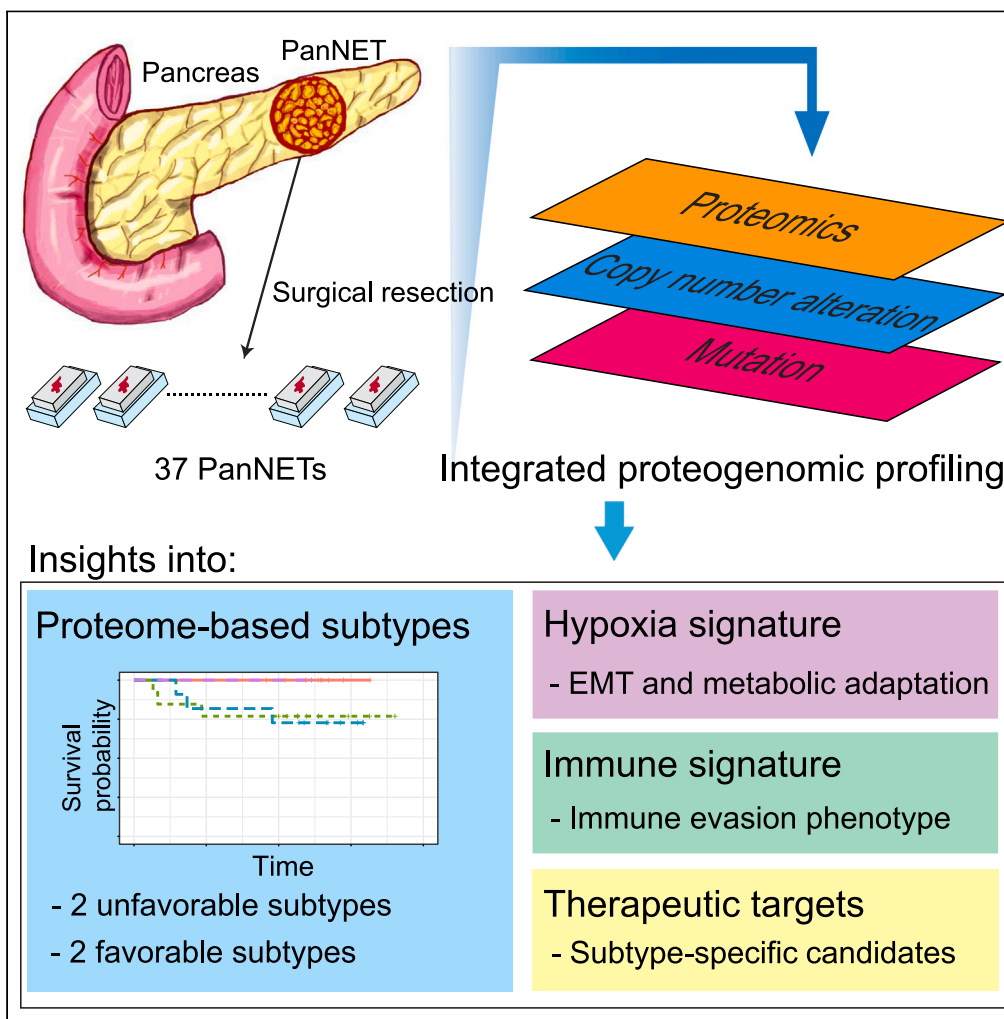Article

# Proteogenomic characterization of pancreatic neuroendocrine tumors uncovers hypoxia and immune signatures in clinically aggressive subtypes



Atsushi Tanaka,
Makiko Ogawa,
Yihua Zhou, ...,
David S. Klimstra,
Julia Y. Wang,
Michael H. Roehrl

michael_roehrl@bidmc.harvard.edu

## Highlights

Proteogenomic characterization of PanNETs identified clinically aggressive subtypes

Aggressive subtypes showed hypoxia signature enrichment with metabolic reprograming

Aggressive subtypes harbored immune suppressive tumor microenvironment

Proteomic analysis identified therapeutic targets specific to PanNET subtypes

# iScience

## Article

# Proteogenomic characterization of pancreatic neuroendocrine tumors uncovers hypoxia and immune signatures in clinically aggressive subtypes

Atsushi Tanaka,[1,2,3] Makiko Ogawa,[1,2,3] Yihua Zhou,[3,4] Yusuke Otani,[1,2] Ronald C. Hendrickson,[5,7] Matthew M. Miele,[5,7] Zhuoning Li,[5] David S. Klimstra,[3,8] Julia Y. Wang,[6] and Michael H. Roehrl[1,2,3,9,*]

## SUMMARY

**Pancreatic neuroendocrine tumors (PanNETs) represent well-differentiated endocrine neoplasms with variable clinical outcomes. Predicting patient outcomes using the current tumor grading system is challenging. In addition, traditional systemic treatment options for PanNETs, such as somatostatin analogs or cytotoxic chemotherapies, are very limited. To address these issues, we characterized PanNETs using integrated proteogenomics and identified four subtypes. Two proteomic subtypes showed high recurrence rates, suggesting clinical aggressiveness that was missed by current classification. Hypoxia and inflammatory pathways were significantly enriched in the clinically aggressive subtypes. Detailed analyses revealed metabolic adaptation via glycolysis upregulation and oxidative phosphorylation downregulation under hypoxic conditions. Inflammatory signature analysis revealed that immunosuppressive molecules were enriched in immune hot tumors and might be immunotherapy targets. In this study, we characterized clinically aggressive proteomic subtypes of well-differentiated PanNETs and identified candidate therapeutic targets.**

## INTRODUCTION

Well-differentiated pancreatic neuroendocrine tumors (PanNETs) are neoplasms originating from endocrine cells of the pancreas and are characterized by variable metastatic propensity, clinical course, and survival. The annual incidence of these tumors in the United States has increased from 1.09 to 6.98 per 100,000 individuals in recent years.[1–3] The latest WHO histological classification of well-differentiated PanNETs comprises three histologic grades (grades 1, 2, and 3) based on the tumor proliferation rate as assessed by Ki67 immunohistochemistry or mitotic count.[4] As this classification scheme has proven prognostic utility,[5] and clinical treatment decisions are currently based on this classification and Tumor-Node-Metastasis (TNM) stage. However, providing appropriate clinical management remains challenging due to disease heterogeneity, especially in low-grade PanNETs (grades 1/2), and a better patient risk stratification scheme is needed for precise clinical management of patients.

In recent years, significant advances have been made in the genomic characterization of PanNETs. Genome sequencing studies have revealed that *MEN1*, *DAXX*, and *ATRX* are among the most frequently altered genes in PanNETs and found that poor outcomes correlate with these alterations.[6–8] The mutational landscape also showed that mTOR pathway genes are frequently altered in PanNETs.[7,8] Previous transcriptome analyses have identified several transcriptomic subtypes with clinical outcome differences.[7,9–11] These reported subtypes harbored characteristic features such as epithelial-mesenchymal transition (EMT) or proliferation. A very limited number of proteomic analyses of PanNETs have been performed.[10,12,13] One of these studies describes transcriptomic subtypes of 35 PanNETs and correlated transcriptomic subtype features at proteome level.[10] However, unsupervised clustering of PanNETs based on mass spectrometry-based proteome, not transcriptome, and detailed systematic characterization of PanNETs have not been accomplished.

Considering that (i) proteins ultimately drive the biochemistry of disease and that (ii) mRNA abundance alone is a poor predictor of actual protein amounts in general,[14,15] a comprehensive proteomic analysis may be a better strategy to study biological processes of tumor

[1]Department of Pathology, Beth Israel Deaconess Medical Center, Boston, MA, USA
[2]Harvard Medical School, Boston, MA, USA
[3]Department of Pathology and Laboratory Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA
[4]ICU Department, Second Affiliated Hospital of Nanchang University, Nanchang, Jiangxi, China
[5]Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA
[6]Curandis, New York, NY, USA
[7]Present address: Sylvester Comprehensive Cancer Center, University of Miami, Florida, USA
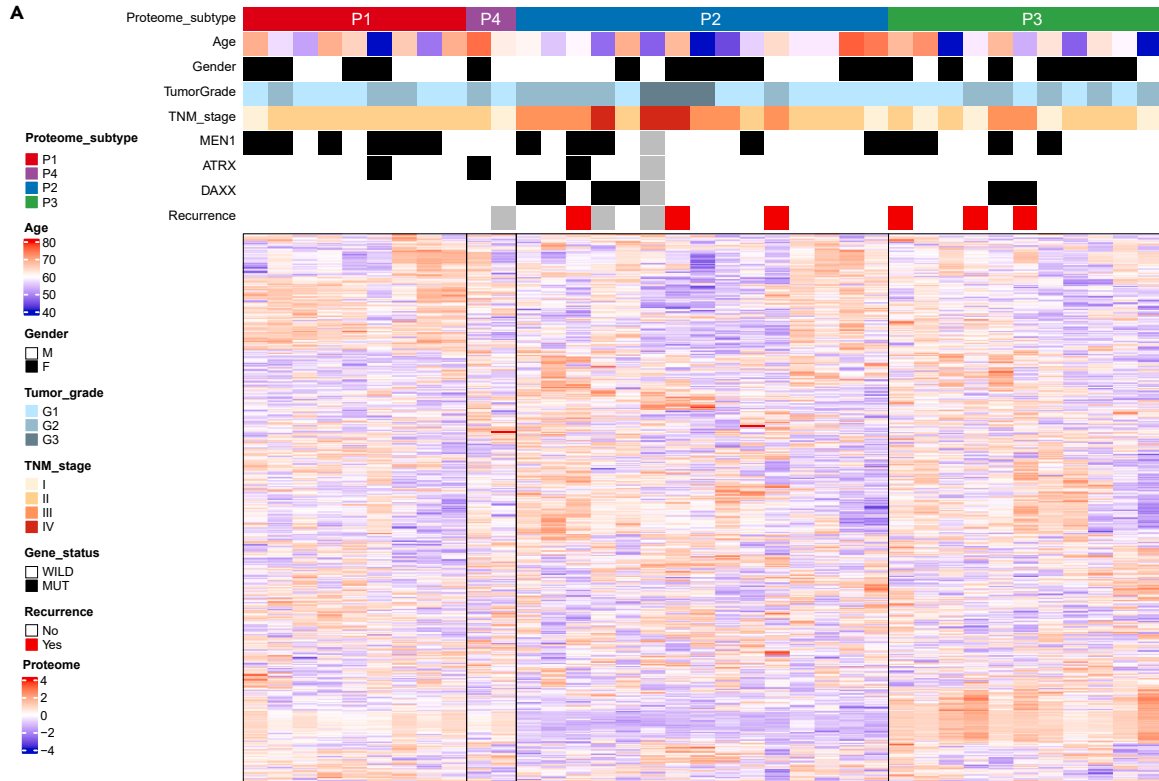[8]Present address: Paige.AI, New York, NY, USA
[9]Lead contact
*Correspondence: michael_roehrl@bidmc.harvard.edu
https://doi.org/10.1016/j.isci.2024.110544

**A**



**B**

| Proteome subtypes | TNM stage | |
|---|---|---|
| | Stages I&II | Stages III&IV |
| P1/4 | 11 | 0 |
| P2/3 | 15 | 11 |

Fisher's exat test p=0.0153
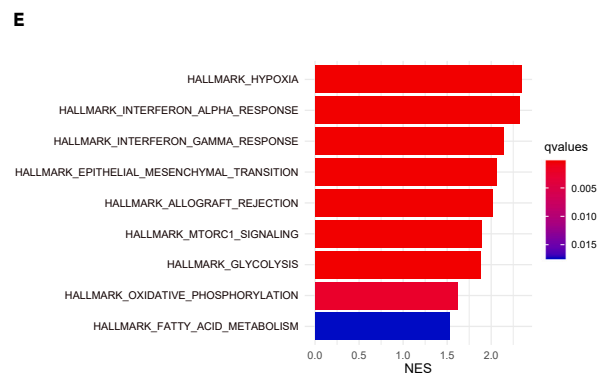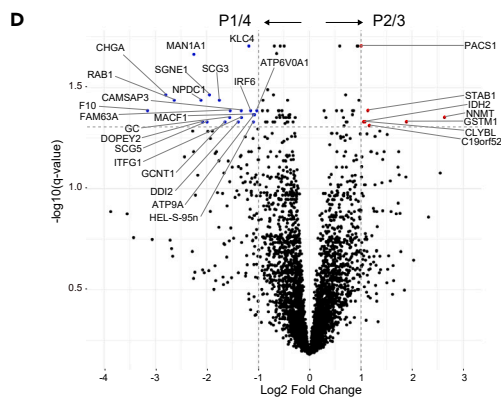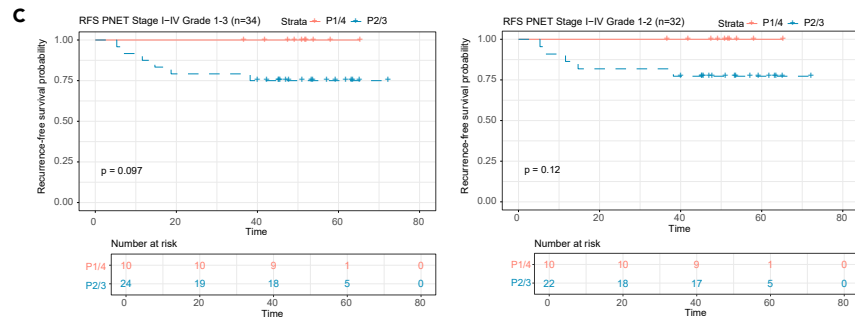
**C**



**D**



**E**

**Figure 1. Proteogenomic profiling with clinical attributes**

(A) Summary plot of the 37 PanNETs with proteogenomic features and clinical attributes. The top 500 proteins with the highest variance are shown in a heatmap. Gray columns indicate the missing data. Note: recurrence events are only observed in the P2/3 subtypes.

(B) Stage III/IV samples are significantly enriched in the P2/3 subtypes.

(C) Kaplan-Meier curve analyses of recurrence-free survival (RFS) time in present study. P2/3 subtypes show shorter RFS time than P1/4 subtypes, however, the difference is not statistically significant. Survival analysis plot with all TNM stages and all tumor grades is shown in left-side. Survival analysis plot with all TNM stages and tumor grades 1 or 2 is shown in right-side.

(D) Volcano plot of the differential expression analysis results between P2/3 and P1/4 subtypes. Gene names with statistical significance (q < 0.05) and two or more-fold changes are shown. The horizontal dashed line corresponds to q = 0.05.

(E) Pathway enrichment analysis with Hallmark gene set signatures between the clinically aggressive and nonaggressive proteome-subtypes (P2/3 vs. P1/4). All the significant results are shown.

aggressiveness. Therefore, we performed integrated proteogenomic analyses of PanNETs to define new proteomic subtypes and features that determine disease aggressiveness and clinical outcome.

## RESULTS

### Unsupervised clustering identifies 4 proteome-based PanNET subtypes with outcome differences

We performed liquid chromatography-tandem mass spectrometry (LC-MS) analyses of 37 primary PanNETs vs. (as control comparison groups) 11 primary pancreatic ductal adenocarcinomas (PDACs) and 9 normal pancreatic tissues. The correlation coefficients between the pooled samples in each LC-MS run were >0.97 (Figure S1A), showing high reproducibility and robustness of our proteomic analyses. We quantified 7,447 proteins in the total cohort, with 4,249 proteins quantified in at least 50% of samples of each histologic group (i.e., PanNETs, PDACs, or normal pancreatic tissues). Unsupervised uniform manifold approximation and projection (UMAP) of the proteome data showed a clear separation between PanNETs, PDACs, and normal pancreatic tissues (Figure S1B).

Among the PanNET samples, the well-differentiated PanNET G3 tumor samples formed a distinct proteomic cluster, whereas the PanNET G1/G2 samples showed no distinct proteomic cluster separation (Figure S1B). To identify proteome-based subtypes of PanNETs, we performed unsupervised clustering[16] and identified four subtypes, P1-P4 (Figures 1A and S1C–S1E). Among these subtypes, PanNET G3 was only observed in P2. Advanced patient cases (TNM stages III or IV) were significantly enriched in P2/3 (11 out of 26) compared to P1/4 (0 out of 11) (Fisher's exact test, $p$ = 0.0153) (Figure 1B). While *MEN1* and *ATRX* mutations did not show clear enrichment in proteomic subtypes, *DAXX* mutations were enriched in P2/P3 (6 out of 26, 23.1%) compared to P1/P4 (0 out of 11, 0.0%). As a marker of clinical aggressiveness, development of subsequent tumor recurrence is an important factor for physicians to predict clinically aggressive behavior and to inform adjuvant, recurrence-preventing treatment. Among 37 PanNETs, 34 PanNETs had both complete surgical resection status and subsequent clinical follow-up data. Follow-up time ranged from 5.3 to 72.2 months (median, 50.0 months). Recurrence events were exclusively observed in P2/3 subtypes (P2/3: 6 out of 24, 25.0%; P1/4: 0 out of 10, 0.0%). Kaplan-Meier analysis showed P2/3 subtypes of all stages had a trend for shorter recurrence-free survival (RFS) times than P1/4 subtypes (no recurrence observed in P1/4), regardless of G3 tumor inclusion status (Figure 1C), but this difference did not reach to statistical significance in the small cohort. Similarly, early stage P2/3 PanNETs of G1/2 showed a trend for shorter RFS time than P1/4 (Figure S1F). To validate clinical outcome difference and trend of our proteome subtypes by using independent cohort from our cohort, we analyzed an independent external RNA-seq dataset of 78 well-differentiated PanNETs (EGAS00001005024) with overall survival (OS) time annotation (RFS data not available for this dataset).[10] TNM stage information was available for 69 PanNETs. Although not statistically significant, the P2/3 subtype was enriched in advanced TNM stages (Figure S1G). By Kaplan-Meier analysis, P2/3 cases had a significantly shorter OS time. Early stage P2/3 PanNETs showed the same OS trend (Figure S1H). These independent results suggest that the P2/3 proteome subtypes are associated with clinically unfavorable outcomes. Based on our data, it is possible that P2/3 may also be an indicator of advanced stage disease (TNM III/IV), but a repeated consistent trend for worse RFS and OS outcomes (although individually not statistically significant) is observed in both our proteomic and the external transcriptomic validation cohorts for early stage P2/3 tumors (Figures S1F and S1H).

To better characterize these clinically aggressive proteomic subtypes (P2/3) relative to the less aggressive subtypes (P1/4), we performed differential protein expression analyses. Comparison between the P2/3 and the P1/4 subtypes revealed 79 significant proteins, including 7 upregulated and 21 downregulated proteins with >2-fold change (Figure 1D; Table S1). Signal pathway enrichment analysis of Hallmark gene sets revealed a hypoxia signature, EMT, and an inflammatory signature enriched in clinically aggressive subtypes (P2/3) (Figure 1E), which is validated by the EGAS00001005024 dataset (Figure S1I).

### Oncogenic somatic mutations and copy number alterations enriched in proteome-subtype P2/3 PanNETs

To elucidate genomic event differences between the P1/4 and P2/3 subtypes, we performed large gene panel sequencing (MSK-IMPACT) of 36 PanNETs and plotted an oncoprint with genes that were reported as recurrent genomic events in a whole genome sequencing cohort.[7] Combined genomic analysis of somatic mutations and copy number alterations revealed that the top five most frequently altered genes in our cohort were *MEN1* (53%), *DAXX* (31%), *BRCA1* (17%), *ATRX* (14%), and *TERT* (11%), which is concordant with previous reports (Figure 2A).[6–8] These genes are mainly involved in four pathways: (i) DNA damage repair molecules, *MUTYH*, *CHEK1/2*, and *BRCA1/2*; (ii) chromatin remodeling, *MEN1* (histone modification, epigenetic gene regulation, and tumor suppressor), *SETD2* (histone methyltransferase), and *ARID1A* (SW1/SNF complex member); (iii) telomere maintenance, *TERT*, *MEN1*, *DAXX*, and *ATRX*; and (iv) regulators of PI3K/AKT/mTOR signaling, PTEN, and TSC1/2, which
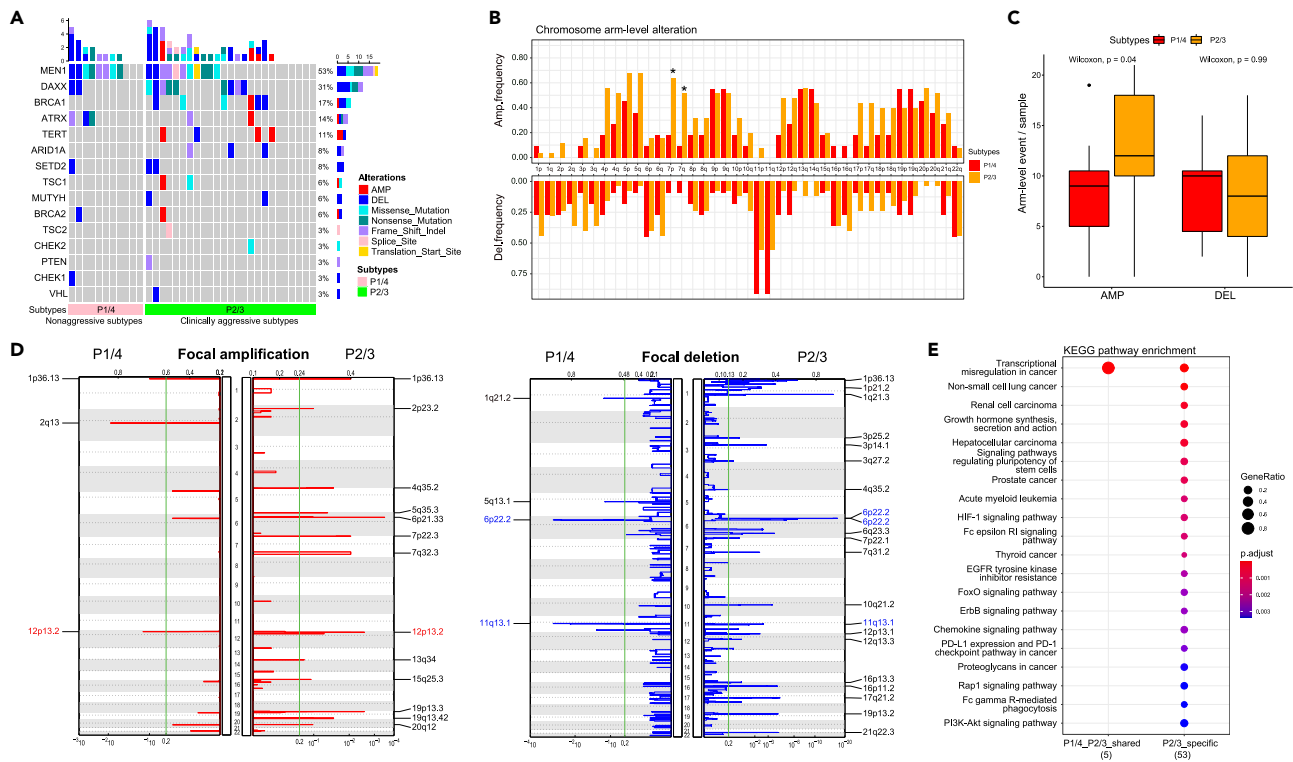
**Figure 2. Somatic recurrent copy number alteration of PanNETs**

(A) Integrated SNV and CNA oncoprint of PanNET based on MSK-IMPACT shows frequent alterations in *MEN1* and *DAXX*. Altered genes are mainly involved in DNA repair, chromatin remodeling, telomere maintenance, and PI3K/mTOR signaling.

(B) Frequency plot of chromosomal arm-level copy number alteration. * indicates statistical differences between the P1/4 and P2/3 subtypes. Only the 7p and 7q events show significant differences in frequency between P1/4 and P2/3.

(C) The arm-level event count per patient shows a significant increase in amplification events in the P2/3 subtypes. Although the arm-level event count of deletion does not show a significant difference, the arm-level event of amplification shows a significantly higher event rate in P2/3 than in P1/4.

(D) Recurrent somatic focal peak analysis result is shown. Chromosomal loci colored in red or blue are shared between the P1/4 and P2/3 subtypes.

(E) KEGG pathway enrichment analysis of cancer-related genes involved in focal peak events. Many oncogenic pathways such as HIF1 signaling are enriched in the clinically aggressive subtypes (P2/3).

are tumor suppressors. We also checked the *VHL* somatic alteration status, as it is a negative regulator of hypoxia signaling.[17] In our cohort, one *VHL* deletion was observed in the P2/3 subtypes, which is compatible with hypoxia pathway enrichment in the P2/3 subtypes. *DAXX* mutation, which is a poor outcome predictor of PanNETs,[6,7] showed a preference for P2/3 over P1/4 (although not statistically significant) (Figures 1B and 1C). In addition, even though the number of gene events was small, DNA damage repair molecules (*BRCA1*), chromatin remodeling (*ARID1A*), telomere maintenance (*TERT*, *DAXX*), and mTOR pathway gene alterations mainly occurred in the P2/3 subtypes. As these pathways are known to be related to malignant features in PanNETs,[6–8] this is compatible with the poor outcome of the P2/3 subtypes.

To study recurrent somatic copy number alterations at the chromosomal arm level and focal peak levels, we performed a GISTIC2.0 analysis.[18] Chromosome arm level amplifications in over 50% of samples affected 4p, 4q, 7p, 7q, 13q, 14q, 19p, 19q, 20p, and 20q in either the P1/4 or P2/3 subtypes (Figure 2B). Only chromosome 7p and 7q amplifications showed a significantly higher frequency in the P2/3 subtypes than in the P1/4 subtypes. Chromosome 7 includes several oncogenes, such as *CDK6* or *EZH2*, which have been reported as predictors of poor outcome and potential treatment targets in neuroendocrine tumors.[19,20] Thus, amplification of chromosome 7 supports P2/3 poor outcomes in our cohort. In contrast, arm level deletions in over 50% of the samples affected 11p and 11q (including the tumor suppressor *MEN1*). There was no significant difference in arm level deletion frequency between the P1/4 and P2/3 subtypes. Frequent loss of chromosome 11 at the arm level among all samples, including *MEN1*, may be an important genomic event in tumorigenesis. When we compared the arm level alteration count per sample between P1/4 and P2/3, arm level amplification was significantly higher in P2/3 (Figure 2C). This suggests that P2/3 could be better characterized by arm level amplification events rather than by deletion events. We then investigated recurrent focal copy number alterations. We found one shared recurrent focal amplification peak (12p13.2, colored red in Figure 2D) and two shared recurrent focal deletion peaks (6p22.22 and 11q13.1, colored blue in Figure 2D) between P1/4 and P2/3. To facilitate understanding of the pathway effect of focal copy number alterations, we performed a KEGG pathway enrichment analysis of genes involved in recurrent focal peak alterations restricted to a cancer-related COSMIC gene list comprised of 723 known oncogenes and tumor suppressors.[21] The analysis found that P2/3-specific focal peaks were enriched in genes related to oncogenic (PI3K-Akt, etc.)
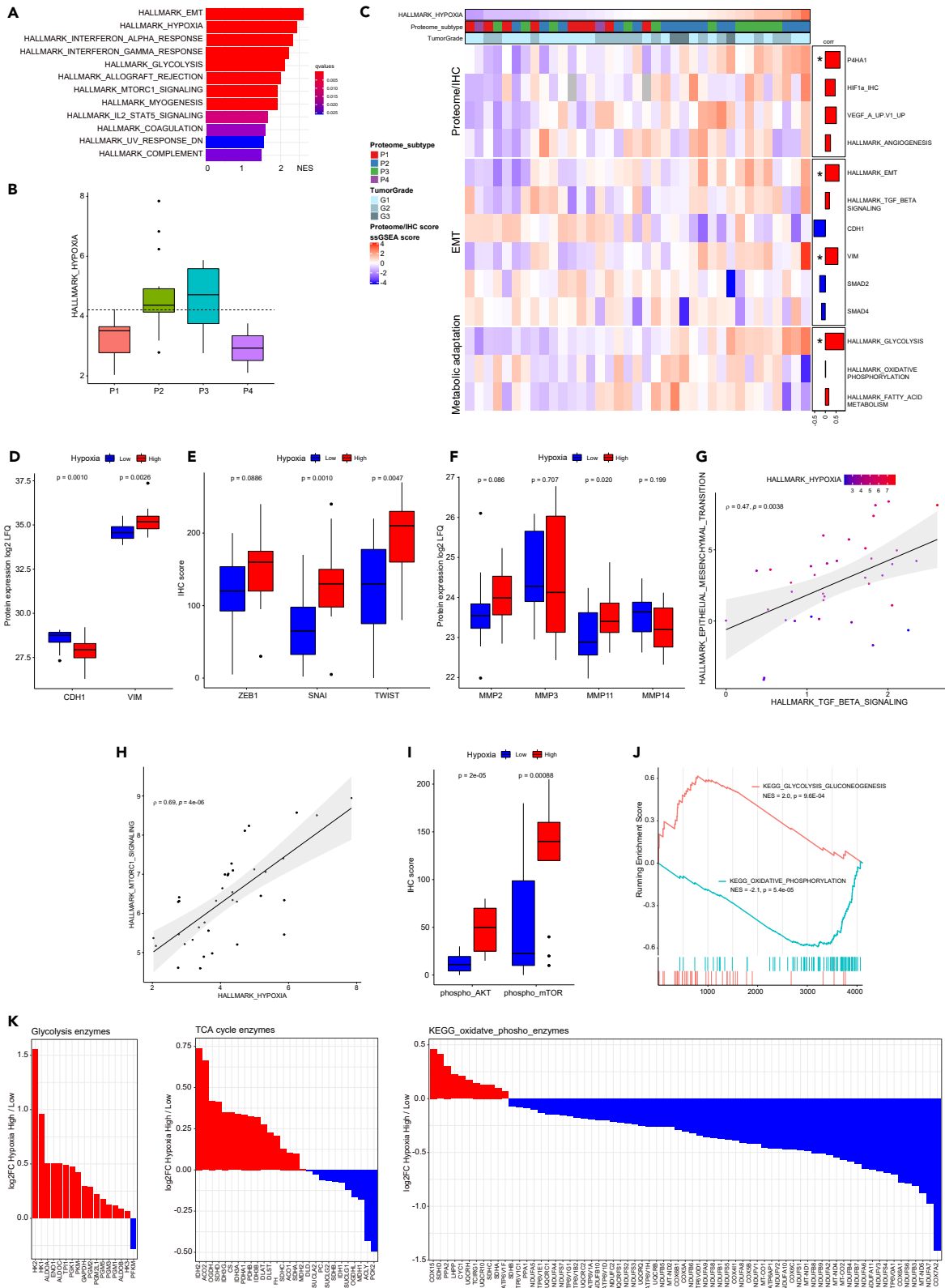
**Figure 3. Proteomic characterization of hypoxia signature in PanNETs**

(A) Pathway enrichment analysis with Hallmark signatures between hypoxia-low and hypoxia-high groups based on ssGSEA Hallmark_Hypoxia score (divided by median value). All statistically significant results are presented. As expected, pathways known to be closely related to hypoxia signature, such as EMT or glycolysis, are enriched in the hypoxia-high group.

(B) Boxplot of hypoxia scores between proteome-based subtypes. The P2/3 subtypes have higher hypoxia scores than the P1/4 subtypes, as expected from pathway enrichment analysis between the P2/3 and P1/4 subtypes. The dashed horizontal line represents the mean hypoxia score across this cohort.

(C) Heatmap of hypoxia-related pathway scores and protein abundances. The samples are ordered according to their hypoxia scores. Spearman coefficients with hypoxia scores are shown on the right-hand side of the heatmap. Hypoxia-related molecules and pathways, such as HIF1α or angiogenesis score, are positively correlated with hypoxia score. EMT-related signatures also show concordant correlations with hypoxia score. As representative EMT markers, E-cadherin (coded by *CDH1*) decreases, and vimentin (coded by *VIM*) increases in EMT, resulting in negative and positive correlations with hypoxia score. Furthermore, glycolysis is upregulated in hypoxia-high status, with high coefficients. * indicates statistical significance.

(D) Boxplots of E-cadherin and Vimentin protein expression in the hypoxia low and hypoxia high groups (quantified by mass spectrometry). Consistent with the hypoxic group status, E-cadherin is significantly downregulated in hypoxia-high group. Vimentin is significantly upregulated in the hypoxia-high group.

(E) Boxplots of the protein expression of EMT-inducing transcription factors (quantified by IHC). SNAI and TWIST are significantly upregulated in the hypoxia-high group.

(F) Boxplots of matrix metalloprotease protein expression in our proteome dataset (quantified by mass spectrometry). MMP11 expression is significantly upregulated in the hypoxia high group.

(G) TGF-β signaling positively correlates with the EMT signature.

(H) Hypoxia signature is positively correlated with mTORC1 signaling.

(I) Phosphorylated Akt and mTOR expression levels (quantified by IHC) are significantly higher in the hypoxia high group than in the hypoxia low group.

(J) GSEA plot of KEGG glycolysis pathway and KEGG oxidative phosphorylation pathway. KEGG, which is a gene set different from the Hallmark gene set, shows again metabolic adaptation.

(K) Fold changes of proteins in the glycolysis pathway, TCA cycle, and oxidative phosphorylation are shown (quantified by mass spectrometry).

and immunosuppressive pathways (PD-L1, etc.) (Figure 2E). KEGG pathway enrichment analysis also revealed HIF1 signaling enrichment in the P2/3 subtypes, concordant with the fact that the most highly enriched pathway of P2/3 is the hypoxia signaling pathway (Figure 1E).

### Hypoxia signature associated with epithelial-mesenchymal transition and crosstalk with TGF-β signaling

In many solid cancers, a hypoxia signature correlates with clinically aggressive phenotypes, such as invasion and metastasis.[22] To investigate hypoxic features in detail at proteome level, we calculated the Hallmark ssGSEA hypoxia score for each sample and separated the cohort into low and high hypoxia score groups, followed by signal pathway enrichment analyses (Figures 3A and 3B). This comparison revealed significant differences in pathway enrichment between hypoxia high and hypoxia low groups. As shown in Figure 3A, the most enriched pathway was the EMT signature, which is known to be related to invasion and metastasis in many solid cancers.[23] Clinically aggressive subtypes P2/3 had higher hypoxia scores than clinically less aggressive subtypes (P1/4). The hypoxia signature was significantly correlated with EMT and metabolic adaptation pathways (i.e., glycolysis pathway) (Figure 3C). As expected, HIF1α IHC score, EMT, and metabolic pathways correlated well with the Hallmark hypoxia score (Figure 3C). P4HA1 protein (a key enzyme involved in collagen synthesis and EMT) level, a stabilizer of HIF1α,[24] is also positively correlated with the hypoxia score. In colorectal cancer, high P4HA1 expression correlates with poor prognosis and facilitates metastasis in a mouse PDX model,[25,26] suggesting that P4HA1 may have a similar biological role in PanNET.

Immunohistochemical assessment confirmed that the EMT phenotype was enriched in the hypoxia high group (Figure 3D). Next, we tested whether EMT-inducing transcription factors (ZEB1, SNAI1/2, and TWIST) were upregulated in the hypoxia high group (Figure 3E). The external well-differentiated PanNET transcriptome datasets (GSE118014 [*n* = 33], EGAS00001005024 [*n* = 78]) confirmed these results (Figure S2A). As tumor cells need to break the basement membrane and intercellular matrix when they invade during EMT, we investigated the expression levels of matrix proteases (MMPs) that were quantified in our proteome dataset. Consistent with a previous report on the correlation between the EMT phenotype and MMP expression,[22,27] MMP11 was expressed at significantly higher levels in the hypoxia high PanNET group (Figure 3F). The external transcriptome datasets showed most MMPs were highly expressed in hypoxia-high group, however, mRNA MMP11 did not reach statistical significance (Figure S2B).

Next, we investigated the crosstalk between EMT signature and TGF-β signaling.[28] The EMT signature score showed a significant positive correlation with TGF-β signaling (Figure 3G). However, SMAD2 and SMAD4 expression levels were negatively correlated with the EMT signature score (Figure 3C). We then elucidated non-SMAD TGF-β signaling, that is, PI3K/AKT pathways.[29] Signal pathway analysis and immunohistochemical assessment of phosphorylated AKT and mTOR revealed that AKT-mTOR pathway activation was enriched in the hypoxia high group (Figures 3H and 3I). To validate our findings, we analyzed the 2 transcriptome datasets (GSE118014, EGAS00001005024) and found the same signatures (Figures S2C–S2E). These findings indicate that the hypoxia-induced EMT signature exhibits crosstalk with TGF-β signaling via the non-SMAD axis. These findings suggest that inhibition of TGF-β signaling may be a therapeutic option for the hypoxia high group (i.e., the P2/3 subtypes).[30]

### Metabolic reprogramming under hypoxia

Energy production in tumors is an important process, especially under hypoxic conditions, in which oxygen and nutrient supplies are very limited. To understand how PanNETs adapt metabolically to hypoxic conditions, we examined ATP-producing pathways, including glycolysis,
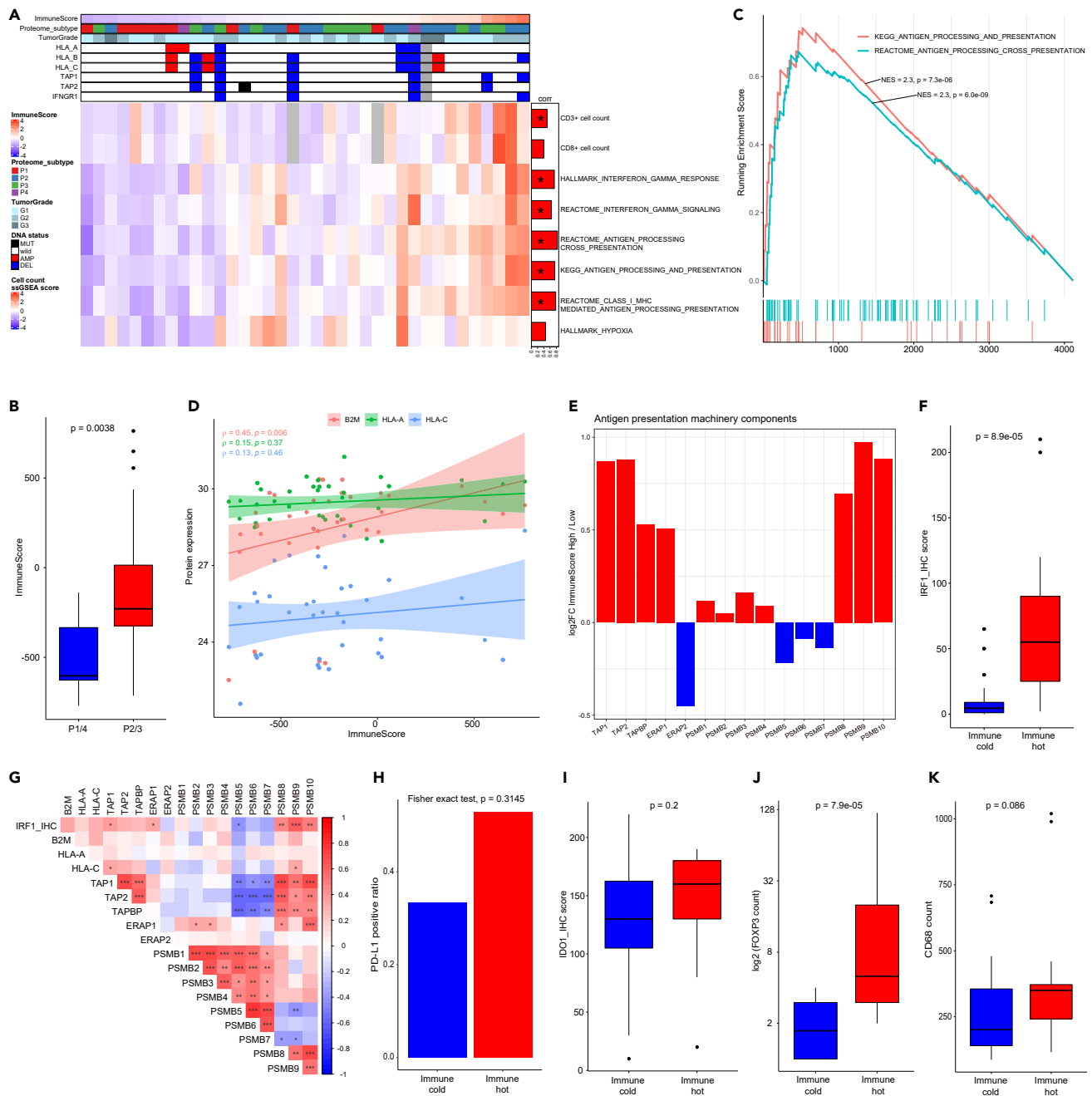
**Figure 4. Integrated immune signature analysis of PanNETs at proteome level**

(A) Immune signature profiling of DNA, proteins, and pathway alterations. Gray columns indicate the missing data. Although genomic events do not have a clear correlation with immune score, immune cell infiltration (CD3 count), INFγ signaling, and antigen presenting system are significantly correlated with immune score. * indicates statistical significance.

(B) PanNET P2/3 subtypes show significantly higher immune scores than P1/4 subtypes.

(C) GSEA between immune cold and hot groups shows significant enrichment of antigen processing-related pathways in immune hot tumors.

(D) Among MHC class I molecules on the cell surface, B2M shows a significant positive correlation with the immune score. HLA-A and HLA-C have a weak positive correlation with the immune score, but not significant. Proteins were quantified by mass spectrometry.

(E) Fold change bar chart of the proteasome components (quantified by mass spectrometry). Most components are upregulated in immune hot tumors compared to immune cold tumors.

(F) IRF1 expression (quantified by IHC) is significantly higher in immune hot tumors than immune cold tumors.

(G) Correlation plot of IRF1, MHC class I, and proteasomecomponents. *, $p < 0.05$; **, $p < 0.01$; $p < 0.001$.

**Figure 4. Continued**

(H) The PD-L1 positivity rate (quantified by IHC) is higher in immune hot tumors than in immune cold tumors, but the difference is not statistically significant.

(I) IDO1 expression (quantified by IHC) is higher in immune hot tumors than in immune cold tumors, but the difference is not statistically significant.

(J) The number of FOXP3 positive cells (quantified by IHC), which are immune suppressors, is significantly higher in immune hot tumors than in immune cold tumors.

(K) CD68 positive cells (quantified by IHC) are enriched in immune hot tumors with close statistical significance.

TCA cycle, and oxidative phosphorylation. Concordant with the signaling pathway analysis result (Figure 3A), KEGG gene set enrichment between hypoxia low and hypoxia high groups showed glycolysis upregulation in the hypoxia high group (Figure 3J). In addition, the oxidative phosphorylation pathway, which is closely connected to the TCA cycle for effective ATP production, was downregulated in the hypoxia high group (Figure 3J). To further elucidate these pathways in detail, we investigated the enzyme expression levels of glycolysis, TCA cycle, and oxidative phosphorylation pathways based on our proteome data. Clearly, glycolysis enzymes were upregulated, and enzymes involved in the oxidative phosphorylation process were downregulated (Figure 3K). Concordantly, some TCA cycle enzymes were downregulated in the hypoxia high group (Figure 3K). The external transcriptome datasets (GSE118014, EGAS00001005024) confirmed glycolysis enrichment in the hypoxia high group (Figure S2F). In general, cancer cells mainly use aerobic glycolysis and do not use the TCA cycle to produce ATP, even though the TCA cycle generates ATP more efficiently (Warburg effect).[31] Our findings suggest that PanNET dynamically shifts the ATP-producing pathway to anaerobic glycolysis under hypoxic conditions, resulting in better adaptation to the tumor microenvironment.

## Immune hot tumors express immune evasion molecules

Since the PanNET P2/3 subtypes showed enrichment of inflammatory pathways, we investigated the immune signature of PanNETs. Recent advances in cancer therapy have demonstrated clear benefits of targeting the immune system. Immune checkpoint inhibition (ICI) and other forms of immunotherapy have led to impressive survival gains in many patients, including those with metastatic disease.[32,33] This critical step is mediated by MHC-T cell receptor (TCR) interactions. Generally, good responders to immunotherapy show higher immune cell infiltration than poor responders by cells such as CD8$^+$ T cells (immune "hot" tumors), which directly attack tumor cells presenting non-self-antigen on MHC class I molecules.[34] To elucidate the tumor immune microenvironment, especially in immune hot tumors, we performed an integrated proteomic analysis with immunohistochemical assessment of immune cell infiltration in tumor tissue (Figure 4A). PanNETs showed a range of CD3$^+$ and CD8$^+$ T cell infiltrations. Concordant with inflammation signature enrichment in P2/3 subtypes (Figure 1E), the proteomic immune score[35] was significantly higher in P2/3 than in P1/4 (Figure 4B) and was positively correlated with CD3$^+$ and CD8$^+$ cell counts in tumor tissue (Figure 4A). Immune hot tumors, defined as an immune score above the median value, showed enrichment of INFγ and antigen processing/ presentation pathways. Gene set enrichment analysis (GSEA) between immune hot and cold tumors showed significant enrichment of the antigen processing/presentation pathway (Figure 4C). Accordingly, the B2M protein expression level, one of the MHC class I molecules quantified in our proteome dataset, showed a positive correlation with the immune score (Figure 4D). In addition, most proteins involved in peptide processing and transfer to the MHC class I complex and antigen presenting machinery components (APM components) showed higher expression in immune hot tumors than in immune cold tumors (Figure 4E), suggesting that immune hot tumors induce immune cell infiltration by presenting antigens to host immune cells. The external transcriptome datasets (GSE118014, EGAS00001005024) confirmed our findings (Figures S3A and S3B). We then examined IRF1 expression, because this transcription factor regulates the expression of MHC molecules and APM components. IRF1 was highly expressed in immune hot tumors (Figure 4F) and was positively correlated with MHC class I and APM components (Figure 4G). PSMB5 was significantly negatively correlated with IRF1, TAP1/2, and TAPBP levels. As PSMB5 is not a member of the immunoproteasome component, this negative correlation is concordant with a previous report.[36] These correlations were confirmed using the GSE118014 and EGAS00001005024 datasets (Figure S3C). To investigate how immune hot tumors survive under high of immune surveillance, we assessed immune suppressor protein expression (i.e., PD-L1, IDO1, LAG3, and C10orf54 [VISTA]) in tumor cells by IHC. Although the transcriptome datasets showed mRNA upregulation of *LAG3* (in the both GSE118014 and EGAS00001005024 datasets) and *C10orf54 (VISTA)* (quantified only in GSE118014 dataset) (Figure S3D), no tumors in our cohort showed LAG3 and C10orf54 (VISTA) protein expression by IHC assessment (Figure S4). PD-L1 and IDO1 showed slightly higher protein expression in immune hot tumors (yet not statistically significant) (Figures 4H and 4I). We also investigated whether suppressive immune cells were increased in immune hot tumor tissues. FOXP3 protein, a marker of regulatory T cells, showed a significant increase in immune hot tumors by IHC (Figure 4J). CD68, a pan-macrophage marker, showed a slight increase in immune hot tumors (Figure 4K), suggesting possible involvement of macrophages in immune reaction regulation. These results suggest that immune hot tumors may survive under high pressure of immune surveillance by utilizing immunosuppressive molecules (PD-L1 or IDO1) and/or by modulating the suppressive immune microenvironment. As shown here, PD-L1 or IDO1 may be good candidates as inhibitory targets in immune hot PanNETs to enhance host immune attack against these tumors. A treatment strategy that aims to activate antigen processing pathway in immune cold tumors may be also effective, which makes immune cold PanNETs visible to host immune system. Since IRF1 is regulated by the IFNγ-JAK-STAT pathway,[37] IFNγ treatment may be useful especially for immune cold tumors.

## Possible drug targets for PanNET treatment

Nonsurgical treatment of PanNETs is currently limited. Only very few drugs, such as somatostatin analogs or everolimus, are available for patients with PanNET. To facilitate drug development and drug repositioning, we performed differential expression analyses between PanNETs and normal pancreatic tissues to determine which proteins were highly and specifically overexpressed in PanNETs. We found
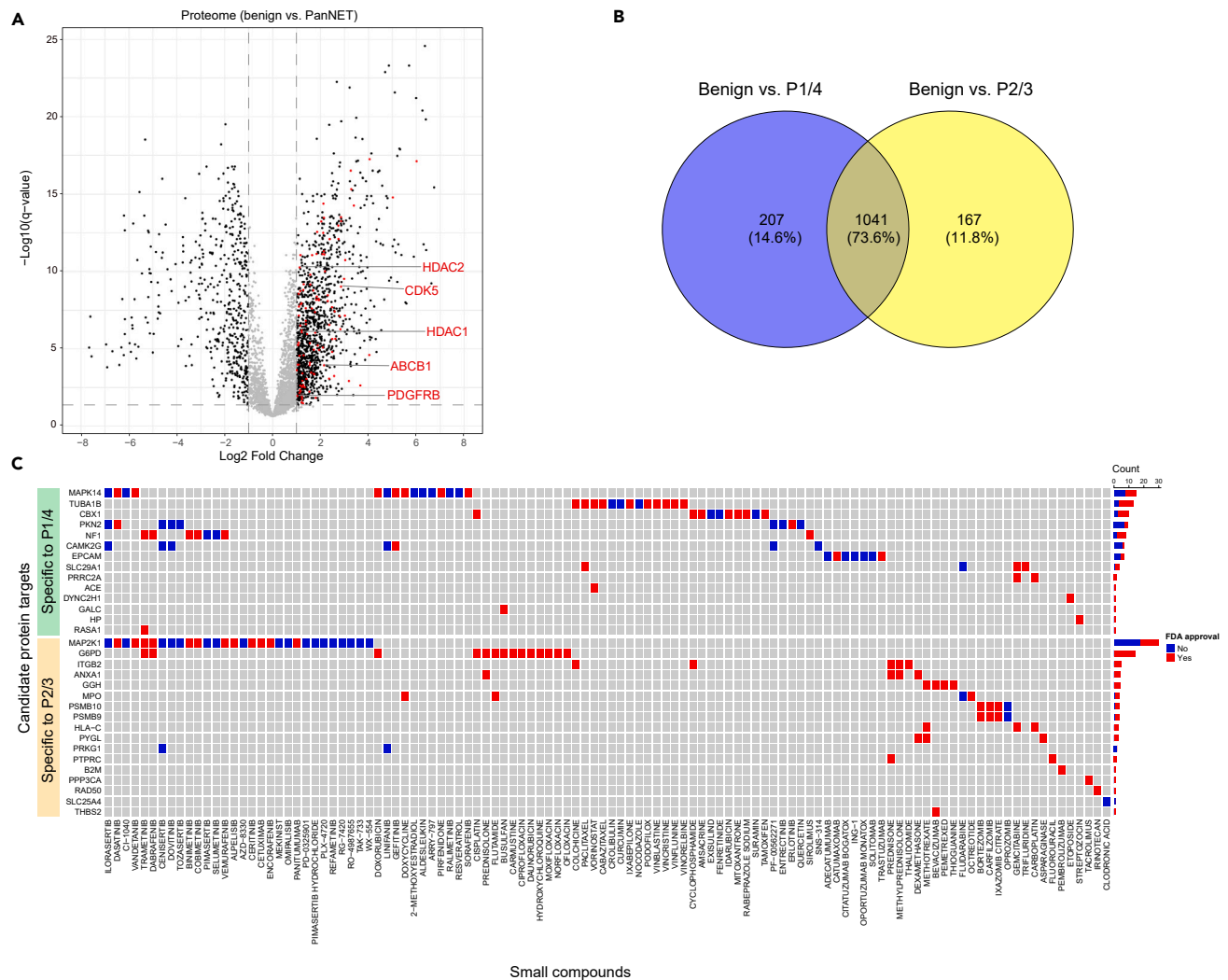
**Figure 5. Possible drug targets for PanNET treatment**

(A) Volcano plot of differential expression between PanNETs and normal pancreatic tissue. Black dots are differentially expressed proteins with statistical significance (FDR <0.05) and > 2-fold change. Red dots represent proteins with inhibitory drugs in the drug-gene interaction database. Labeled proteins (in red) are shown in this plot, highlighting that this analysis detects known therapeutic candidates.

(B) Venn diagram of differentially expressed proteins (benign vs. P1/4 and benign vs. P2/3). Only proteins with FDR <0.05 and >2-fold change are counted.

(C) Possible therapeutic drug candidates targeting subtype-specific proteins with fold change increase >2.

that 1,226 proteins were significantly upregulated (FDR <0.05) with >2-fold change compared to normal pancreatic tissue (Figure 5A; Table S2). We searched for currently available inhibitory drugs against these proteins and found 352 inhibitory drugs against 73 protein targets, including 139 FDA-approved drugs against 38 protein targets (Table S3). For example, HDAC1/2, which has been shown to be a therapeutic target for PanNETs,[38–40] was re-identified as a druggable target in our analysis. ABCB1 is a well-known target which could reverse cancer multidrug resistance, which suggest that ABCB1 inhibition of PanNET may sensitize this tumor to many different types of anticancer drugs.[41] PDGFRB and CDK5 have also been shown to be effective targets in several cancers.[42,43] To identify overexpressed protein targets specific for our proteomic subtypes of PanNET, we performed differential protein expression analyses between benign tissue vs. P1/4 and benign tissue vs. P2/3. We found 207 and 167 specifically overexpressed proteins in P1/4 and P2/3, respectively (FDR <0.05, >2-fold change) (Figure 5B). We then searched for small inhibitory compounds against these subtype-specific protein targets, resulting in 14 and 17 targetable proteins in P1/4 and P2/3, respectively (Figure 5C). For example, two different MAPK family members may be targetable in P1/4 (MAPK14) and P2/3 (MAP2K1, MEK1), respectively. EPCAM is a specific inhibitory target of P1/4, which is concordant with the finding that P1/4 showed relatively higher epithelial phenotype than P2/3 (i.e., EMT low, see Figure 1E). In contrast, G6PD, which is a main enzyme of glycolysis, is a specific inhibitory target of P2/3. This is concordant with a metabolic adaptation feature of P2/3 (Figures 1E and 3). This analysis provides a list of possibly promising subtype-specific candidates for future targeted treatment of PanNETs.

## DISCUSSION

In this study, we describe unsupervised proteomic profiling of well-differentiated PanNETs to discover clinically relevant subtypes and characterize these subtypes with a focus on hypoxia signatures and immune landscapes using integrated proteogenomic methods and IHC validation. Unsupervised proteomic classification identified four subtypes, which form two major groups (P2/3 and P1/4) with clinical outcome differences. SNVs and CNAs related to DNA repair, chromatin remodeling, telomere maintenance, and the mTOR pathway were enriched in the clinically aggressive subtypes (P2/3). A recurrent copy number alteration analysis showed that arm-level amplification of chromosome 7, which harbors several oncogenes, was significantly enriched in clinically aggressive subtypes (P2/3). Recurrent focal peak amplifications and deletions were also observed at the P2/3 subtypes. In addition, KEGG pathway enrichment analysis of genes involved in recurrent focal peaks revealed hypoxia signatures and immunosuppressive pathways enriched in P2/3 subtypes. Differential protein expression analysis between the P1/4 and the P2/3 subtypes also revealed that hypoxia and immune-related signatures were enriched in the P2/3 subtypes. A more detailed characterization of the hypoxia signature in PanNETs revealed association with EMT and metabolic adaptation to a hypoxic microenvironment. While proteomic immune landscape characterization revealed high expression of antigen presenting machinery including MHC class I in the P2/3 subtypes, it also highlighted high expression of immunosuppressive molecules in the tumor and an immunosuppressive microenvironment (regulatory T cell and macrophage infiltration) in the P2/3 subtypes. We also used our proteomic characterization of PanNETs to identify possible drug targets for future therapeutic development or repurposing of currently approved drugs.

Our proteomic subtypes form two major groups: the P1/4 (clinically less aggressive subtypes) and the P2/3 subtypes (clinically aggressive/higher stage subtypes). While the proteomic subtyping was obtained via entirely unsupervised clustering, we found that all clinical recurrence events occurred exclusively in the P2/3 subtypes. We found that hypoxia and immune signatures characterize the P2/3 subtypes and may shape their aggressive clinical phenotypes. Based on the biological characteristics of P2/3 and comparison of inferred P2/3 subtypes in the transcriptomic EGAS00001005024 dataset (Table S5), P2/3 proteome subtypes seem to be partially similar to the previously described "metastasis-like primary" PanNET subtype[9,11] and the "stromal/mesenchymal" subtype.[10] P1/4 may partially overlap with the "Alpha cell-like" subtype.[10] However, the prior classifications were all based on transcriptome analyses only, limiting direct comparisons.

Compatible with a whole-genome sequencing study of 97 PanNETs,[7] somatic mutations in DNA mismatch repair, chromatin remodeling, telomere maintenance, and mTOR pathway pathways showed a higher prevalence of the aggressive P2/3 subtypes relative to the more indolent P1/4 subtypes. Most of the mutations were loss-of-function, suggesting that PanNET tumorigenesis is driven mainly by the loss of tumor suppressor genes. In contrast, recurrent somatic copy number alteration analysis revealed that arm-level amplification of chromosome 7 was significantly enriched in P2/3, suggesting that clinically aggressive phenotypes may be partially derived from genes in this chromosome. Several genes on chromosome 7, such as *CDK6* and *EZH2*,[19,20] are annotated as oncogenic in the COSMIC census gene list.[21] However, functional studies of gene amplifications in PanNETs have not yet been conducted. Recurrent somatic focal copy number analysis revealed copy number alterations in the P2/3 subtypes, which were enriched in HIF1α signaling, PI3K-mTOR pathway, and immune evasion pathway. These signaling pathways were recurrently affected by different genomic alterations (SNVs and CNAs), suggesting their tumorigenic importance.

Previous studies have suggested hypoxia signature enrichment in a subset of PanNETs,[7,10,11] but detailed pathway analyses focusing on metabolic adaptation under hypoxic conditions have not been performed before, especially at the proteomic level. Our proteomic analyses revealed that the hypoxic tumor microenvironment is associated with EMT via TWIST and SNAI upregulation and accompanied by non-SMAD TGF-β signaling. In addition, tumor cells dynamically shift their energy-producing pathways to anaerobic glycolysis in order to adapt to hypoxic conditions. These findings suggest that targeting the EMT signature (EMT transcription factors, MMPs, etc.) and/or key molecules of cancer metabolic reprogramming may be effective, especially for PanNET subtypes P2/3 (hypoxia high tumors). Such therapeutic strategy may be also effective for metastatic lesions since these lesions are generally under limited oxygen and nutrient resources.[44]

As shown in our proteomic study and in a previous transcriptomic study,[11] immune signatures were enriched in the poor survival group. Our study showed upregulation of antigen presentation machinery in immune hot tumors with an increase in inflammatory cell invasion in tumor tissue. IHC-based immune cell profiling revealed an increase of immunosuppressive regulatory T cells in immune hot tumors, which may partially explain clinical tumor aggressiveness even under high immune surveillance pressure. In addition, macrophages, whose role in cancer immunity and outcome correlation are context dependent,[45,46] were increased in immune hot tumors (poor outcome group), perhaps suggesting an immunosuppressive role of macrophages in the PanNETs. Concordantly, a previous transcriptomic study of PanNETs showed macrophage infiltration of immune hot tumors.[11] In addition, we found that immune hot tumors showed higher PD-L1 (CD274) and IDO1 expression, both of which are immunosuppressive molecules. This indicates that immune hot PanNETs (hypoxia high, poor outcome subtypes) may respond to immune checkpoint targeting therapy. Further studies on patient stratification biomarkers for immunotherapy are required.

Finally, we investigated the possibility of drug repositioning by comparing protein expression differences with normal adjacent pancreatic tissues. We were motivated by the facts that currently virtually all drugs target proteins directly and that targeting highly expressed proteins in tumor tissues compared with normal tissue may reduce the possibility of side effects. Our analysis identified 352 putatively inhibitory drugs against 73 PanNET proteins, including 139 FDA-approved drugs against 38 possible target proteins. Our analysis identified HDAC1/2 as a possible drug target that has already been shown to be targetable in PanNETs.[38–40] In addition, we describe possible subtype-specific protein targets, which could be beneficial for developing future proteome subtype-based treatments. Further functional studies will have to interrogate these putative targets.

In summary, we demonstrated that proteogenomic profiling identified clinically relevant subtypes of PanNETs and characterized molecular pathways enriched in poor outcome subtypes. We then characterized a hypoxia signature that features metabolic adaptation

and immune evasion and that is enriched in poor outcome subtypes. Based on our results, targeting metabolic adaptation pathways and immune evasion mechanisms may be beneficial. In addition, our proteogenomic dataset of well-differentiated PanNETs provides a unique resource to the scientific community for studying these tumors.

### Limitations of the study

One limitation is the number of patients in the cohort and the focus on formalin-fixed paraffin-embedded (FFPE) tissue. Both of these will need to be addressed in prospective cohorts of fresh or fresh frozen tissues. Another limitation is the absence of mechanistic studies of our findings. However, functional studies are currently challenging due to the shortage of robust cell culture or organoid models of PanNETs. In addition, FFPE tissue limited our ability to interrogate post-translation modifications such as phosphorylation events that are preferably measured in fresh or fresh frozen samples. The relatively small size of the proteomic cohort makes statistically meaningful association studies between proteomic subtypes and genomic alterations difficult. Nevertheless, our study represents the largest proteomic investigation of well-differentiated PanNETs to date and creates a uniquely valuable dataset for this intriguing disease.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Clinical specimens and pathological data
  - MSK-IMPACT sequencing
  - Recurrent somatic copy number alteration detection
  - Tissue proteome extraction and MS sample preparation from FFPE tissue
  - Quality control sample preparation for mass spectrometry analysis
  - Proteome sequencing by mass spectrometry
  - Immune cell profiling
  - Tissue microarray (TMA) construction
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Protein identification, quantification, and differential expression analysis
  - Transcriptome data processing and differential expression analyses
  - Gene set enrichment analysis
  - Unsupervised consensus clustering for proteomic data
  - Proteome subtype inference of external dataset
  - Immunohistochemistry (IHC) validation
  - Immune score calculation
  - General statistical analysis

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.110544.

### AUTHOR CONTRIBUTIONS

A.T. carried out experiments, performed data analyses, and wrote a draft of the manuscript. M.O., Y.Z., and Y.O. helped with experiments and data analysis. R.H., M.M., and Z.L. supported mass spectrometric experiments and data analyses. D.K. consulted on the study and provided partial funding. J.W. consulted on the study, analyzed data, and edited the manuscript. M.R. designed and supervised the study, analyzed data, wrote and edited the manuscript, and provided funding.

## DECLARATION OF INTERESTS

D.S.K. is now an employee of Paige.AI. M.H.R. is a member of the Scientific Advisory Boards of Azenta Life Sciences and Universal Diagnostics (UDX). None of these companies had any role in the support, design, execution, data analysis, or any other aspect of this study.

## REFERENCES

1. Yao, J.C., Hassan, M., Phan, A., Dagohoy, C., Leary, C., Mares, J.E., Abdalla, E.K., Fleming, J.B., Vauthey, J.N., Rashid, A., and Evans, D.B. (2008). One hundred years after "carcinoid": epidemiology of and prognostic factors for neuroendocrine tumors in 35,825 cases in the United States. J. Clin. Oncol. 26, 3063–3072. https://doi.org/10.1200/jco.2007.15.4377.

2. Korse, C.M., Taal, B.G., van Velthuysen, M.L.F., and Visser, O. (2013). Incidence and survival of neuroendocrine tumours in the Netherlands according to histological grade: experience of two decades of cancer registry. Eur. J. Cancer 49, 1975–1983. https://doi.org/10.1016/j.ejca.2012.12.022.

3. Dasari, A., Shen, C., Halperin, D., Zhao, B., Zhou, S., Xu, Y., Shih, T., and Yao, J.C. (2017). Trends in the Incidence, Prevalence, and Survival Outcomes in Patients With Neuroendocrine Tumors in the United States. JAMA Oncol. 3, 1335–1342. https://doi.org/10.1001/jamaoncol.2017.0589.

4. Digestive System Tumours (2019). WHO Classification of Tumours, 5th Edition.

5. Yang, M., Zeng, L., Ke, N.W., Tan, C.L., Tian, B.L., Liu, X.B., Xiang, B., and Zhang, Y. (2020). World Health Organization grading classification for pancreatic neuroendocrine neoplasms: a comprehensive analysis from a large Chinese institution. BMC Cancer 20, 906. https://doi.org/10.1186/s12885-020-07356-5.

6. Chan, C.S., Laddha, S.V., Lewis, P.W., Koletsky, M.S., Robzyk, K., Da Silva, E., Torres, P.J., Untch, B.R., Li, J., Bose, P., et al. (2018). ATRX, DAXX or MEN1 mutant pancreatic neuroendocrine tumors are a distinct alpha-cell signature subgroup. Nat. Commun. 9, 4158. https://doi.org/10.1038/s41467-018-06498-2.

7. Scarpa, A., Chang, D.K., Nones, K., Corbo, V., Patch, A.M., Bailey, P., Lawlor, R.T., Johns, A.L., Miller, D.K., Mafficini, A., et al. (2017). Whole-genome landscape of pancreatic neuroendocrine tumours. Nature 543, 65–71. https://doi.org/10.1038/nature21063.

8. Jiao, Y., Shi, C., Edil, B.H., de Wilde, R.F., Klimstra, D.S., Maitra, A., Schulick, R.D., Tang, L.H., Wolfgang, C.L., Choti, M.A., et al. (2011). DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. Science 331, 1199–1203. https://doi.org/10.1126/science.1200609.

9. Sadanandam, A., Wullschleger, S., Lyssiotis, C.A., Grötzinger, C., Barbi, S., Bersani, S., Körner, J., Wafy, I., Mafficini, A., Lawlor, R.T., et al. (2015). A Cross-Species Analysis in Pancreatic Neuroendocrine Tumors Reveals Molecular Subtypes with Distinctive Clinical, Metastatic, Developmental, and Metabolic Characteristics. Cancer Discov. 5, 1296–1313.

10. https://doi.org/10.1158/2159-8290.Cd-15-0068.

10. Yang, K.C., Kalloger, S.E., Aird, J.J., Lee, M.K.C., Rushton, C., Mungall, K.L., Mungall, A.J., Gao, D., Chow, C., Xu, J., et al. (2021). Proteotranscriptomic classification and characterization of pancreatic neuroendocrine neoplasms. Cell Rep. 37, 109817. https://doi.org/10.1016/j.celrep.2021.109817.

11. Young, K., Lawlor, R.T., Ragulan, C., Patil, Y., Mafficini, A., Bersani, S., Antonello, D., Mansfield, D., Cingarlini, S., Landoni, L., et al. (2021). Immune landscape, evolution, hypoxia-mediated viral mimicry pathways and therapeutic potential in molecular subtypes of pancreatic neuroendocrine tumours. Gut 70, 1904–1913. https://doi.org/10.1136/gutjnl-2020-321016.

12. Gorai, P.K., Bharti, P.S., Kumar, S., Rajacharya, G.H., Bandyopadhyay, S., Pal, S., Dhingra, R., Kumar, R., Nikolajeff, F., Kumar, S., and Rani, N. (2023). C1QA and COMP: plasma-based biomarkers for early diagnosis of pancreatic neuroendocrine tumors. Sci. Rep. 13, 21021. https://doi.org/10.1038/s41598-023-48323-x.

13. Song, Y.L., Yu, R., Qiao, X.W., Bai, C.M., Lu, C.M., Xiao, Y., Zhong, D.R., Chen, J., Zhao, Y.P., Zhang, T.P., et al. (2017). Prognostic relevance of UCH-L1 and alpha-internexin in pancreatic neuroendocrine tumors. Sci. Rep. 7, 2205. https://doi.org/10.1038/s41598-017-02051-1.

14. Gillette, M.A., Satpathy, S., Cao, S., Dhanasekaran, S.M., Vasaikar, S.V., Krug, K., Petralia, F., Li, Y., Liang, W.W., Reva, B., et al. (2020). Proteogenomic Characterization Reveals Therapeutic Vulnerabilities in Lung Adenocarcinoma. Cell 182, 200–225.e35. https://doi.org/10.1016/j.cell.2020.06.013.

15. Zhang, B., Wang, J., Wang, X., Zhu, J., Liu, Q., Shi, Z., Chambers, M.C., Zimmerman, L.J., Shaddox, K.F., Kim, S., et al. (2014). Proteogenomic characterization of human colon and rectal cancer. Nature 513, 382–387. https://doi.org/10.1038/nature13438.

16. Wilkerson, M.D., and Hayes, D.N. (2010). ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics 26, 1572–1573. https://doi.org/10.1093/bioinformatics/btq170.

17. Kapitsinou, P.P., and Haase, V.H. (2008). The VHL tumor suppressor and HIF: insights from genetic studies in mice. Cell Death Differ. 15, 650–659. https://doi.org/10.1038/sj.cdd.4402313.

18. Mermel, C.H., Schumacher, S.E., Hill, B., Meyerson, M.L., Beroukhim, R., and Getz, G. (2011). GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human

cancers. Genome Biol. 12, R41. https://doi.org/10.1186/gb-2011-12-4-r41.

19. April-Monn, S.L., Andreasi, V., Schiavo Lena, M., Sadowski, M.C., Kim-Fuchs, C., Buri, M.C., Ketkar, A., Maire, R., Di Domenico, A., Schrader, J., et al. (2021). EZH2 Inhibition as New Epigenetic Treatment Option for Pancreatic Neuroendocrine Neoplasms (PanNENs). Cancers 13, 5014. https://doi.org/10.3390/cancers13195014.

20. Shi, Y., Qian, Z.R., Zhang, S., Li, W., Masugi, Y., Li, T., Chan, J.A., Yang, J., Da Silva, A., Gu, M., et al. (2017). Cell Cycle Protein Expression in Neuroendocrine Tumors: Association of CDK4/CDK6, CCND1, and Phosphorylated Retinoblastoma Protein With Proliferative Index. Pancreas 46, 1347–1353. https://doi.org/10.1097/mpa.0000000000000944.

21. Sondka, Z., Bamford, S., Cole, C.G., Ward, S.A., Dunham, I., and Forbes, S.A. (2018). The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. Nat. Rev. Cancer 18, 696–705. https://doi.org/10.1038/s41568-018-0060-1.

22. Ruan, K., Song, G., and Ouyang, G. (2009). Role of hypoxia in the hallmarks of human cancer. J. Cell. Biochem. 107, 1053–1062. https://doi.org/10.1002/jcb.22214.

23. Mittal, V. (2018). Epithelial Mesenchymal Transition in Tumor Metastasis. Annu. Rev. Pathol. 13, 395–412. https://doi.org/10.1146/annurev-pathol-020117-043854.

24. Xiong, G., Stewart, R.L., Chen, J., Gao, T., Scott, T.L., Samayoa, L.M., O'Connor, K., Lane, A.N., and Xu, R. (2018). Collagen prolyl 4-hydroxylase 1 is essential for HIF-1alpha stabilization and TNBC chemoresistance. Nat. Commun. 9, 4456. https://doi.org/10.1038/s41467-018-06893-9.

25. Tanaka, A., Zhou, Y., Shia, J., Ginty, F., Ogawa, M., Klimstra, D.S., Hendrickson, R.C., Wang, J.Y., and Roehrl, M.H. (2020). Prolyl 4-hydroxylase alpha 1 protein expression risk-stratifies early stage colorectal cancer. Oncotarget 11, 813–824. https://doi.org/10.18632/oncotarget.27491.

26. Agarwal, S., Behring, M., Kim, H.G., Bajpai, P., Chakravarthi, B.V.S.K., Gupta, N., Elkholy, A., Al Diffalha, S., Varambally, S., and Manne, U. (2020). Targeting P4HA1 with a Small Molecule Inhibitor in a Colorectal Cancer PDX Model. Transl. Oncol. 13, 100754. https://doi.org/10.1016/j.tranon.2020.100754.

27. Radisky, E.S., and Radisky, D.C. (2010). Matrix metalloproteinase-induced epithelial-mesenchymal transition in breast cancer. J. Mammary Gland Biol. Neoplasia 15, 201–212. https://doi.org/10.1007/s10911-010-9177-x.

28. Deshmukh, A.P., Vasaikar, S.V., Tomczak, K., Tripathi, S., den Hollander, P., Arslan, E., Chakraborty, P., Soundararajan, R., Jolly,

M.K., Rai, K., et al. (2021). Identification of EMT signaling cross-talk and gene regulatory networks by single-cell RNA sequencing. Proc. Natl. Acad. Sci. USA *118*, e2102050118. https://doi.org/10.1073/pnas.2102050118.

29. Zhang, Y.E. (2009). Non-Smad pathways in TGF-beta signaling. Cell Res. *19*, 128–139. https://doi.org/10.1038/cr.2008.328.

30. Huynh, L.K., Hipolito, C.J., and Ten Dijke, P. (2019). A Perspective on the Development of TGF-β Inhibitors for Cancer Treatment. Biomolecules 9, 743. https://doi.org/10.3390/biom9110743.

31. Liberti, M.V., and Locasale, J.W. (2016). The Warburg Effect: How Does it Benefit Cancer Cells? Trends Biochem. Sci. *41*, 211–218. https://doi.org/10.1016/j.tibs.2015.12.001.

32. Robert, C., Schachter, J., Long, G.V., Arance, A., Grob, J.J., Mortier, L., Daud, A., Carlino, M.S., McNeil, C., Lotem, M., et al. (2015). Pembrolizumab versus Ipilimumab in Advanced Melanoma. N. Engl. J. Med. *372*, 2521–2532. https://doi.org/10.1056/NEJMoa1503093.

33. Larkin, J., Chiarion-Sileni, V., Gonzalez, R., Grob, J.J., Cowey, C.L., Lao, C.D., Schadendorf, D., Dummer, R., Smylie, M., Rutkowski, P., et al. (2015). Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma. N. Engl. J. Med. *373*, 23–34. https://doi.org/10.1056/NEJMoa1504030.

34. Kraehenbuehl, L., Weng, C.H., Eghbali, S., Wolchok, J.D., and Merghoub, T. (2022). Enhancing immunotherapy in cancer by targeting emerging immunomodulatory pathways. Nat. Rev. Clin. Oncol. *19*, 37–50. https://doi.org/10.1038/s41571-021-00552-7.

35. Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W., Treviño, V., Shen, H., Laird, P.W., Levine, D.A., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. Nat. Commun. *4*, 2612. https://doi.org/10.1038/ncomms3612.

36. Ferrington, D.A., and Gregerson, D.S. (2012). Immunoproteasomes: structure, function, and antigen presentation. Prog. Mol. Biol. Transl. Sci. *109*, 75–112. https://doi.org/10.1016/b978-0-12-397863-9.00003-1.

37. Der, S.D., Zhou, A., Williams, B.R., and Silverman, R.H. (1998). Identification of genes differentially regulated by interferon alpha, beta, or gamma using oligonucleotide arrays. Proc. Natl. Acad. Sci. USA *95*, 15623–15628. https://doi.org/10.1073/pnas.95.26.15623.

38. Schmitz, R.L., Weissbach, J., Kleilein, J., Bell, J., Hüttelmaier, S., Viol, F., Clauditz, T., Grabowski, P., Laumen, H., Rosendahl, J., et al. (2021). Targeting HDACs in Pancreatic Neuroendocrine Tumor Models. Cells *10*, 1408. https://doi.org/10.3390/cells10061408.

39. Jin, N., Lubner, S.J., Mulkerin, D.L., Rajguru, S., Carmichael, L., Chen, H., Holen, K.D., and LoConte, N.K. (2016). A Phase II Trial of a Histone Deacetylase Inhibitor Panobinostat in Patients With Low-Grade Neuroendocrine Tumors. Oncol. *21*, 785–786. https://doi.org/10.1634/theoncologist.2016-0060.

40. Baradari, V., Huether, A., Höpfner, M., Schuppan, D., and Scherübl, H. (2006). Antiproliferative and proapoptotic effects of histone deacetylase inhibitors on gastrointestinal neuroendocrine tumor cells. Endocr. Relat. Cancer *13*, 1237–1250. https://doi.org/10.1677/erc.1.01249.

41. Engle, K., and Kumar, G. (2022). Cancer multidrug-resistance reversal by ABCB1 inhibition: A recent update. Eur. J. Med. Chem. *239*, 114542. https://doi.org/10.1016/j.ejmech.2022.114542.

42. Mandl, M.M., Zhang, S., Ulrich, M., Schmoeckel, E., Mayr, D., Vollmar, A.M., and Liebl, J. (2017). Inhibition of Cdk5 induces cell death of tumor-initiating cells. Br. J. Cancer *116*, 912–922. https://doi.org/10.1038/bjc.2017.39.

43. Yang, L., Li, N., Xue, Z., Liu, L.R., Li, J., Huang, X., Xie, X., Zou, Y., Tang, H., and Xie, X. (2020). Synergistic therapeutic effect of combined PDGFR and SGK1 inhibition in metastasis-initiating cells of breast cancer. Cell Death Differ. *27*, 2066–2080. https://doi.org/10.1038/s41418-019-0485-4.

44. Tarragó-Celada, J., and Cascante, M. (2021). Targeting the Metabolic Adaptation of Metastatic Cancer. Cancers *13*, 1641. https://doi.org/10.3390/cancers13071641.

45. Zhang, Q.W., Liu, L., Gong, C.Y., Shi, H.S., Zeng, Y.H., Wang, X.Z., Zhao, Y.W., and Wei, Y.Q. (2012). Prognostic significance of tumor-associated macrophages in solid tumor: a meta-analysis of the literature. PLoS One 7, e50946. https://doi.org/10.1371/journal.pone.0050946.

46. Komohara, Y., Jinushi, M., and Takeya, M. (2014). Clinical significance of macrophage heterogeneity in human malignant tumors. Cancer Sci. *105*, 1–8. https://doi.org/10.1111/cas.12314.

47. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods *13*, 731–740. https://doi.org/10.1038/nmeth.3901.

48. Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS *16*, 284–287. https://doi.org/10.1089/omi.2011.0118.

49. Cheng, D.T., Mitchell, T.N., Zehir, A., Shah, R.H., Benayed, R., Syed, A., Chandramohan, R., Liu, Z.Y., Won, H.H., Scott, S.N., et al. (2015). Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): A Hybridization Capture-Based Next-Generation Sequencing Clinical Assay for Solid Tumor Molecular Oncology. J. Mol. Diagn. *17*, 251–264. https://doi.org/10.1016/j.jmoldx.2014.12.006.

50. Marchione, D.M., Ilieva, I., Devins, K., Sharpe, D., Pappin, D.J., Garcia, B.A., Wilson, J.P., and Wojcik, J.B. (2020). HYPERsol: High-Quality Data from Archival FFPE Tissue for Clinical Proteomics. J. Proteome Res. *19*, 973–983. https://doi.org/10.1021/acs.jproteome.9b00686.

51. Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372. https://doi.org/10.1038/nbt.1511.

52. Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. *10*, 1794–1805. https://doi.org/10.1021/pr101065j.

53. Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol. Cell. Proteomics *13*, 2513–2526. https://doi.org/10.1074/mcp.M113.031591.

54. Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proc. Natl. Acad. Sci. USA *98*, 5116–5121. https://doi.org/10.1073/pnas.091062498.

55. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

56. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.

57. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140. https://doi.org/10.1093/bioinformatics/btp616.

58. Barbie, D.A., Tamayo, P., Boehm, J.S., Kim, S.Y., Moody, S.E., Dunn, I.F., Schinzel, A.C., Sandy, P., Meylan, E., Scholl, C., et al. (2009). Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature *462*, 108–112. https://doi.org/10.1038/nature08460.

59. Kuhn, M. (2020). Tidymodels: A Collection of Packages for Modeling and Machine Learning Using Tidyverse Principles.

60. R Core Team (2021). R: A Language and Environment for Statistical Computing (R foundation for statistical computing Vienna, Austria).

# STAR★METHODS

## KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| **Antibodies** | | |
| Rabbit polyclonal anti-HIF1a (dilution 1:100) | Sigma-Aldrich | Cat. # HPA001275; RRID: AB_1079057 |
| Rabbit monoclonal anti-ZEB1 (dilution 1:300) | Sigma-Aldrich | Cat. # HPA027524; RRID: AB_1844977 |
| Rabbit monoclonal anti-SNAI (dilution 1:10,000) | Abcam | Cat. # ab85936; RRID: AB_1925448 |
| Mouse monoclonal anti-Twist (dilution 1:4000) | Abcam | Cat. # ab175430; RRID: AB_2928035 |
| Rabbit monoclonal anti-pAKT (dilution 1:50) | Abcam | Cat. # ab81283; RRID: AB_2224551 |
| Rabbit monoclonal anti-p-mTOR (dilution 1:200) | Cell Signaling | Cat. # 2976; RRID: AB_490932 |
| Rabbit monoclonal anti-IRF1 (dilution 1:100) | Abcam | Cat. # ab243895; RRID: AB_2832955 |
| Rabbit polyclonal anti-IDO1 (dilution 1:50) | Sigma-Aldrich | HPA023072; RRID: AB_1846220 |
| Mouse monoclonal anti-FOXP3 (dilution 1:500) | Abcam | Cat. # ab20034; RRID: AB_445284 |
| Mouse monoclonal anti-CD3 (dilution 1:200) | Leica Biosystems | Cat. # NCL-L-CD3-565; RRID: AB_563541 |
| Mouse monoclonal anti-CD8 (dilution 1:1500) | Dako | Cat. # M7103; RRID: N/A |
| Mouse monoclonal anti-CD68 (dilution 1:200) | Dako | Cat. # M0876; RRID: N/A |
| Recombinant Anti-PD-L1 antibody (dilution 50) | Abcam | Cat. #ab282458 |
| VISTA (D1L2G™) XP Rabbit mAb (dilution 1:50) | Cell Signaling | Cat. # 64953; RRID: AB_2799671 |
| Mouse monoclonal anti-LAG3 (dilution 1:500) | Novus Biologicals | Cat. # NBP1-97657; RRID: AB_11162489 |
| **Chemicals, peptides, and recombinant proteins** | | |
| DL-Dithiothreitol | Sigma-Aldrich | Cat. # 43815 |
| Trypsin/Lys-C | Promega | Cat. # V5073 |
| Acetonitrile | Fisher chemical | Cat. # A955 |
| Empore SPE Disks | Sigma-Aldrich | Cat. # 66883-U |
| Formic acid | Fisher chemical | Cat. # A117-10X1AMP |
| S-Trap | ProtiFi | Cat. # C02-mini |
| Iodoacetamide | Sigma-Aldrich | Cat. # I1149 |
| Sodium dodecyl sulfate | Sigma-Aldrich | Cat. # L3771 |
| **Deposited data** | | |
| Proteome data | This paper | PXD034571 |
| Transcriptome data of PanNETs | (Chan et al.)[6] | GSE118014 |
| Transcriptome data of PanNETs | (Yang et al.) | EGAS00001005024 |
| **Software and algorithms** | | |
| MaxQuant (v1.6.4) | (Cox et al.) | https://maxquant.org/maxquant/ |
| Perseus (v1.6.15) | (Tyanova et al.)[47] | https://maxquant.org/perseus/ |
| clusterProfiler (version 3.16.1) | (Yu et al.)[48] | https://github.com/YuLab-SMU/clusterProfiler |
| single sample GSEA | Broad Institute | https://github.com/broadinstitute/ssGSEA2.0 |
| ConsensusClusterPlus (v1.52.0) | (Wilkerson et al.)[16] | https://git.bioconductor.org/packages/ConsensusClusterPlus |
| ESTIMATE (v1.0.13) | (Yoshihara et al.)[35] | https://sourceforge.net/projects/estimateproject/ |

## RESOURCE AVAILABILITY

### Lead contact

Further information requests should be directed to and will be fulfilled by the lead contact, Michael H. Roehrl (michael_roehrl@bidmc.harvard.edu).

## Materials availability

This study did not generate new reagents.

## Data and code availability

- The mass spectrometry proteomics raw data and processed protein expression data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD034571. The data are publicly available. In the present study, we used 2 external datasets for validation. One contains transcriptome data from 33 PanNETs (GSE118014).[6] The mRNA log2TPM expression data was downloaded from supplementary files of the published paper. The other contains transcriptome data from 84 pancreatic neuroendocrine neoplasms (EGAS00001005024).[10] Bam files were downloaded and processed to calculate gene expression levels as described below.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this study is available from the lead contact upon request.

## EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

The Institutional Review Board granted approval for the study. Formalin-fixed paraffin-embedded (FFPE) tissue blocks were obtained from the institutional biobank. The clinical data of these samples, such as age and gender, were collected from medical records anonymously and are summarized in Table S4. Ancestry/race/ethnicity information was not available from the medical record.

## METHOD DETAILS

### Clinical specimens and pathological data

We selected 37 well-differentiated pancreatic neuroendocrine tumors, 11 pancreatic ductal adenocarcinomas, and 9 normal pancreatic tissues (surgically resected due to non-PanNET conditions and histologically unremarkable) with sufficient tissue for protein extraction. Formalin-fixed paraffin-embedded (FFPE) blocks were obtained and selected from the clinical sample repository after pathological review. All data were retrospectively acquired and anonymized. Clinical data, including patient demographics, treatment history, recurrence status, and clinical MSK-IMPACT targeted sequencing results (10 PanNET and 11 PDAC cases available) were retrieved from the medical records. Histological type and other pathological parameters including Ki67 immunohistochemical assessment were extracted from diagnostic pathology reports, and tumor content ratios for all samples were assessed by gastrointestinal subspecialty pathologists. Detailed clinicopathological features and prognostic information are summarized in Table S4, including gender, age, and TNM stage.

### MSK-IMPACT sequencing

MSK-IMPACT sequencing was performed as previously described.[49] Among 37 PanNETs, 10 PanNETs already had available clinical MSK-IMPACT sequencing data. We checked formalin-fixed paraffin-embedded (FFPE) tissue availability of the remaining 27 PanNETs, and found that 26 PanNETs had sufficient FFPE tissue available for sequencing. We used FFPE tumor blocks that included the tumor center. Briefly, ten unstained 10-μm sections each were prepared from tumor and normal FFPE blocks. We macroscopically dissected tumor areas to enrich for tumor content and extracted genomic DNA using the DNeasy Tissue kit and the EZ1 Advanced XL system (Qiagen, Valencia, CA). Extracted DNA was sheared using a Covaris E200 instrument (Covaris, Woburn, MA, USA). Custom DNA probes were designed for targeted sequencing of all exons and selected introns of 505 genes. Probes were synthesized using the NimbleGen SeqCap EZ library custom oligo system and biotinylated to allow sequence enrichment by capture using streptavidin-conjugated beads. Sequencing libraries were prepared using the KAPA HTP protocol (Kapa Biosystems, Wilmington, MA, USA) and the Biomek FX system (Beckman Coulter, Brea, CA, USA). Pooled libraries containing the captured DNA fragments were sequenced using Illumina HiSeq 2500 to obtain high, uniform coverage (>500x median coverage). All classes of genomic alterations, including substitutions, indels, copy number alterations, and rearrangements, were determined and called against the patient's matched normal sample. The testing was performed in a CLIA-certified laboratory.

### Recurrent somatic copy number alteration detection

Gene-level SCNAs and significant SCNAs in the discovery cohort were identified using Genomic Identification of Significant Targets in Cancer (GISTIC, version 2.0.23) to determine which SCNA regions were significantly amplified or deleted than expected by chance with q value at 0.2.[18] Segmented copy number outputs of MSK-IMPACT were used as input for GISTIC2.0. The following parameters were used: Amplification Threshold = 0.1, Deletion Threshold = 0.1, Confidence Level = 0.90, Arm Level Peel = ON. The default values for other factors were used.

### Tissue proteome extraction and MS sample preparation from FFPE tissue

Sample preparation of FFPE tissues (37 PanNETs, 11 PDACs, 9 normal pancreatic tissues) for mass spectrometric analysis was performed according to a previous protocol for FFPE tissue with modifications.[50] In brief, ten unstained 10-μm sections each were made from the formalin-fixed paraffin-embedded (FFPE) tissues. Adjacent 4-μm sections were subjected to H&E staining to confirm tumor content, after which the tumor areas were macrodissected from the corresponding 10 sections. After dewaxing, lysis buffer containing 100 mM Tris and

5% SDS was added to the samples. Then tissue mixture was sonicated and incubated at 98°C for 20 min at 1000 rpm, and subsequently incubated at 80°C for 2 h at 1000 rpm. After centrifugation at 14,000 g for 30 min at 4°C, the supernatant containing all soluble proteins was collected. Protein concentration was determined using a BCA assay (Pierce). Proteomes (100 μg from each sample) from FFPE tissue were reduced (dithiothreitol), alkylated (iodoacetamide), and double-digested (Trypsin/Lys-C mix, mass spec grade, Promega) and processed using the S-Trap system (ProtiFi) according to the manufacturer's instructions. The samples were desalted using a laboratory-made $C_{18}$ StageTip. Desalted peptides were dried in a SpeedVac vacuum concentrator, re-dissolved in 3% acetonitrile/0.1% formic acid, and stored at -80°C until mass spectrometry analysis.

### Quality control sample preparation for mass spectrometry analysis

As a quality control measure, we pooled all protein samples into one reference sample. All individual patient samples were processed side-by-side with pooled quality control samples in batches. In each mass spectrometry run, we ran one pooled quality control sample per sample batch and calculated the pairwise correlation coefficients between each batch to ensure instrument inter-run stability (Figure S1A).

### Proteome sequencing by mass spectrometry

Desalted peptides were dissolved in 3% acetonitrile/0.1% formic acid and injected into a C18 capillary column (Peptide BEH 1.7μm x 75μm x 250 mm) on a nano ACQUITY UPLC system (Water), which was coupled to a Q Exactive Plus mass spectrometer (Thermo Scientific) via a Proxeon 2 nano electrospray source. Peptides were eluted with a non-linear 200 min gradient of 2-35% buffer B (0.1% (v/v) formic acid, 100% acetonitrile) at a flow rate of 300nl/min. After each gradient, the column was washed with 90% buffer B for 5 min and re-equilibrated with 98% buffer A (0.1% formic acid and 100% HPLC-grade water). MS data were acquired with an automatic switch between a full scan and 10 data-dependent MS/MS scans (TopN method). Target value for the full scan MS spectra was $1 \times 10^6$ ions in the 380-1600 m/z range with a maximum injection time of 50 ms and resolution of 70,000 at 200 m/z with data collected in profile mode. Precursors were selected using a 1.5 m/z isolation width. Precursors were fragmented by higher-energy C-trap dissociation (HCD) with a normalized collision energy of 27 eV. MS/MS scans were acquired at a resolution of 17,500 at 200 m/z with an ion target value of $5 \times 10^4$, maximum injection time of 50 ms, dynamic exclusion for 15 s, and data collected in centroid mode.

### Immune cell profiling

To assess immune cell profile of each sample, we performed IHC using immune cell markers (CD3, CD8, FOXP3, and CD68), and counted the number of positive cells in each TMA core (each sample had two cores). We used the total number (i.e., cell number per two TMA cores) for the downstream analysis.

### Tissue microarray (TMA) construction

To facilitate IHC assessment, we constructed TMA of PanNETs. Among the samples in the cohort, FFPE blocks from 35 PanNETs were available for TMA construction. Three separate 2-mm tissue cores from each tumor sample were drilled out from each donor paraffin block and transferred to tissue array blocks using a robotic TMA arrayer (TMA Grand Master, 3DHistech). Tumor areas were selected based on rigorous review of individual histologic slides for each donor block and electronic image-based coring target area selection using the TMA Grand Master software.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Protein identification, quantification, and differential expression analysis

Label-free protein quantification was carried out with MaxQuant[51,52] with human UniProt FASTA (downloaded 09/2018). "Match between runs" option of MaxQuant was used. A minimum of 1 peptide was required for protein identification, but 2 peptides were required to calculate a protein level ratio. Protein quantification values were normalized using the MaxLFQ algorithm[53] as implemented in MaxQuant. After quantification and data normalization in MaxQuant, protein expression data were processed using Perseus software. In this step, only proteins whose expression values were valid in over 50% of samples from at least one group (i.e., pancreatic neuroendocrine tumors, pancreatic ductal adenocarcinomas, and normal pancreatic tissues) were kept and missing values were imputed with shifted Gaussian distribution, followed by differential expression analyses using Perseus software.[47,54]

### Transcriptome data processing and differential expression analyses

We downloaded 84 RNA-seq bam files which were outputs generated by STAR[55] from EGAS00001005024.[10] Among 84 samples, 78 samples were well-differentiated PanNETs and used for further analyses. Gene counts were obtained using featureCounts[56] with Ensembl (v87) transcript annotations. Transcripts per million (TPM) values were calculated from the gene counts using the formula:

$$\frac{(Count/Gene\ length) \times 1,000,000}{Sum\ (Count/Gene\ length)}$$

To perform differential expression analyses between specified groups, we input the gene counts data into edgeR[57] and obtained significantly expressed genes after filtration of low count data (minimum gene count >10).

### Gene set enrichment analysis

To identify enriched pathways/biological processes, we used clusterProfiler (version 3.16.1) with a ranked gene list ordered by fold-change values.[48] Fold-change values of gene expression between specified groups were calculated using edgeR and Perseus. We conducted enrichment tests of Hallmark and KEGG gene sets obtained from the Molecular Signatures Database (MSigDB version 7.4). For functional characterization of each sample by single sample Gene Set Enrichment Analysis (ssGSEA), we calculated normalized enrichment scores (NES) of cancer-relevant gene sets by projecting the matrix of normalized RNA expression (TPM value) and protein expression onto Hallmark gene sets using ssGSEA implementation available on https://github.com/broadinstitute/ssGSEA2.0 with following parameters: sample.norm.type = "log", weight = 0.75, statistic = "area.under.RES", output.score.type = "NES", nperm = 1000, min.overlap = 3, correl.type = "z.score". To elucidate gene set enrichment in each sample, we applied single sample Gene Set Enrichment Analysis (ssGSEA) to normalized RNA expression data and proteome data.[58] Using ssGSEA, we calculated the normalized enrichment score (NES) for MSigDB terms.

### Unsupervised consensus clustering for proteomic data

To infer proteome-based subtype, we applied consensus clustering on normalized protein expression matrix of 37 PanNETs using ConseususClusterPlus R package with the following parameters: maxK = 10, reps = 1,000 bootstraps, pItem = 0.8, pFeature = 1, clusterAlg = "hc", distance = "peason".[16] The number of clusters was mainly determined by four factors: the average pairwise consensus matrix within consensus clusters, delta plot of the relative change in the area under the cumulative distribution function (CDF) curve, and tracking plot. We selected a consensus matrix with k = 4, which seemed to be the cleanest separation (Figure S1).

### Proteome subtype inference of external dataset

To validate biological differences between our proteome subtypes (P1/4 vs P2/3), we constructed a classifier of our proteome subtypes for transcriptome data (78 PanNETs) from EGAS00001005024. To minimize expression profile difference between transcriptome and proteome, we calculated mRNA-protein correlation coefficients of 79 differentially expressed genes between P1/4 and P2/3 and selected 27 genes which showed significant positive correlation between mRNA and protein abundance for classifier construction. We used the tidymodel R package for this modeling.[59] To construct a classifier for our proteome subtypes, we split our cohort (37 PanNETs) into a training dataset (8 P1/4, 19 P2/3 samples) and a test dataset (3 P1/4, 7 P2/3 samples) while retaining the subtype ratio after scaling. We then trained a random forest classifier with the training dataset accompanied by 10-fold cross validation. We applied the trained classifier to the test dataset and got a 1.0 F-means value (which means all sample subtypes were correctly predicted). We then applied the classifier to the EGAS00001005024 dataset after data preprocessing and inferred proteomic subtypes in the dataset (Table S5). We checked for overlap between inferred proteome subtype and reported transcriptome subtype.[10] P1/4 samples (n=15) were comprised of 3 PDX1-high, 11 Alpha cell-like, and 1 Stromal/Mesenchymal subtype(s). P2/3 samples (n=63) were comprised of 17 PDX1-high, 17 Alpha cell-like, 23 Stromal/Mesenchymal, and 6 Proliferative subtypes.

### Immunohistochemistry (IHC) validation

Formalin-fixed paraffin-embedded tissues were cut into 4-μm sections. Paraffin was removed with xylene, and antigens were retrieved by heat-mediated epitope retrieval (pH 6.0). IHC staining was performed using a Leica BOND-MAX automated system (Wetzlar, Germany) with appropriate positive and negative controls of IHC target molecules. Used antibodies with their product IDs and staining dilutions are listed in the key resources table. IHC results (except CD3, CD8, FOXP3, CD68) were scored using a semiquantitative approach. Staining intensity of individual tumor cells was determined and assigned intensities of 0, 1+, 2+, or 3+ (averaged across 3 independent tissue cores per case). The total weighted IHC score (IHC H-score) of a sample was calculated by multiplying the expression intensity of individual tumor areas (score, 0-3+) by their relative contribution (0-100%) to total tumor area and adding these to yield a total weighted sum. Thus, IHC H-scores have a theoretical range of 0–300. Scoring of all tissue samples was performed independently by two pathologists. In cases of discrepancies in immunohistochemical assessment between the two pathologists, the cases were reviewed together, and a consensus score was determined. Representative IHC images were shown in Figure S4.

### Immune score calculation

Using the ESTIMATE[35] package in R, we calculated immune scores from normalized protein expression matrix of the 37 PanNET samples. Similarly, the immune scores of the GSE118014 cohort and EGAS00001005024 cohort were calculated based on normalized mRNA expression data. We then divided our cohort and the external dataset cohorts (GSE118014, EGAS00001005024) into two groups based on the median immune score in each cohort. We named these groups "immune cold" (low immune score group) and "immune hot" (high immune score group).

### General statistical analysis

The Wilcoxon signed-rank test was used for continuous variables unless otherwise specified. Fisher's exact test was used for categorical variables. Kaplan-Meier survival analysis was used for survival analyses. All statistical analyses were performed using R (v4.3.2).[60]