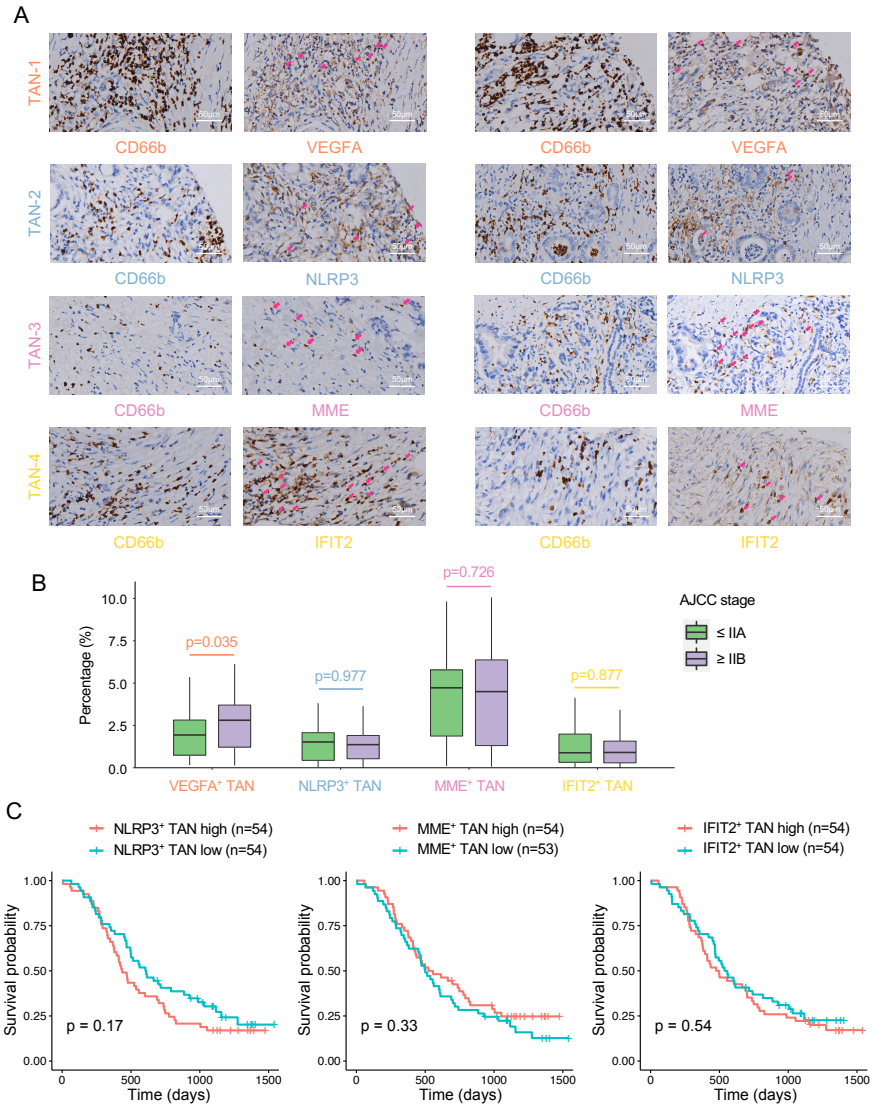

Figure S6. Spatial distribution and clinical relevance of TAN subclusters, related to Figure 3
(A) IHC images of additional representative PDAC tissues stained for marker genes of neutrophils (CD66b) and TAN subclusters (VEGFA for TAN-1, NLRP3 for TAN-2, MME for TAN-3, and IFIT2 for TAN-4) on serial slides. Neutrophils were identified according to CD66b staining on serial slides, and the polynucleated morphology. Pink arrows highlight the neutrophils expressing TAN subcluster markers. (B) Boxplot summarizing the percentages of VEGFA$^+$ TANs, NLRP3$^+$ TANs, MME$^+$ TANs, and IFIT2$^+$ TANs among total cells in PDAC tissues from patients with early-stage disease (AJCC stage ≤IIA) and late-stage disease (AJCC stage ≥IIB). The middle lines represented median values, the boxes represented inter-quartile ranges, and the whiskers extended to the furthest data point within 1.5 times the inter-quartile ranges. Difference between the groups was analyzed by Mann-Whitney U test. (C) Kaplan-Meier survival curve presenting the overall survival of PDAC patients in IHC analysis. The patients were divided equally into two groups according to the percentage of NLRP3$^+$ TANs, MME$^+$ TANs, and IFIT2$^+$ TANs among total cells in PDAC tissues. Abbreviations: TAN, tumor-associated neutrophils.

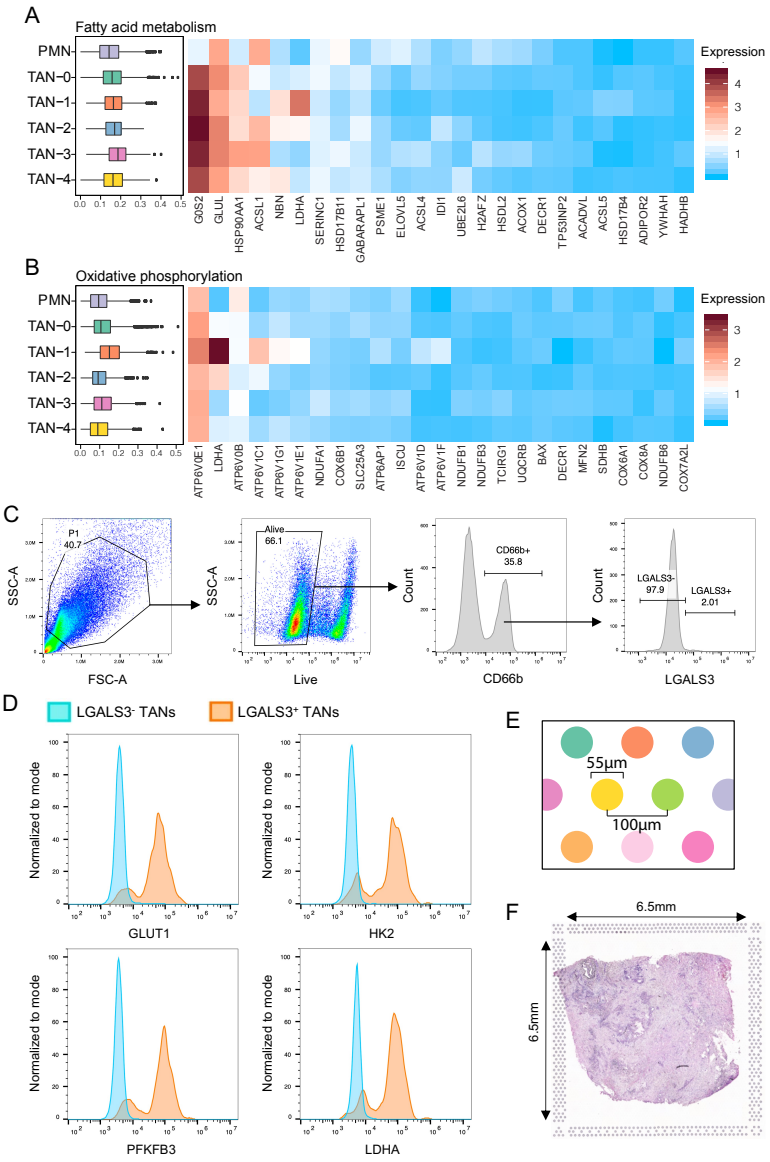

Figure S7. Analysis of metabolic features of TAN subclusters, related to Figure 4

(A-B) Boxplots of expression of hallmark fatty acid metabolism (A) and oxidative phosphorylation (B) signature in each neutrophil subcluster (left panel), and heatmap of average normalized expression of genes in signature (right panel). (C) Gating strategy of flow cytometry analysis of the expression of GLUT1, HK2, PFKFB3 and LDHA in LGALS3^{+} and LGALS3^{-} TANs from PDAC tissues in Figure 4C. (D) Flow cytometry analysis of the expression of GLUT1, HK2, PFKFB3 and LDHA in LGALS3^{+} and LGALS3^{-} TANs from one representative PDAC tissue. (E) Graphical scheme describing the spot size and resolution of spatial transcriptomics. (F) Hematoxylin & Eosin stained PDAC tissue on spatial transcriptomics array. Abbreviations: PMN, polymorphonuclear leukocytes; TAN, tumor-associated neutrophils; FSC, forward scatter; SSC, side scatter.

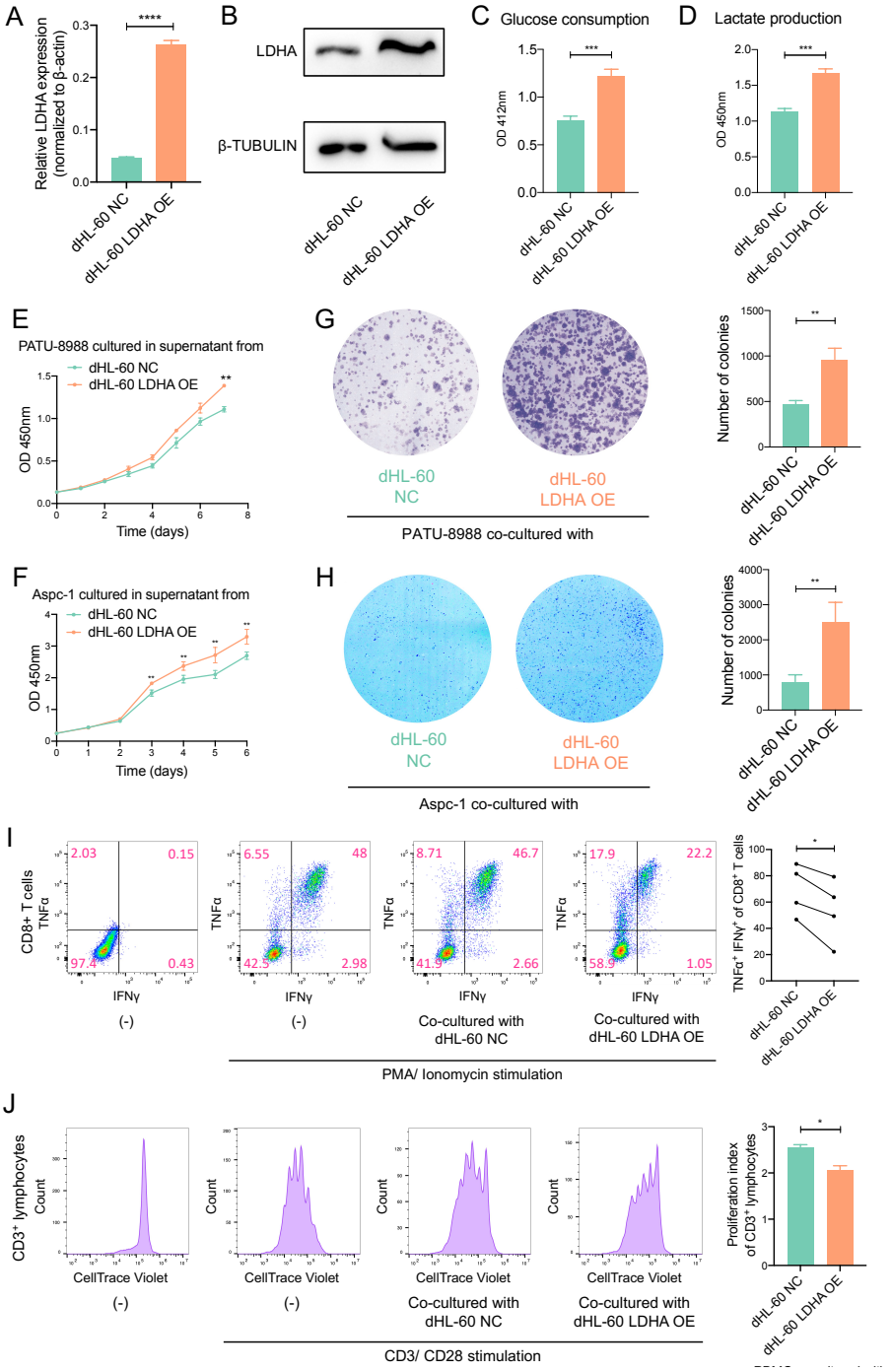

Figure S8. Glycolytic switch enhances pro-tumor functions in TANs

(A-B) Validation of LDHA overexpression in neutrophil-like differentiated HL-60 (dHL-60) cells by qPCR and Western blot. qPCR data were normalized to fold over β-actin (housekeeping gene), and represented as

mean with standard deviation. (C) Bar plot showing glucose consumption in control and LDHA overexpressed dHL-60 cells, assessed by glucose uptake assay kit. Data were represented as mean with standard deviation. (D) Bar plot showing lactate production in control and LDHA overexpressed dHL-60 cells, assessed by L-lactate assay kit. Data were represented as mean with standard deviation. (E-F) PATU-8988 cells (E) and Aspc-1 cells (F) were cultured in the supernatant derived from control and LDHA overexpressed dHL-60 cells, and the proliferation of PATU-8988/ Aspc-1 cells along time course was evaluated by CCK8 assay. (G-H) Colony formation assay, in which control and LDHA overexpressed dHL-60 cells were cultured in upper chambers, and PATU-8988 cells (G) or Aspc-1 cells (H) were cultured in lower chambers. The numbers of colonies were represented as mean with standard deviation. (I) Peripheral blood mononuclear cells (PBMCs) were co-cultured with control and LDHA overexpressed dHL-60 cells for three days, stimulated with PMA/ionomycin, and the percentage of $IFN\gamma^+$ and $TNF\alpha^+$ cells among stimulated $CD8^+$ T cells was analyzed by flow cytometry. Data from 4 different donors were summarized on the right panel. (J) PBMCs were stained with CellTrace Violet, stimulated with anti-human CD3/CD28 for four days in the absence or presence of dHL-60 cells (control or LDHA-overexpressed), and the proliferation of $CD3^+$ lymphocytes was analyzed by flow cytometry. The proliferation index of triplicate cultures were represented as mean with standard deviation. Abbreviations: dHL-60, neutrophil-like differentiated HL-60; NC, negative control; OE, overexpression. * ($p<0.05$); ** ($p<0.01$); *** ($p<0.001$); **** ($p<0.0001$)

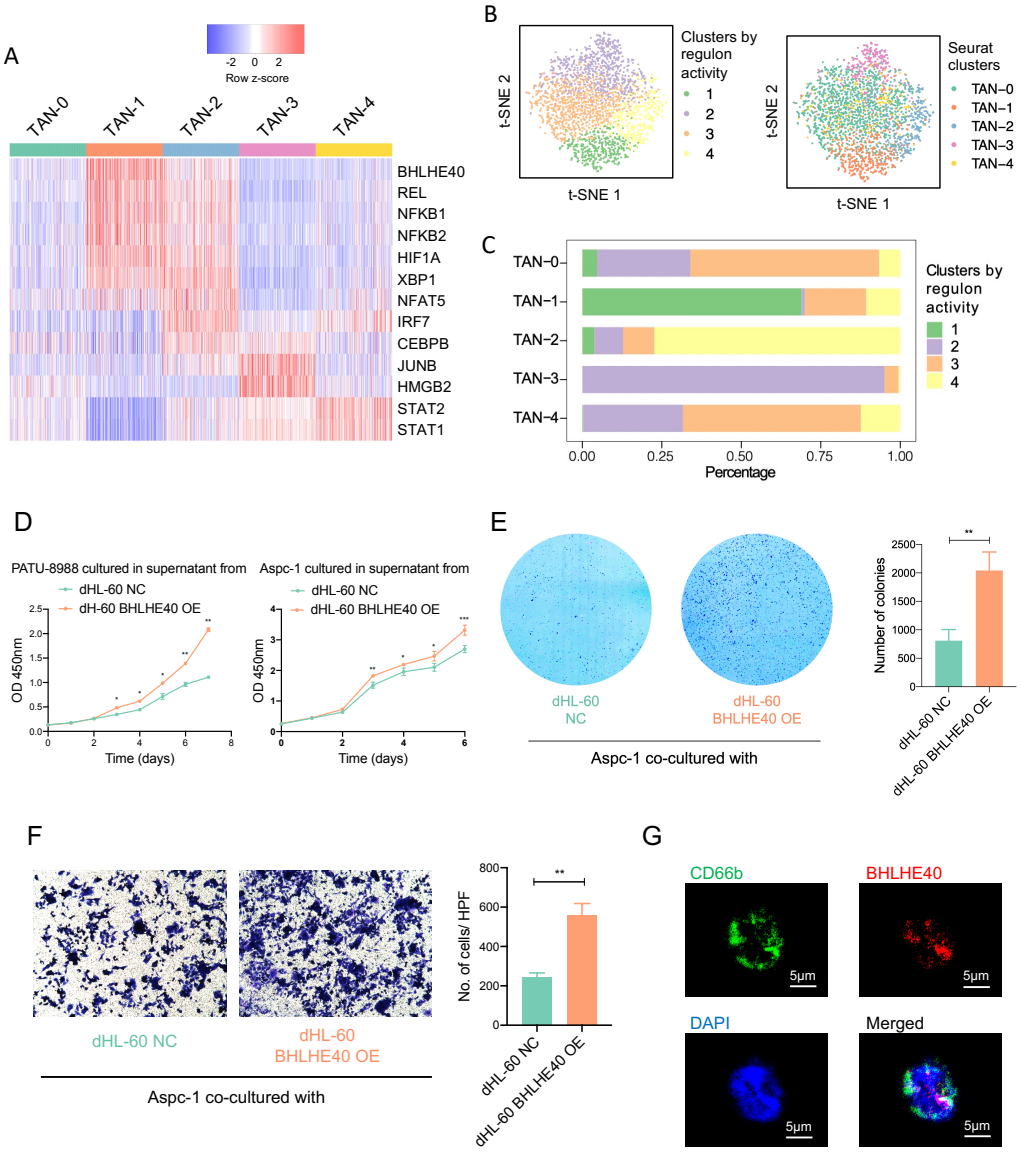

Figure S9. BHLHE40 is a major driver for the pro-tumor phenotype of TAN-1, related to Figures 5-6

(A) Heatmap depicting area under the curve (AUC) scores of representative transcription factors in each TAN subcluster, analyzed using SCENIC. (B) Unsupervised clustering of TANs based on regulon activities, visualized with t-SNE plots, with each cell color-coded for clusters by regulon activity (left panel), and the

original TAN subclusters based on RNA profiles (right panel). (C) Bar plot showing the association between clusters by regulon activity and the original TAN subclusters based on RNA profiles. (D) PATU-8988 cells (left panel) and Aspc-1 cells (right panel) were cultured in the supernatant derived from control and BHLHE40 overexpressed dHL-60 cells, and the proliferation of PATU-8988/ Aspc-1 cells along time course was evaluated by CCK8 assay. (E) Colony formation assay, in which control and BHLHE40 overexpressed dHL-60 cells were cultured in upper chambers, and Aspc-1 cells were cultured in lower chambers. The numbers of colonies were represented as mean with standard deviation. (F) Aspc-1 cells were co-cultured with control and BHLHE40 overexpressed dHL-60 cells for three days, and their migration capacity was assessed by transwell assay. The numbers of migrated cells were represented as mean with standard deviation. (G) Immunofluorescence (IF) staining of CD66b and BHLHE40 on PDAC tissue. Abbreviations: TAN, tumor-associated neutrophils; t-SNE, t-distributed stochastic neighbor embedding; dHL-60, neutrophil-like differentiated HL-60; NC, negative control; OE, overexpression; HPF, high power field. * ($p<0.05$); ** ($p<0.01$); *** ($p<0.001$)

Supplementary methods

Human subjects

For the primary cohort, we enrolled 2 healthy donors, 2 patients with treatment-naïve chronic pancreatitis (CP), and 5 patients with PDAC. Additional 138 PDAC patients were involved in the validation cohort (24 patients in validation cohort 1, and 114 patients in validation cohort 2). All the PDAC patients had untreated, resectable, non-metastatic pancreatic tumors that were confirmed to be PDAC according to pathologist assessment, and the patients with infectious diseases, rheumatic diseases, or other malignancies were excluded. In addition, blood samples from 6 healthy donors were collected for *in vitro* immune functional assays. The samples were obtained from Ruijin Hospital, Shanghai Jiaotong University School of Medicine. The study protocol was approved by the Ethics Committee of Ruijin Hospital affiliated to School of Medicine, Shanghai Jiaotong University. All enrolled participants consented to attend this cohort study and signed written informed consent. It was not appropriate or possible to involve patients or the public in the design, or conduct, or reporting, or dissemination plans of our research

Sample collection, processing, and purity assessment

4ml peripheral blood was collected using EDTA anticoagulant tubes prior to surgery, and processed within 15 minutes after collection. Red blood cells were depleted by gravity sedimentation through HetaSep (StemCell). Peripheral blood leukocytes were harvested and labeled with CD66b-phycoerythrin (PE) monoclonal antibody (BioLegend, 305106), and then neutrophils were isolated by anti-PE beads and MACS column (Miltenyi), according to manufacturer's protocol.

For PDAC tumor tissues, following resection in the operating room, small tissue blocks from ≥3 different sites of tumor were collected and immersed in tissue storage solution (Miltenyi), and transported to the research facility on ice immediately. On arrival, tissues were rinsed with PBS, and necrotic foci, hemorrhagic foci, and blood vessels were removed. Tissues were then minced in into small pieces (<1mm in diameter), and then transferred into 2ml digestion medium containing 100ul enzyme H, 50ul enzyme R, and 12.5ul enzyme A (Tumor Dissociation Kit, Miltenyi, 130-095-929) in RPMI-1640. Tissues were enzymatically digested on a shaker at 37℃ for 20min. The dissociated cells were collected at the interval of 10min to increase viability, and filtered through a 40μm nylon cell strainer to harvest single-cell suspension. Dead cells were removed by dead cell removal kit (Miltenyi). For the primary cohort, $CD45^+$ immune cells were isolated using CD45 microbeads (Miltenyi), according to manufacturer's protocol. For validation cohort 1, $CD66b^+$ neutrophils were isolated using the same protocol as that for peripheral blood. Throughout the dissociation and isolation procedure, cells were maintained on ice whenever possible.

For the neutrophils isolated from peripheral blood or tumor tissues using CD66b-PE monoclonal antibody and anti-PE beads, the purity was assessed by flow cytometry performed with Beckman CytoFlex S and analyzed using FlowJo software.

Single-cell RNA-sequencing

Single-cell RNA sequencing was performed on the single-cell suspensions with viability >70%, including $CD66b^+$ PMNs from peripheral blood and $CD45^+$ immune filtrates from tumor tissues. Single-cell capture was achieved by BD Rhapsody system via a limited dilution approach, in which ~10,000 cells were randomly distributed across >200,000 microwells. Beads with oligonucleotide barcodes were added to microwells so that each bead was paired with a cell. Cells were lysed in microwells, and poly-adenylated RNA molecules were hybridized to the beads. All the beads were collected from microwells into a single tube. Reverse transcription was performed, and each cDNA molecule was tagged on the 5′ end with a unique molecular identifier (UMI) and a cell label barcode. Whole transcriptome libraries were prepared according to the BD Rhapsody single-cell whole-transcriptome amplification workflow. In brief, second strand cDNA was synthesized, and WTA adaptor was ligated for universal amplification. The adaptor-ligated cDNA products were amplified by 18 cycles of PCR. Random priming PCR of the whole-

transcriptome amplification products was performed to enrich the 3′ end of the transcripts linked with the UMI and barcode. Sequencing libraries were then quantified by a High Sensitivity DNA chip (Agilent) on a Bioanalyzer 2200 and the Qubit High Sensitivity DNA assay (Thermo Fisher Scientific), and sequenced using HiSeq Xten (Illumina, San Diego, CA) on 150 bp paired-end run.

Single-cell RNA-seq data processing and quality control

Raw data was processed using fastp with default parameter to filter adaptor sequences and remove low-quality reads [1]. Cell barcode whitelist was identified by UMI-tools [2]. The UMI-based clean data was mapped to human reference genome (Ensemble version 91) using STAR algorithm [3] with customized parameter from UMI-tools standard pipeline. UMI count matrices were generated for each sample, and imported into Seurat R toolkit (version 3.2.3) [4]. Low quality cells (<200 genes/cell or >20% mitochondrial genes) were excluded.

Unsupervised clustering, marker identification, and cell type annotation

Seurat package was applied for normalization and scaling of the expression matrix, using default settings. Mitochondrial contamination was regressed out by setting "vars.to.regress" parameter. In the process of subclustering PMNs and TANs, batch effect was removed with ComBat [5]. To reduce the dimensionality of the expression matrix, principal component analysis (PCA) was performed based on 2,000 highly variable genes. Unsupervised cell clusters were acquired by graph-based clustering approach (The top 30 PCs were selected, resolution = 0.5), and visualized by UMAP dimensionality reduction. The clusters were annotated to known biological cell types according to the expression of canonical markers. Marker genes of each cluster were identified by FindAllMarkers function under the following criteria: logFC > 0.25; min.pct > 0.25; adjusted P value < 0.05.

Comparison of single-cell RNA profiles between PMNs from healthy controls, CP patients and PDAC patients, and TANs from PDAC tumor tissues

The averaged expression profiles of PMNs and TANs from each sample was visualized by Barnes-Hut implementation of t-distributed stochastic neighbor embedding (t-SNE) using Rtsne package [6] (1000 iterations, perplexity = 4, trade-off $\theta = 0.5$). To compare the expression profiles of neutrophils, differentially expressed genes were identified using ''FindMarkers'' function (logFC > 1; min.pct > 0; adjusted P value < 0.05) implemented in the Seurat package, and pathway analysis was performed by gene set enrichment analysis (GSEA) [7].

Definition of cell signatures

To investigate the functions of each neutrophil subcluster, we selected the gene sets associated with neutrophil functions (Table S10), and the expression levels of each gene set were estimated for each individual cell with gene set variation analysis (GSVA) [8] implemented in the GSVA package. Linear models analyzing difference between neutrophils from one cluster and neutrophils from all other clusters were generated for comparison of pathway activities between TAN subclusters. The metabolism signature of each cell was defined as the average normalized expression of the genes involved in hallmark metabolic pathways.

Constructing single-cell trajectories of neutrophils

Pseudotime transitional trajectory of neutrophils was constructed utilizing Monocle2 package (version 2.18.0) with default parameters [9]. Top 2,000 highly variable genes in neutrophils were selected as input, and dimensionality reduction was performed by "DDRTree" method. Differentially expressed genes along the pseudotime trajectory were identified by the ''differentialGeneTest'' function with a q-value < 10^{-20}, and visualized by "plot_pseudotime_heatmap" function. Pathway enrichment analysis was performed for the differentially expressed genes with Database for Annotation, Visualization and Integrated Discovery (DAVID).

Cell-cell communication analysis

CellPhoneDB (http://www.cellphonedb.org/) was applied to analyze cell-cell communication between different cell types in tumor microenvironment [10], based on the normalized expression matrix from Seurat. Mean expression of each receptor-ligand pair was calculated as the mean of the average expression of receptor in one cluster and the average expression of ligand in the other cluster, and the p-value indicates the cell type-specificity of the crosstalk.

SCENIC analysis

To construct gene regulatory networks that predicts the association between transcription factors and target genes in TANs, SCENIC analysis [11] was run with default settings, using the 20-thousand motifs database for RcisTarget and GRNboost. The activities of regulons (transcription factors and target genes) were calculated with AUCell. Unsupervised cluster analysis of regulon activities was performed by graph-based clustering approach (The top 20 PCs were selected, resolution = 0.5), and visualized by t-SNE dimensionality reduction.

Bulk RNA sequencing and data analysis

Isolated PMNs and TANs were lysed with TRIzol (Invitrogen), and total RNA was extracted with RNeasy Mini Kit (QIAGEN, Germany), according to the manufacturer's protocol. RNA sequencing libraries were generated with KAPA RNA library preparation kit (Illumina) and KAPA dual-indexed adapter kit (Illumina), and sequenced on Illumina HiSeq X Ten platform (2 x 150 bp). The raw RNA-Seq reads were aligned to the human reference genome hg38 using STAR (v2.7.0d) [3]. The human reference was downloaded from GENCODE (https://www.gencodegenes.org/). Name-sorted and indexed BAM files were generated by Samtools (v1.8-47) [12]. To quantify gene expression levels, Salmon (v0.13.1) [13] was used to determine transcripts per kilobase million (TPM) value. Transcript sequences (v34) and the annotation file were downloaded from GENCODE. We used fragments per kilobase million (FPKM) to evaluate expression levels of individual genes. The R package limma [14] was used to calculate differentially expressed genes between PMNs and TANs, and pathway analysis was performed by gene set enrichment analysis (GSEA) [7].

Proteomic assay

Protein extraction and digestion

$1x10^6$ CD66b+ cell pellets were resuspended in 150 uL 1% SDS lysis solution (Beyotime, P0013G) for 15 min on ice, vortexed at the highest speed for 10 seconds, then incubated for an additional 15 min, followed by centrifugation for 10 min at 13500g and 4°C. The supernatant was collected and detected protein concentration by BCA assay according to the manufacturer's instructions (ThermoFisher, 23225). Extracted proteins were reduced in 10 mM dithiothreitol (ThermoFisher, 20290) at 56°C for 30 min, then alkylated in 15 mM iodoacetamide (BBI Life Sciences, A600539) at room temperature for 30 min in darkness. The protein was precipitated with acetone at -20°C for 4 hours. Protein samples was digested by trypsin at 37°C for 16 hours. Peptide mixture was desalted by SepPak C18 cartridge (Waters, 186002319) and vacuum-dried by SpeedVac (ThermoFisher).

Nano-LC-MS/MS

EASY-nLC 1200 coupled to Orbitrap QExactive HFX (Thermo Fisher Scientific) were used for proteome analysis. iRT peptides (Biognosys AB) were spiked-in samples according to the manufacturer's instructions (Biognosys AB). Buffer A: 0.1 % Formic acid in water; buffer B: 0.1 % Formic acid in 80 % ACN. Sample peptides were resolved in buffer A, loaded onto a PepMap C18 analytical column (2 μm particle, 250 mm × 75 μm NVFS 1200 bar, ThermoFisher Scientific, 164941) and separated at a flow rate of 300 nL/min with 120 min gradient. The mass spectrometer was operated under data-independent acquisition (DIA) mode with multiple MS1 scans interspersed with 20 MS/MS scans per scan cycle. Precursor ion signals was quantified by full MS scans with ion target value of 3 × 106 charges in the 350–1,250 m/z range, maximum injection time of 60 ms and resolution of 120,000; the MS2 scans with ion target value of 1×10^6 charges for the precursor window, and resolution of 30,000. MS/MS isolation windows were set 9 m/z -25 m/z unit.

Data analysis

All DIA raw files were processed directDIA analysis with Spectronaut software version 14.11.210528.47784 (Biognosys, Zurich, Switzerland). Database selected the human UniProt reference proteome with 20,379 sequences downloaded in May 2020. Both MS1 and MS2 data were used for peptide identification while the parameters of the quantification process were solely derived from the MS1 data. Trypsin/P proteolytic cleavage rule was used, allowing up to two miscleavages and a peptide length of 7–52 amino acids. Carbamidomethylation set as a fixed modification, acetylation of the protein N-terminus and oxidation of methionines as variable modifications. Protein intensities were normalized using the "Global Normalization". The identification was carried out using a kernel density estimator and false discovery rate (FDR) cut-off 0.01 for precursor and protein levels.

Glycolysis/gluconeogenesis metabolites assay

Sample preparation

$1x10^6$ $CD66b^+$ cell pellets were mixed with 1 ml ice-cold 50% methanol water solution in 1.5 mL of Eppendorf tube, vortexed for 30 s, incubated in liquid nitrogen and thawed on ice, followed by repetition vortex–freeze–thaw cycle until cells were completely broken. For protein precipitation, the samples was put in the −20 °C refrigerator for 1 h and then centrifuged for 20 min at 16300g and 4°C. The supernatant was collected and lyophilized, then redissolved in methanol /acetonitrile/ water solution (2:2:1). Vortexed until all the precipitates were dissolved. The solutions were centrifuged for 5 min at 16300g and 4 °C, the supernatants were transferred to the loading bottles for LC-MS/MS analysis.

Multiple reaction monitoring (MRM) analysis by HPLC-QTRAP/MS

A HPLC System (SHIMADZU LC-30A, JAPAN) coupled with triple quadrupole mass spectrometer (SCIEX QTRAP6500, USA) were used to establish a new quantitative method for 16 glycolysis/gluconeogenesis metabolites. The glycolysis/gluconeogenesis metabolite library was purchased from Sigma (Sigma, ML0013). 2 μL sample was separated on an Acquity UPLC®BEH HILIC column (2.1 × 100 mm I.D., 1.7 μm) with solvent A (20 mM of ammonium acetate in 95% H2O/ACN, pH = 9) and solvent B (20 mM of ammonium acetate in 95% ACN/H2O, pH = 9). The column temperature was kept at 40 °C. The total flow rate was 0.3 mL/min. The gradient was: t = 0.0 min, 90% B solvent; t = 0.5 min, 90% B solvent; t = 7.0 min, 10% B solvent; t =8 min, 10% B solvent; t = 9 min, 90% B solvent; t = 10.0 min, 90% B solvent. The instrument was operated in MRM mode, the parameters were listed as follows: temperature of electrospray ion source was 400 °C; capillary voltage was 4500 V (positive mode) and -5500 V (negative mode). Declustering Potential (DP), Entrance Potential (EP), Collision Energy (CE), Collision Cell Exit Potential (CXP) were optimized for each metabolite by FIA mode in both positive and negative mode.

Data analysis

LC-MRM MS data were acquired in Analyst 1.6 software (AB SCIEX, Framingham, U.S.A.), quantitative analysis was performed by MultiQuant 3.0.1 (AB SCIEX, Framingham, U.S.A.). The metabolites with RSD below 30% were used for further analysis. Nonparametric Mann–Whitney U test performed in SPSS 18 (SPSS, Chicago, IL) was applied to evaluate the statistic alterations of metabolites' level among two groups with p value set at 0.05.

Immunohistochemistry (IHC) staining of tissue microarrays

PDAC tumor tissues were obtained from surgical resection, and fixed with 4% formalin immediately. Tissue microarrays were constructed from paraffin-embedded tissue specimens (three 1mm cores of different regions from each tissue). IHC staining was performed according to following protocols: Tissue slides were de-paraffinized with xylenes and graded ethanol. Antigen retrieval was performed in 0.01M citrate buffer (pH 6.0), heated in microwave. After blocking endogenous peroxidase with 3% H2O2 solution and blocking nonspecific binding with 5% BSA, the slides were incubated with primary antibodies (as listed below), followed by HRP-conjugated secondary antibodies (Dako), and diaminobenzidine (DAB, Dako). The slides were then counterstained with hematoxylin (Dako). The antibodies used are listed as following: CD66b (BioLegend, 305102); VEGFA (Proteintech, 19003-1-AP); NLRP3

(Proteintech, 19771-1-AP); MME (Proteintech, 18008-1-AP); IFIT2 (Proteintech, 12604-1-AP); BHLHE40 (Novus Biologicals, NB100-1800).

We analyzed the percentage of CD66b (+) and TAN subcluster marker (+) double-positive cells (assessed based on serial sections) among total cells in 10 high power fields of each core, and calculated the average proportion of double-positive cells in three tissue cores from each patient. Therefore, the patients were divided equally into two groups based on the infiltration level of each TAN subcluster, and its association with clinical parameters was further investigated.

Immunofluorescence (IF) staining of PDAC tissues

PDAC tissues were embedded in optimal cutting temperature (OCT) medium and immediately frozen at -80°C. Frozen tissues were cut into 6μm thick sections with cryostat, and layered on gelatin-coated slides. The slides were blocked with 10% normal goat serum for 1h at room temperature, and then incubated with primary antibodies at 4°C overnight and washed with PBS, followed by incubation with fluorochrome-conjugated secondary antibody for 2hr at room temperature. After washing with PBS, the slides were mounted in aqueous mounting media with DAPI. The tissues slides were visualized with confocal microscopy (Zeiss LSM 900). The following antibodies were used for IF staining: CD66b (BioLegend, 305102); CD68 (Servicebio, GB113150); p65 (abcam, ab16502); c-JUN (Proteintech, 24909-1-AP); VEGFA (Proteintech, 19003-1-AP); NLRP3 (Proteintech, 19771-1-AP); MME (Proteintech, 18008-1-AP); IFIT2 (Proteintech, 12604-1-AP); BHLHE40 (Novus Biologicals, NB100-1800); goat-anti-rabbit (Servicebio, GB21303); goat-anti-mouse (Servicebio, GB22301).

Flow cytometry analysis of PDAC tissues

Single-cell suspensions from PDAC tissues were obtained as described above. For staining of cell surface markers, the cells were incubated with Zombie NIR™ fixable viability dye (BioLegend, 423105) and Fc block (BioLegend) for 15min before staining with fluorochrome-conjugated antibodies for 30 min at 4°C. For intracellular staining, the cells were fixed and permeabilized with BD Cytofix/Cytoperm™ Fixation/Permeabilization Kit according to manufacturer's protocol, incubated with primary antibody for 30min at 4°C, followed by PE-conjugated secondary antibody for 30min at 4°C. Flow cytometry analysis was performed with Beckman CytoFlex S or Fortessa X-20 flow cytometer. Flow cytometry staining was performed with the following antibodies: anti-CD45 (BD, 563792); anti-CD15 (BioLegend, 323006); anti-CD66b (BD, 564679); anti-CD66b (BioLegend, 305118); anti-LGALS3 (BioLegend, 125418); anti-GLUT1 (abcam, ab115730); anti-HK2 (abcam, ab209847); anti-PFKFB3 (abcam, ab181861); anti-LDHA (Proteintech, 19987-1-AP); and anti-rabbit (abcam, ab72465).

Analysis of public RNA-seq data

Gene expression data and corresponding clinical information of TCGA-PAAD cohort were downloaded from UCSC Xena. Gene expression profiles of the published dataset by Cao et al. were downloaded from Genomic Data Commons. The signature scores of neutrophils/ macrophages/ TAN-1 in each sample were calculated by GSVA, according to the marker genes identified in single-cell sequencing (logFC > 1 and adjusted p value < 0.05). The treatment-naïve patients with standard PDAC histology were divided equally into two groups based on the expression of TAN-1 signature, and Kaplan-Meier survival curves and log-rank tests were used to compare overall survival in different patient groups. Pearson correlation analysis was performed to assess the association between the expression of neutrophil signature and macrophage signature.

Spatial transcriptomics

PDAC tissue was embedded in OCT compound on dry ice and stored at -80°C. PDAC tissue was cut into 6μm thick sections with cryostat, and layered within the frames of capture areas on Visium Spatial slides (10X Genomics). After fixing with methanol and staining with hematoxylin and eosin (H&E), the slide was scanned and visualized using Pannoramic MIDI (3DHISTECH). Optimization of permeabilization time was performed using Visium Spatial Tissue

Optimization Slides & Reagent Kits (10X Genomics), according to manufacturer's instructions. Tissue permeabilization and spatial transcriptomic sequencing were performed with Visium Spatial Gene Expression Slides & Reagent Kits. Briefly, after permeabilization, the slide was incubated in RT Master Mix for 45 min at 53℃, followed by incubation with Second Strand Mix for 15 min at 65℃. Barcoded cDNA was purified and amplified, and then fragmented, A-tailed, ligated with adaptors and index PCR amplified. The concentration and size distribution of Visium library were analyzed with Qubit High Sensitivity DNA assay (Thermo Fisher Scientific) and High Sensitivity DNA chip on a Bioanalyzer 2200 (Agilent). The library was sequenced by Illumina sequencer on a 150 bp paired-end run.

Raw data was processed using fastp with default parameter to filter adaptor sequences and remove low-quality reads [1]. The clean data was mapped to human genome (GRCh38) using SpaceRanger v1.1.0 to obtain feature-barcode matrix. Seurat package was applied for normalization and sctransform. Neutrophil signatures of each spatial spot were calculated with ssGSEA [15] based on the marker genes of neutrophil cluster identified from single-cell RNA profiles of PDAC tissues, and the spots with top 10% neutrophil signatures (ssGSEA score >0.53) were defined as neutrophil-enriched spots. Similarly, glycolysis signatures of each spatial spot were calculated with ssGSEA [15] according to hallmark glycolysis pathway. Tumor and stromal regions were annotated according to brightfield image of H&E staining. The regions of spots in stromal area adjacent to tumor regions were defined as "adjacent-tumor area".

Cell culture

HL-60, Aspc-1, and PATU-8988 cell lines were obtained from the Cell Bank of Shanghai Institute of Biotechnology, Chinese Academy of Sciences (Shanghai, China). Aspc-1 and PATU-8988 were cultured in DMEM medium (Meilunbio) supplemented with 10% FBS (Gibco) and antibiotics (Beyotime, penicillin, 100U/ml; streptomycin, 0.1mg/ml). HL-60 cells were cultured in RPMI-1640 medium with 10% FBS and antibiotics. Neutrophil-like differentiated HL-60 (dHL-60) cells were induced by supplementing 1% DMSO in culture media of HL-60 for 6 days. For *in vitro* induction of TAN-1 phenotype, dHL-60 cells were stimulated with thapsigargin (Abcam), IL1β (PreproTech), TNFα (PeproTech), or cultured in hypoxic conditions (1% O_2) for 24h. In co-culture experiments, we used 24mm transwell plates with 0.4 μm pore polyester membrane insert (Corning), in which dHL-60 cells were cultured in upper chambers, and PDAC cell lines were cultured in lower chambers.

Transient overexpression and knockdown of genes in neutrophil-like differentiated HL-60 (dHL-60) cells

The full-length overexpression plasmids of LDHA and BHLHE40 and knockdown plasmids of BHLHE40 were commercially obtained from Gene Chem (Shanghai, China). The sequence of shRNA of BHLHE40 was as following: 5'-GCAGTGGTTCTTGAACTTA-3'. The shRNA sequence was ligated into pCLenti-U6-shRNA-CMV-Puro-WPRE vector. Empty vector was also used as negative control. The plasmids were transiently transfected into dHL-60 cells with Lipofectamine 2000 (Invitrogen), according to the manufacturer's protocol.

Real-time RT-PCR (qPCR) analysis

Total RNA was extracted with TRIzol (Invitrogen). 1 mg RNA was used as template for cDNA conversion with HiScript III RT SuperMix (+gDNA wiper) (Vazyme). Real-time RT-PCR analysis of mRNA expression was performed in triplicate with AceQ Universal SYBR qPCR Master Mix (Vazyme). The primers for target genes were listed in Table S11. Gene expression was normalized to β-actin as endogenous control.

Western blotting

Total protein was extracted using RIPA buffer (Solarbio, China) with protease inhibitor cocktail (Roche Applied Science, Switzerland), and quantitated by bicinchoninic acid (BCA) protein assay (Beyotime). Western blotting was performed as described previously [16].

Metabolism assays

For glucose uptake assay, dHL-60 cells were starved overnight, and seeded in 96-well plate (3000 cells/well). Glucose consumption was measured by glucose uptake assay kit (abcam, ab136955), according to manufacturer's protocol. For analysis of lactate production, dHL-60 cells were cultured at the density of $1x10^6$ in 24-well plate for 24h, and the concentration of lactate in culture supernatant was measured by L-lactate assay kit (abcam, ab65331).

Cell proliferation assay

Aspc-1 and PATU-8988 were seeded in 96-well plates, and cultured in the presence of supernatants from control and LDHA/BHLHE40 overexpressed dHL-60 cells. The number of viable cells were assessed at each 24h interval using cell counting kit 8 (Dojindo, Japan) according to the manufacturer's protocol. For colony formation assay, Aspc-1 and PATU-8988 cells were seeded in 6-well plates (1000 cells/well), and control and LDHA/BHLHE40 overexpressed dHL-60 cells were cultured in upper chambers in transwell plates. After 2 weeks of co-culture, colonies were stained with crystal violet (Beyotime, C0121).

Migration assay

After co-culturing with control and BHLHE40 overexpressed dHL-60 cells for 3 days, Aspc-1 and PATU-8988 cells were loaded in the upper chamber of transwell plates (8μm pore size; Corning), and incubated at 37℃ for 36~48h. Migrated cells were stained with crystal violet, and counted in 3 random high power fields.

In vitro immune functional assays

Peripheral blood mononuclear cells (PBMCs) were isolated from peripheral blood of healthy donors with standard Ficoll-Hypaque procedures. For analysis of cytokine expression, PBMCs were co-cultured with control and LDHA/BHLHE40 overexpressed dHL-60 cells for three days using transwell plates (0.4 μm pore size, Corning), and then stimulated with PMA and ionomycin, in the presence of brefeldin A and monensin (Cell Stimulation Cocktail, Thermo Fisher Scientific) for 4h. The percentage of $IFN\gamma^+$ and $TNF\alpha^+$ cells among stimulated $CD8^+$ T cells was analyzed by flow cytometry, as described above. The following antibodies were used for flow cytometry staining: anti-CD8 (BioLegend, 344748), anti-IFNγ (BioLegend, 502506), anti-TNFα (BioLegend, 502909). For lymphocyte proliferation assay, PBMCs were labelled with 5μM CellTrace Violet, seeded in 96-well round-bottom plates ($1x10^5$ cells/well), and co-cultured with control and LDHA/BHLHE40 overexpressed dHL-60 cells at 1:1 ratio. PBMCs were activated with coated anti-CD3 antibody (1μg/ml, clone OKT3, eBioscience) and soluble anti-CD28 antibody (5μg/ml, clone CD28.2, eBioscience) for 4 days. At the end of culture, cells were stained with Zombie NIR™ fixable viability dye and PE-conjugated anti-CD3 (BioLegend, 300441) as described above, and the CellTrace signals of lymphocytes was analyzed with flow cytometry. The proliferation index of lymphocytes was calculated using ModFit software.

Chromatin immunoprecipitation (ChIP)

For each sample, 100 million crosslinked cells were lysed to prepare nuclear extracts. After chromatin shearing by sonication for 10 min, lysates were incubated on a sky wheel overnight at 4 °C with protein A Dynabeads (Invitrogen, USA) coupled with 3-6 μg of BHLHE40 antibody (Novus Biologicals, NB100-1800). After immunoprecipitation, the beads were washed and recovered using magnets. DNA was reverse-cross-linked at 55 °C for 4 hours, followed by purification of the DNA product with a QIAquick PCR Purification Kit (QIAGEN, USA). DNA was quantitated using the Qubit dsDNA HS assay and a Qubit 2.0 Fluorimeter (Invitrogen). ChIP products were used to amplify the PCR products for 40 cycles using the following PCR conditions: 95 °C for 3 min and thermocycling at 95 °C for 15 s, 60 °C for 60 s, followed by 4 °C forever. The ChIP DNA Library was sequenced with Illumina HiSeq X Ten using the paired-end module and with 150 bp reads on each end (Novogene Biotech, China).

Statistical analysis

The association between two variables was assessed by Pearson correlation analysis. Kaplan-Meier survival curves and log-rank tests were used to compare overall survival in different patient groups. Linear models were generated for

comparison of pathway activities between TAN subclusters. Comparisons between two groups were performed with Student's t test (for normally distributed data) or Mann-Whitney U test (for non-normally distributed data). Benjamini & Hochberg correction was performed for multiple comparison.

References:

1 Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 2018;**34**:i884-i90.

2 Smith T, Heger A, Sudbery I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. Genome Res 2017;**27**:491-9.

3 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S*, et al.* STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;**29**:15-21.

4 Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, 3rd*, et al.* Comprehensive Integration of Single-Cell Data. Cell 2019;**177**:1888-902.e21.

5 Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 2012;**28**:882-3.

6 Krijthe J. Rtsne: T-Distributed Stochastic Neighbor Embedding using Barnes-Hut Implementation. 2015.

7 Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA*, et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;**102**:15545-50.

8 Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 2013;**14**:7.

9 Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M*, et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol 2014;**32**:381-6.

10 Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat Protoc 2020;**15**:1484-506.

11 Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G*, et al.* SCENIC: single-cell regulatory network inference and clustering. Nat Methods 2017;**14**:1083-6.

12 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N*, et al.* The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;**25**:2078-9.

13 Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 2017;**14**:417-9.

14 Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W*, et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;**43**:e47.

15 Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF*, et al.* Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature 2009;**462**:108-12.

16 Wu J, Li H, Shi M, Zhu Y, Ma Y, Zhong Y*, et al.* TET1-mediated DNA hydroxymethylation activates inhibitors of the Wnt/β-catenin signaling pathway to suppress EMT in pancreatic tumor cells. J Exp Clin Cancer Res 2019;**38**:348.