CANCER

Multiplexed glycan immunofluorescence identification of pancreatic cancer cell subpopulations in both tumor and blood samples

Braelyn Binkowski[1]†, Zachary Klamer[1]†, ChongFeng Gao[1], Ben Staal[1], Anna Repesh[1], Hoang-Le Tran[1], David M. Brass[1], Pamela Bartlett[2], Steven Gallinger[3], Maria Blomqvist[4,5], J. Bradley Morrow[2], Peter Allen[6], Chanjuan Shi[6], Aatur Singhi[7], Randall Brand[7], Ying Huang[8], Galen Hostetter[1], Brian B. Haab[1]*

Copyright © 2025 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

Pancreatic ductal adenocarcinoma (PDAC) tumor heterogeneity impedes the development of biomarker assays for early disease detection. We hypothesized that PDAC cell subpopulations could be identified by aberrant glycan signatures in both tumor tissue and blood samples. We used multiplexed glycan immunofluorescence to distinguish between PDAC and noncancer cell subpopulations within tumor tissue, and we developed hybrid glycan sandwich assays to determine whether the aberrant glycan signatures could be detected in blood samples. We found that PDAC cells were identified by signatures of glycans detected by four glycan-binding proteins (VVL, CA19-9, sTRA, and GM2) and that there are three types of glycan-defined PDAC tumors: sTRA type, CA19-9 type, and intermixed. In patient-matched tumor and blood samples, the PDAC tumor type could be determined by the aberrant glycans in the blood. As a result, the combined assays of aberrant glycan signatures were more sensitive and specific than any individual assay. Our results demonstrate a methodology to detect and stratify PDAC.

INTRODUCTION

The early detection and treatment of pancreatic ductal adenocarcinoma (PDAC) results in better patient outcomes for this extremely aggressive cancer, but only ~15% of patients are diagnosed at a resectable stage (IA, IB, or IIA) where the tumor has not yet spread to major blood vessels, regional lymph nodes, or distant organs (*1*). There are currently no biomarker assays that will identify PDAC at an early, localized stage when curative resection is an option. A barrier to the development of these assays is the heterogeneity between PDAC tumors in the subpopulations of cancer cells they contain. Previous studies revealed that PDAC tumors diverge widely in the intratumor makeup of the PDAC subpopulations and macromolecular compositions. Two primary PDAC cancer cell subpopulations, termed classical and basal-like (or squamous), have been identified by gene expression profiling on whole-tumor mRNA (*2*–*4*), and other, less well-characterized PDAC subpopulations have been identified by single-cell and spatial transcriptomics analyses (*5*, *6*). One study divided the primary subtypes into five categories (*5*), while another categorized the cancer cells by their subclonal lineages and their states (cycling, adhesive, etc.) (*6*). Some tumors contain mostly basal-like cancer cells or mostly classical cancer cells, while other tumors have a mix of the two (*7*, *8*), together with varying numbers of cancer cells that have both classical and basal-like characteristics (*7*, *9*, *10*). Because of this heterogeneity, no single biomarker assay identifies PDACs across their range of subpopulation makeup.

A potential solution to this problem is to use assays that are more specific to each subpopulation in combination to complementarily detect the individual cancer cells subpopulations that make up PDAC tumors. Various approaches have been developed for detecting different cancer cell subpopulations within PDAC tumors. These assays include a 16 gene signature (*11*), multi-protein immunochemical assays (*7*, *12*, *13*), and individual immunochemical assays used to GATA-binding factor 6 (GATA6) (*14*) and keratin 17 (KRT17) (*15*, *16*). These assays could be useful in clinical applications to differentiate between cancer cell subpopulations and guide treatment, such as to predict differential drug responses between the classical and basal-like subtypes (*17*, *18*). The assays, however, are not suitable for the diagnostic use of subpopulation-linked biomarkers for the complementary detection of divergent tumors, as they are not readily detectable in an easily accessible specimen such as blood plasma. An additional challenge is the potential detection of noncancer cells and resulting false-positive identifications of the current protein biomarkers. The previous studies did not determine detection sensitivity relative to all cancer cells or the detection specificity relative to noncancer cells, although multiple types of noncancer cells can also express the transcription factor and keratin protein biomarkers that have been identified. Therefore, we sought to identify molecular patterns that could distinguish between PDAC subpopulations and noncancer cells and that could be detected in blood samples (serum or plasma).

All cells are decorated with complex carbohydrate structures known as glycans. The cell surface glycome, or the complete array of glycans on cell membranes, results through complex biosynthetic processes involving multiple enzymes, substrates, and metabolites. Hence, the cell surface glycome is not linked to the expression levels of any gene or protein but is instead regulated by multiple inputs that are specific to each cell type. This specificity is illustrated by glycans that are already widely used as cell-type markers, such as the glycans bound by CD15 antibodies for neutrophils, CD65s antibodies for

[1]Van Andel Institute, Grand Rapids, MI, USA. [2]Trinity Health, Grand Rapids, MI, USA. [3]University Health Network, Toronto, ON, Canada. [4]Department of Laboratory Medicine, Institute of Biomedicine, University of Gothenburg, Gothenburg, Sweden. [5]Department of Clinical Chemistry, Sahlgrenska University Hospital, Gothenburg, Sweden. [6]Duke University School of Medicine, Durham, NC, USA. [7]University of Pittsburgh Medical Center, Pittsburgh, PA, USA. [8]Fred Hutchinson Cancer Research Center; Seattle, WA, USA.
*Corresponding author. Email: brian.haab@vai.org
†These authors contributed equally to this work.

myeloid differentiation, and the TRA-1-60 antibody for induced pluripotent stem cells, among others. As cells transform from normal cells to abnormal dysplastic cells and then to cancer cells, they activate biological programs of gene expression and metabolism that are not present in the originating cells and that affect the biosynthesis of glycans (*19*–*21*). PDAC cells have glycosylation distinct from noncancer cells and between subsets of PDAC cells, as shown by matrix-assisted laser desorption/ionization (MALDI) glycan imaging (*22*). Furthermore, cancer-associated glycans can be detected using immunochemical assays on both cancer cell surfaces and on cellular secretions in the peripheral blood plasma (*23*), making them ideal for improved cancer identification in surveillance and monitoring applications. Therefore, identifying how the cancer glycome is composed and develops could increase our ability to identify and stratify patients with early-stage PDAC by their unique glycan signatures.

The glycans detected by the CA19-9 antibody and the sialyl tumor-related antigen (sTRA) assay are established PDAC biomarkers that provided impetus to define glycan biomarkers of PDAC subpopulations. The CA19-9 plasma assay is used to confirm diagnoses and monitor the trajectory of PDAC (*24*), but at least half of PDAC cancer cell subpopulations in tumors do not produce CA19-9 (*25*). The glycan detected by the sTRA assay is structurally related to the CA19-9 glycan but is produced by different subpopulations of PDAC cells, so that separate subpopulations of PDAC cells display one, both, or neither of these two glycans. The CA19-9 exclusive cells are epithelial and mostly classical, while the sTRA exclusive cells include both epithelial/classical and mesenchymal/basal-like (*26*), indicating a general but not exact correspondence between expression-based and glycan-based classifications. Owing to the complementarity of the glycans, a subset of tumors may be detected with CA19-9, while another subset is better detected with sTRA. The plasma detection of CA19-9 in combination with sTRA identified significantly more patients with PDAC than by either individual biomarker alone, while maintaining a matched low rate of false-positive identification (*23, 25*–*28*). Therefore, the cancer glycome and the signatures of subpopulation-specific glycans may better identify diverse PDAC tumors than any other approach.

We explored whether distinct PDAC subpopulations display different cancer-associated glycan signatures that enable their specific detection. We approached this question using multiplexed glycan immunofluorescence on primary human PDAC tumors. To our knowledge, this study is the first analysis of cancer-associated glycan signatures on specific cell populations in primary tumors. We found that PDAC cell subpopulations displayed aberrant signatures of glycans that rarely occurred on noncancer cells and that PDAC tumors displayed distinct types of intratumoral heterogeneity in their glycan-defined PDAC subpopulations. The glycan signatures were detectable both in the tumor tissue and in the peripheral blood, such that we could use blood-based assays to identify the distinct subpopulations of PDAC cells in the tumor. As a result, the combined assays for the aberrant glycan signatures improved the sensitivity and specificity of PDAC detection over any individual assay.

RESULTS

PDAC cells display aberrant glycan signatures

To begin to address the challenge presented by the heterogeneity of PDAC cancer cell subpopulations, we asked whether the glycans that decorate the cellular plasma membrane—the glycan signature of the cell—differ between PDAC cancer cells and noncancer cells (Fig. 1A). To develop optimized assays for each of these glycan motifs, we analyzed glycan microarray data from the CarboGrove database (*29*) to select glycan-binding antibodies and lectins [glycan-binding proteins (GBPs)] (Fig. 1B and table S1) that specifically bind to PDAC-associated glycan motifs in addition to the glycans bound by the CA19-9 and sTRA assays. For example, the Tn antigen is overproduced in many cancers including PDAC (*30, 31*); glycan motifs with a terminal N-acetylglucosamine (*22, 32*) have previously been linked to subsets of PDAC cells; the epitope of the TRA-1-60 antibody was found as a marker of chemoresistant, stem-like PDAC cells (*33, 34*), and the GM2 ganglioside—a glycolipid typically associated with neuroendocrine cells—has been identified on PDAC cells with neuroendocrine characteristics (*35*). The glycan array analysis of the GBPs (Supplementary Methods) showed that each GBP specifically binds glycan motifs that encompass the targeted cancer-associated glycan (fig. S1).

These glycans have been investigated individually in PDAC tissues but never in combination. We therefore asked whether individual cancer-associated glycan motifs considered together could better distinguish cancer from noncancer than any single biomarker alone. We used the GBPs that were selected using the CarboGrove data to localize their target glycans in whole block primary tissue specimens using multiplexed glycan immunofluorescence (Fig. 1C), and we used SignalFinder to automate the thresholding of primary image data (*25, 36, 37*). The SignalFinder algorithm is robust to variation in backgrounds and staining intensities, as is common in immunofluorescence datasets, which enabled robust comparisons of signals across tissue specimens. The output of SignalFinder—a binary map of pixels that are above the background (rather than a gradient of signal)—was aligned with the bright-field image of the hematoxylin and eosin (H&E)–stained tissue using the Warpy FiJI plugin (Fig. 1C) (*38, 39*). The use of dichotomized fluorescence data in the data transformations needed for alignment with the bright-field H&E images reduces the pixel shearing artifacts for nonaffine transformations that can occur when using raw fluorescence data. We then quantified the percent of each cell within each region of interest (ROI) that was positive for each glycan.

We analyzed whole block sections from 34 PDAC tumors and 4 noncancerous pancreas specimens (Table 1 and table S2). We used whole block sections to enable comparisons across the wide range of PDAC and noncancer cell subpopulations in PDAC tumors. In a training set of 15 PDAC tumors, we tested the associations between the glycan signatures and cell types annotated by the study pathologist (G.H.) (Fig. 2A and fig. S2). The pathologist annotated 13 types of histological features (table S3 and fig. S2) within ROIs on the bright-field images of the H&E-stained tissue sections, which contained 39 to 111 ROIs in each tumor for a total of 1149 ROIs and 666,544 detected cells. We asked whether we could distinguish between cells in the cancer ROIs and cells in the noncancer ROIs based on the presence of any single glycan alone or the presence of an aberrant permutation, or signature, of glycans. Any given ROI in a complex tumor is not purely of one cell type, and methods are not available for definitively identifying every cancer cell. However, by comparing multiple ROIs that have no or minimal cancer cells to multiple ROIs that have many cancer cells, we could unambiguously test for features that are associated with cancer cells.

We used recursive partitioning to ask whether specific subpopulations of cells in the cancer ROIs display certain glycan signatures

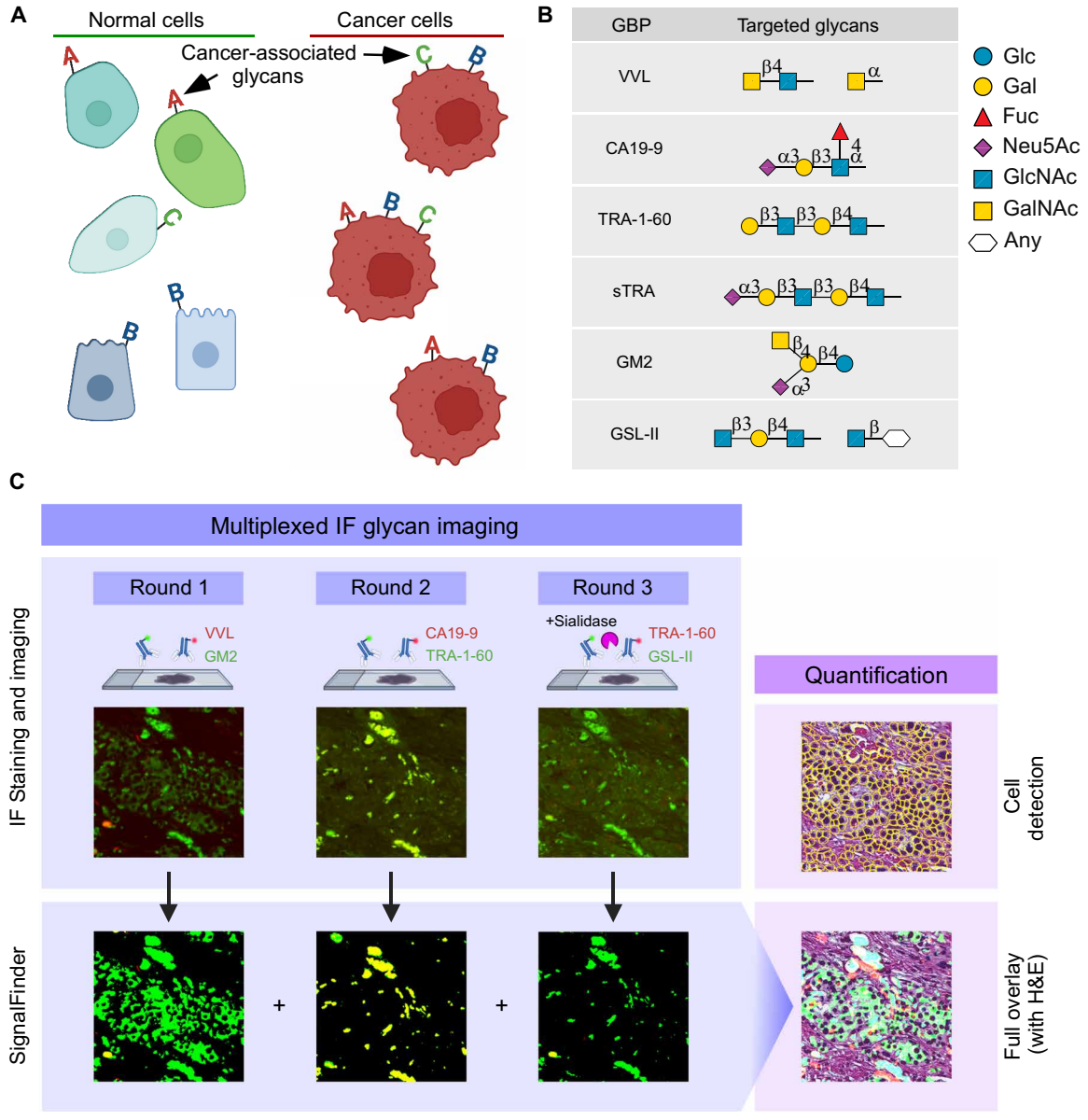

Fig. 1. Multiplexed IF glycan imaging enables identification and quantification of cellular glycan signatures. (**A**) Model of cell surface glycan signatures. (**B**) Glycan-binding proteins (GBPs) and their targets. (**C**) Multiplexed immunofluorescence with automated signal thresholding by SignalFinder, followed by assignments of glycan signatures to individual cells. (A) and (C) created with BioRender.com.

more frequently than the cells in the noncancer ROIs. We chose recursive partitioning because it tests the interactions of all markers without requiring definitions of all interaction terms (63 one-to-five-way terms) and enables straightforward interpretation of high and low production of each marker. We determined which of the six glycan motifs could contribute to distinguishing cancer from noncancer by calculating the importance scores using several classification approaches (table S4). The combined analysis showed that four glycan motifs—those bound by the Vicia Villosa Lectin (VVL), CA19-9, sTRA, and GM2 GBPs—were mainly contributing to the classification of cancer cells (table S4). Because a reduced set of these four glycans performed nearly equivalently to the full set (table S4), we performed the subsequent analyses using the reduced set to limit the number of potential signatures and maximize the power of the analysis.

To enable delineation of the associations between specific glycan signatures and cancer cells, we specified the glycan signature of each cell by classifying each cell as positive or negative for each of the four glycans [using the cutoffs in the fraction-positive values derived from recursive partitioning (fig. S3)] and then using a four-digit binary code to represent the glycan signature. In this binary code system, the glycans retain the order of the GBPs VVL, CA19-9, sTRA, and GM2. Therefore, the code 0100 indicates that CA19-9–detected glycans alone are displayed by a cell. We used these analyses to quantify the percentages of cells in each ROI that are positive for each glycan signature as represented by the four-digit binary code (Fig. 2B).

Table 1. Tumor specimens. N/A, not available; –, not applicable; *, two specimens obtained from one patient.

		Duke	UHN	UPMC	VAI	Total
Specimens		21	4	9	4	38
Sex	Male	9	2	5	2*	16
	Female	12	2	4	1	18
Age at diagnosis	Mean	66.48	57.75	71	–	66.65
	SD	10.66	15.39	8.34	–	11.05
	Median	68	55.5	73	–	68.5
	Min	39	43	56	–	39
	Max	80	77	83	–	83
Specimen type	PDAC	21	4	9	0	34
	Normal	0	0	0	4	4
Site	Head	14	N/A	7	–	
	Tail	7	N/A	2	–	
Margins	Positive	7	0	2	–	9
	Negative	15	4	7	–	26
Lymph nodes	Positive	13	1	2	–	16
	Negative	8	3	7	–	18
Neoadjuvant treated	Yes	0	0	9	–	9
	No	21	4	0	–	29

Using our four-digit binary code, we analyzed the signatures of the glycans displayed on the cells in the analyzed ROIs. We found that cells in the noncancer ROIs were predominantly positive for only one of the four glycans. For example, cells in acinar ROIs were predominantly VVL-bound glycans alone (code 1000) or GM2-bound glycans alone (code 0001) (Fig. 2C). Islet cells and ductal cells also were predominantly positive for only one of the four glycans (fig. S4). In contrast to the noncancer ROIs, the PDAC ROIs included a broader mix of glycan signatures, including cells that produced signatures of 2, 3, or 4 glycans (Fig. 2C). To identify glycan signatures that associate with the cancer cells, we used recursive partitioning to compare the dichotomized glycan values of the cells in the cancer ROIs to the cells in the noncancer ROIs. We found that eight different glycan signatures were more frequently on cancer cells that on noncancer cells (fig. S3). Seven were positive for two to four glycans, and one was positive for only 1 glycan (CA19-9–bound glycans, code 0100) (fig. S3).

In direct comparisons of the ROIs, the PDAC ROIs had significantly higher percentages of cells with all eight of the identified signatures relative to the normal ductal, acinar, or islet ROIs [$P < 0.05$, Wilcoxon test, false discovery rate (FDR) correction] (Fig. 2, D and E). In pooled comparisons of the cells, cells with the eight identified glycan signatures had a significantly higher probability ($P < 0.05$, log odds) of occurring in a cancer ROI than in a normal ductal, acinar, or islet ROI (fig. S4). The PDAC ROIs were less distinct from the ductal proliferation ROIs (a category that broadly encompasses abnormal lesion types), but the pooled comparisons of the cells showed significant associations ($P < 0.05$, log odds) with PDAC of the eight identified glycan signatures (fig. S4). Representative images (Fig. 2F and fig. S5) indicated that normal ductal cells were mainly positive for CA19-9 alone and islet cells were positive mainly for GM2 alone and that the cancer cells were positive for two or three of the glycans. One of the noncancer-associated glycan signatures was higher in the PDAC ROIs relative to the acinar ROIs (code 0101, Fig. 2D), but this signature was higher in all normal-cell ROI relative to acinar ROIs (fig. S4) and therefore was indicative of low levels in acinar cells. The code 0000 signature (low in all four glycans) appeared in all cell types and was not statistically associated with cancer ROIs (Fig. 2D), and a search among these cells using the initial set of six glycans did not reveal additional associations with cancer ROIs.

To determine whether we could use these two to four glycan signatures to differentiate between cancer and noncancer in PDAC tissues, we analyzed a test set of the remaining 19 PDAC specimens and 4 noncancer specimens (Fig. 3A). We first identified regions in each tumor with spatially dense clusters of each of the glycan signatures based on the method described previously (*27*). This method calculates the spatial density of each glycan signal and converts the overlap of the spatially dense regions to distinct ROIs with the corresponding four-digit binary code. We randomly sampled the ROIs at an average of 112 ROIs per tissue (range 20 to 165), for a total of 2135 ROIs (average 308 cells/ROI) (table S5). Using annotations made on the images of the H&E-stained tumors by a pathologist (G.H.) who was blinded to the glycan data, we compared the prevalence of each glycan signature between the annotated cancer ROIs (combined well, moderate, and poor differentiation, 461 ROIs) and the noncancer ROIs (lymphoid, islet, acini, connective tissue, peripheral nerve, blood vessels, normal duct, and macrophages, 791 ROIs) (table S5 and complete list in table S6). We excluded ROIs with tissue-fold artifacts, high levels of debris, or non-pancreatic histology. Each cancer-associated glycan signature was represented in ROIs from multiple tumors, with variation between the tumors in types and variety of glycan signatures (Fig. 3B). Some of the CA19-9–negative cancer ROIs were positive for other cancer-associated glycan signatures (as in tumors 10-16673 and 18137) (Fig. 3B). In addition, 3 of the 19 tumors were completely

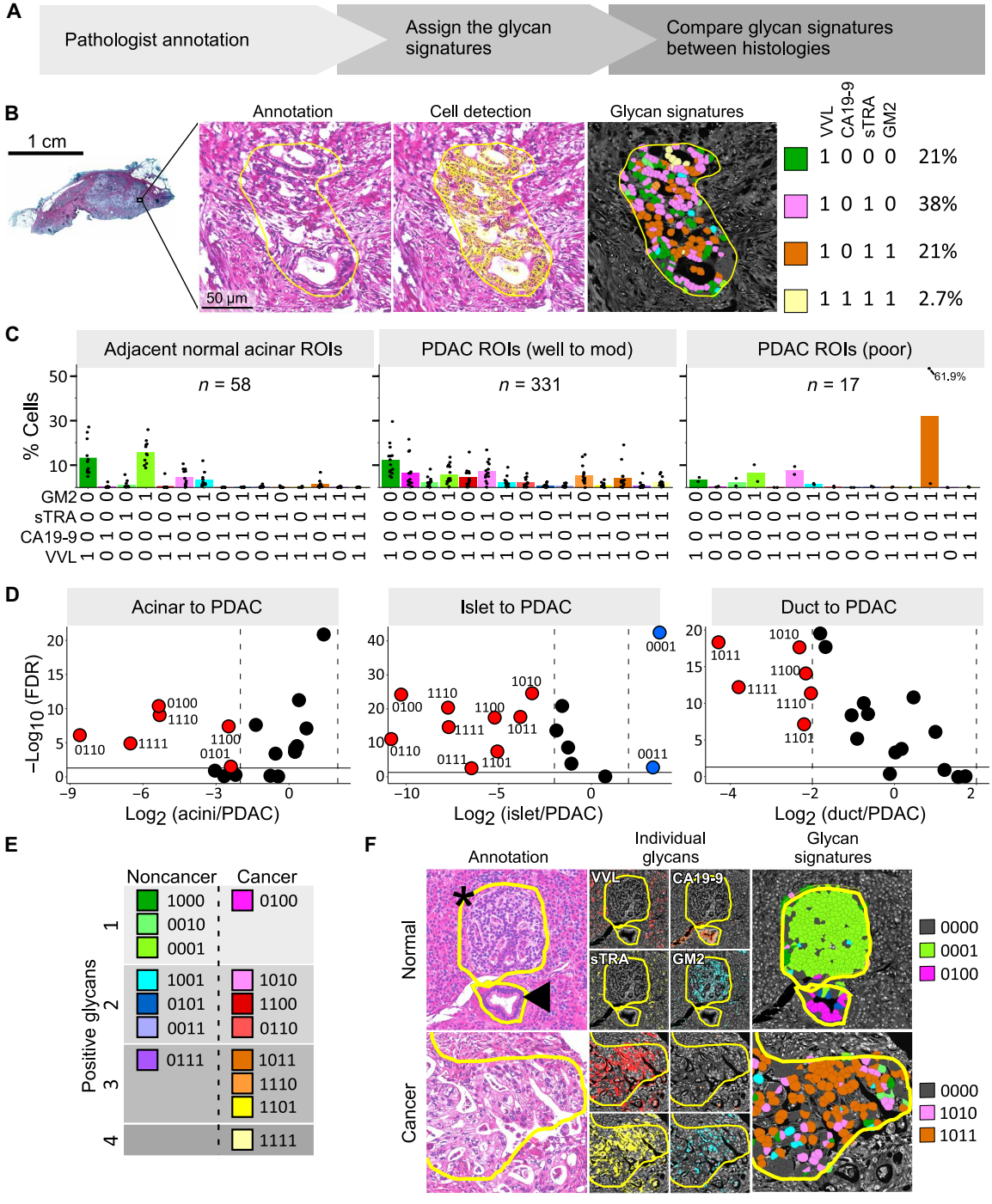

Fig. 2. A training set of PDAC tumor specimens shows cancer association glycan signatures. (**A**) Training design. (**B**) Quantification of cells with each glycan signature in ROIs. (**C**) Distributions of cell surface glycan signatures among acinar and PDAC cells. (**D**) Volcano plots of pairwise comparisons between cell types. (**E**) Cancer-associated and noncancer-associated glycan signatures. (**F**) Representative images and quantifications of noncancer and cancer ROIs. The asterisk marks normal islet cells and the arrowhead marks a normal duct.

negative for CA19-9 (18-460, 33524, and 09-5741); all three had cancer ROIs that were positive for putative cancer-associated glycan signatures (41 to 87% of the cancer ROIs) with 0% positivity in the noncancer ROIs (Fig. 3B).

Of the eight putative cancer-associated glycan signatures determined from the training set, the test set demonstrated that six were significantly associated with PDAC with odds ratios (ORs) from 4.10 to 13.27 [$P = 0.016$ and $P < 0.0001$, respectively, inversion of bootstrap confidence interval (CI) of OR (*40*)] (Fig. 3C). Of the eight predicted noncancer glycan signatures, three were significantly associated with noncancer with ORs of 0.19 or less ($P < 0.0001$, inversion of bootstrap CI of OR) (Fig. 3C and table S7). The ORs of

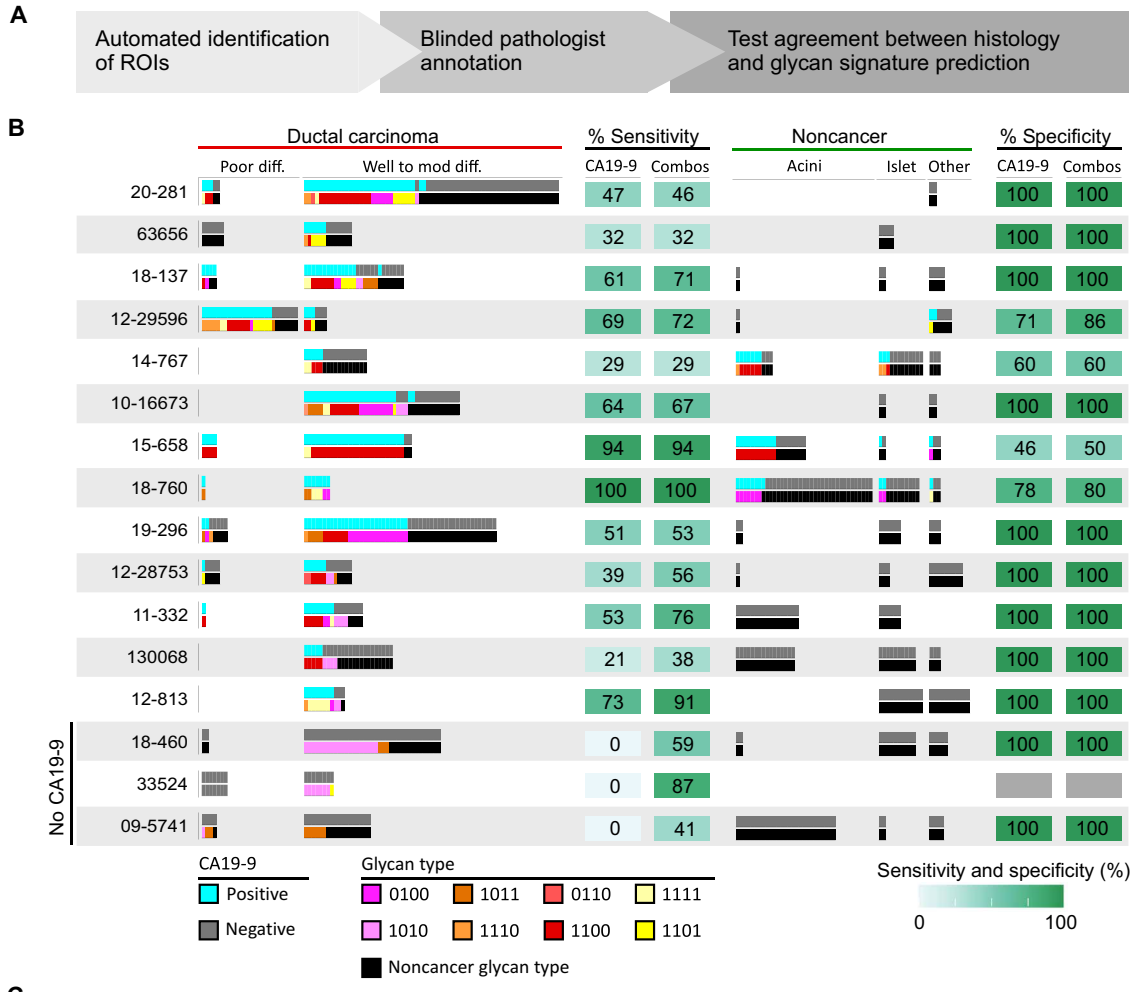

Fig. 3. A test set of PDAC tumor specimens validates the putative PDAC-associated glycan signatures. (**A**) Test set design. (**B**) Patterns of positivity of the putative cancer-associated glycan signatures compared to CA19-9 in individual ROIs in each specimen. (**C**) Prevalence and association with cancer of the predicted cancer-associated (left) and noncancer-associated (right) glycan signatures.

the two nonsignificant predicted cancer signatures (1.66 and 2.74) were higher than all ORs of the predicted noncancer signatures (Fig. 3C), showing that the direction and magnitude of all eight of the predicted cancer-associated glycan signatures were aligned with the cancer ROIs.

To further test the associations with cancer of the eight identified signatures, we asked whether cells in cancer precursor lesions [pancreatic intraductal neoplasia (PanIN)] show a mix of signatures intermediate between normal ductal cells—their closest normal counterpart—and PDAC cells. A study pathologist who did not make the initial histology classifications (C.S.) reviewed the lesions classified as ductal proliferation and identified those that appeared to be PanINs ($n = 25$) across six tumors of the training set. The cells in the PanIN lesions produced glycan signatures typical of normal ducts (the single-glycan codes 1000, 0100, and 0001) and signatures associated with PDAC (1010 and 1011) (fig. S4). The code 1011 cells had a significantly higher probability of occurring in PanIN relative to the normal duct and in PDAC relative to PanIN ROIs ($P < 0.05$, log odds) (fig. S4). This result is consistent with PanINs producing glycan signatures that are intermediate between normal ducts and PDAC.

We therefore investigated whether the use of the glycan signatures could improve the identification of cancer cells over the use of the CA19-9 glycan alone. We determined the number of ROIs that were positive for any of the eight predicted cancer glycan signatures or for any of the eight predicted noncancer glycan signatures and compared these numbers to the pathologist-annotated histology of the ROIs. We also made the same comparison using the numbers of ROIs positive or negative for CA19-9 alone. A significantly higher percentage of cancer ROIs were positive for the predicted cancer glycan signatures (59.7%, CI 51.7 to 69.3) than for CA19-9 alone [46%, CI 32.2 to 52, $P < 0.0001$, inversion of bootstrap CI for difference in sensitivity (*40*)], and an equivalent percentage of noncancer ROIs were positive for the noncancer glycan signatures (89.9%, CI 81.2 to 96.6) as were negative for CA19-9 alone (92.6%, CI 83.4 to 99.4, $P = 0.24$, inversion of bootstrap CI for difference in specificity) (table S8). We made similar observations in comparisons with the other individual glycans (table S8). Thus, these analyses validated the associations with cancer of the eight glycan signatures identified in the training set.

The glycan-defined subpopulations have segregated distributions in spatial clusters and tumors

Previous research showed that PDAC subpopulations tend to segregate into spatially defined clusters within tumors and that PDAC tumors variously contain either a single dominant subpopulation of PDAC cells or a mixture of subpopulations (Fig. 4A) (*7*, *8*). This spatial segregation is indicative of clonal outgrowths of cancer cells with diverging trajectories of DNA mutations (*41, 42*). The images of selected cancer ROIs suggested clonal outgrowth patterns with a general preponderance of one or two glycan-defined subpopulations, rather than a random mix (Fig. 4B). We therefore asked whether the spatial clusters of cancer cells within tumors were relatively homogeneously predominated by one or two cell subpopulations or whether the cancer cell clusters contained random mixtures of cell subpopulations.

Using the whole set of 34 PDAC tumor samples, we identified regions within each tumor harboring spatially clustered cancer-associated glycan signatures using the method described previously (*27*) and determined the glycan signatures of all cells within the clusters (average 30,740 cells per tumor, total of 1,045,142 cells). First examining the intratumoral distributions of the subpopulations without reference to spatial clusters, we observed that the tumors grouped into three distinct types of intratumoral distributions (Fig. 4C). One group contained mainly the PDAC cell subpopulations displaying sTRA in the absence of CA19-9 (codes 1010 and 1011, sTRA exclusive cells), and another group of tumors predominantly contained the subpopulations displaying CA19-9 in the absence of sTRA (codes 1100, 1101, and 0100, CA19-9 exclusive cells). A third group of tumors featured a heterogenous mix of subpopulations, including both the sTRA exclusive and the CA19-9 exclusive cells, as well as the dual expressing cells that display both CA19–9 and sTRA (codes 0110, 1110, and 1111). The Gini index of distribution inequality was different between the tumor groups, ranging from average 0.86 in the sTRA-type tumors to 0.63 in the CA19-9–type tumors to 0.54 in the mixed-type tumors (Fig. 4C) (where 1 is completely homogeneous). Such a range between homogeneous and heterogeneous intratumoral distributions in cancer cell subpopulations mirrors the range observed with expression-defined subpopulations (*7*).

We then examined the distribution of the PDAC cell subpopulations within the spatial clusters of cancer cells. Using the pathologist-annotated ROIs with PDAC histology (total of 2135 ROIs, average of 122 per tumor), we found that the clusters, similar to the tumors, have distributions that range from relatively uniform (containing a single, predominant subpopulation) to mixed (fig. S6). We therefore asked whether only certain glycan-defined cell subpopulations comprise the uniform clusters. We used a mixed effects model to determine the probability of cells with each of the glycan signatures appearing within a relatively uniform spatial cluster (Gini index ≥ 0.8) versus within a mixed spatial cluster (Gini index < 0.8) (Fig. 4D). We found that cells belonging to the sTRA exclusive (codes 1010 and 1011) and one of the CA19-9 exclusive (code 0100) subpopulations were significantly more probable ($P < 0.05$, log odds) to appear in uniform clusters, but that cells belonging to the dual expressing subpopulations were not associated with uniform clusters (Fig. 4D).

Classical and basal-like PDAC subpopulations as defined by gene expression have patterns of homogeneity and heterogeneity within spatial clusters and tumors (*7*) that are similar to those observed here. These patterns suggested a lineage model of cancer progression from the classical to the mixed classical/basal-like to the basal-like subpopulations. We therefore asked whether our data are consistent with a lineage model of progression among the glycan-defined PDAC cell subpopulations. We determined the apparent lineages of the glycan-defined PDAC subpopulations using a network graph of their co-occurrences within spatial clusters (Fig. 4E). The relationships indicated that the sTRA exclusive and CA19-9 exclusive glycan signatures (codes 0100 and 1010, respectively) are independent of each other and rarely intermixed within clusters. The signatures that are positive for the additional glycans were subsets of the CA19-9 exclusive and sTRA exclusive cell subpopulations and rarely intermixed within the same clusters (e.g., code 1101 with code 1011). In addition, when we determined the apparent lineages at the tumor level using the percent of cells displaying each glycan signature within each tumor, we observed identical relationships (Fig. 4E). Thus, within this sample set, PDAC tumors occur in defined types that are either predominant in a single glycan-defined cell subpopulation or that have several subpopulations evenly intermixed among them. The sTRA exclusive and CA19-9 exclusive cells segregated into separate spatial clusters and tumors, consistent with a

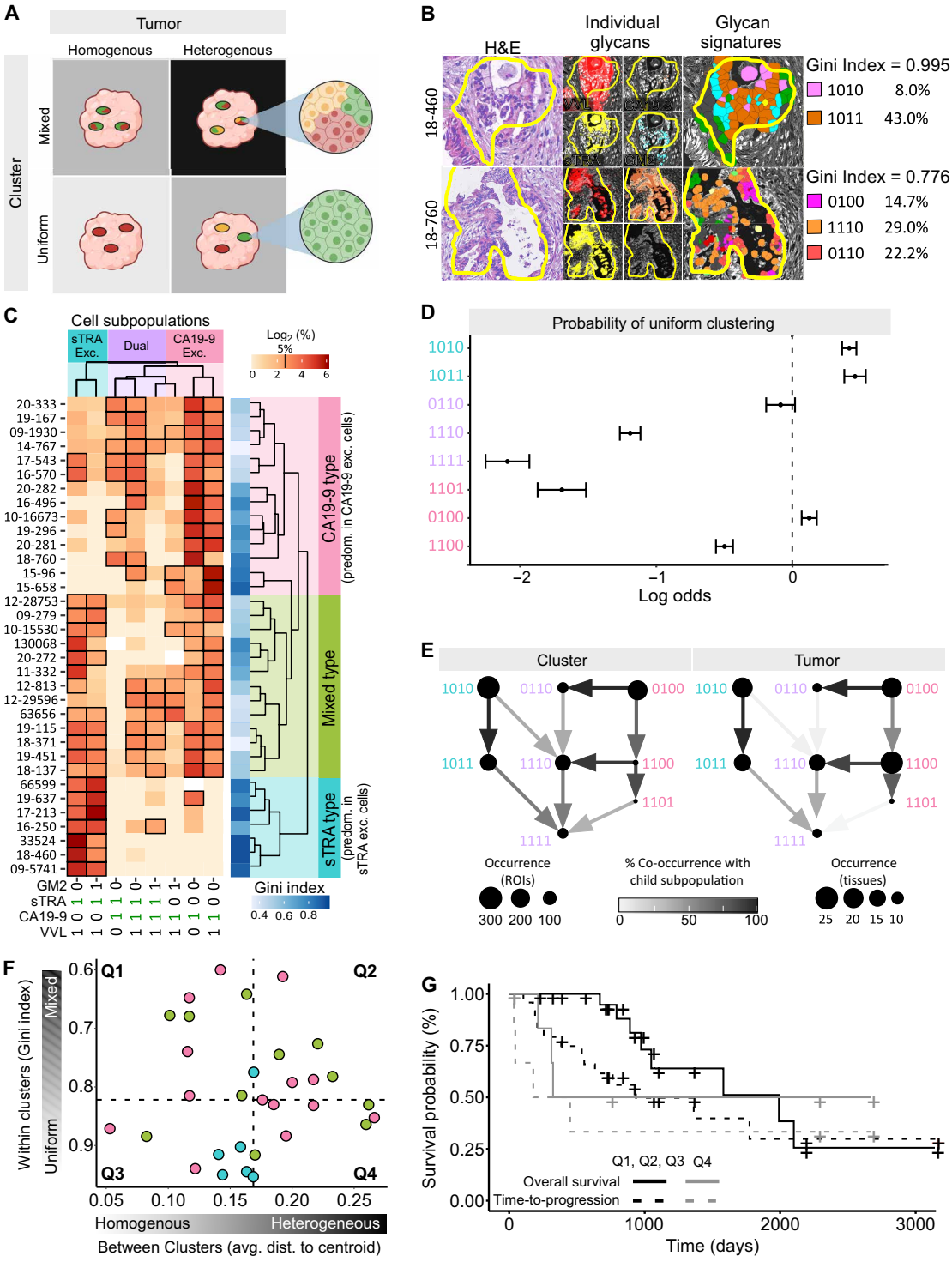

Fig. 4. The glycan-defined PDAC cell subpopulations have unequal distributions among spatial clusters and tumors. (**A**) Cluster and tumor heterogeneity. Clusters may be classified as uniform or mixed, and tumors may be classified as heterogeneous or heterogeneous. Created with BioRender.com. (**B**) Two examples of within-cluster subpopulation purity, or uniformity, in cancer-cell clusters. (**C**) Tumor type classifications based on the predominance of cells with each glycan signature type—CA19-9 exclusive, sTRA exclusive, or dual expressing—in the tumor. (**D**) Probability of association with pure or uniform clusters. Positive log odds are associated with pure clusters. (**E**) Apparent lineages of glycan-defined PDAC cell subpopulations in the cell clusters and tumors. (**F**) Relationship between average subpopulation uniformity in the clusters and the average centroid differences between clusters within each tumor. Each data point is a tumor, color coded by tumor type from (C). The dashed lines are the median values in each axis. (**G**) Kaplan-Meijer plot of patients with PDAC in the different groups defined in (F).

lineage model of progression from the sTRA exclusive and CA19-9 exclusive subpopulations to the dual expressing subpopulations.

If the glycan signatures identify subpopulations that align with clonal outgrowths of PDAC cells, then one also would expect consistency with previous analyses, suggesting that heterogeneity in clonal outgrowths of cancer cells is associated with poor outcome (*43*). We accordingly asked whether any tumors contain different kinds of uniform PDAC cell clusters—postulated to be clonal outgrowths of different PDAC subpopulations—together in the same tumor, as modeled in Fig. 4A. With this model, there are four classes of tumors: homogenous-uniform, homogenous-mixed, heterogeneous-uniform, and heterogeneous-mixed. For each tumor, we calculated both the average within-cluster uniformity (Gini index) and the between-cluster variance (average distance to centroid; fig. S6). We found that a group of eight tumors had both higher than median differences between the PDAC cell clusters and higher than median average purity within the clusters (Fig. 4F). We saw no general relationship between within-cluster homogeneity and between-cluster heterogeneity or with the tumor types defined in Fig. 4C, indicating a broad range of characteristics among PDAC tumors. On the basis of the prediction that outgrowths of different clonal trajectories are predicted to be associated with poor outcomes (*44*), we compared survival probabilities between the patients with heterogenous-uniform tumors and the tumors in the other three groups (Fig. 4G). The time to 50% survival was shorter in the heterogeneous-uniform group (324 days versus 1991 days, respectively), and the 2-year survival rate was lower (50% versus 95%, respectively). In addition, survival after progression was significantly shorter in the heterogeneous-uniform group relative to the other three groups (average 197 days versus 470 days, respectively, $P = 0.048$, Wilcoxon-rank sum test). Overall survival was not significantly different between the groups (log-rank test), potentially owing to a small sample size or not accounting for additional subpopulations not detected by the current signatures, but the substantial difference in 50% and 2-year survival uniquely in the heterogeneous-uniform group suggests a link between clonal outgrowths of different PDAC subpopulations and poor survival.

In vitro assays for cancer-associated glycan signatures enable plasma-based identification of different PDAC tumor types

We next asked whether the aberrant glycan signatures used in tumor tissue to distinguish between cancer cell subpopulations and tumor types could also be used in plasma samples in an analogous way. Different glycans displayed on the same cells can be released together on proteins or cell fragments such as extracellular vesicles (Fig. 5A) (*45*). A standard sandwich immunoassay captures and detects the same antigen, as in the CA19-9 assay. In contrast, a hybrid sandwich immunoassay uses different lectins or antibodies for capture and detection to bind and quantify the glycans present in a sample (Fig. 5B) (*28, 46, 47*). These assays could provide improved cancer specificity by detecting glycans that typically are not coproduced by noncancer cells, such as by capturing the CA19-9 antigen and detecting the sTRA antigen (referred to as the CA19-9.sTRA assay) (Fig. 5B).

We first tested whether PDAC cancer cell lines release glycan-containing moieties that are detectable by in vitro assays using conditioned media (fig. S7). We used 24 different cell lines with a variety of classical or basal-like gene expression traits and a variety of glycan production traits (*26*). We cultured the cell lines on 12-well glass microscope slides and processed these slides through our multiplexed glycan immunofluorescence and analysis pipeline. We confirmed the accuracy of the immunofluorescent glycan detection by comparisons with Western blots for the glycans (fig. S7). At the cellular level, some cell lines were predominant in cells displaying the sTRA exclusive glycan signatures, others were predominant in cell displaying the CA19-9 exclusive glycan signatures, and others included cells displaying the dual expressing glycan signatures (fig. S7). We tested the hybrid glycan assays of all combinations of capture and detection antibodies and lectins (16 assays total, substituting anti-Tn antibody for VVL as capture antibody). The hybrid glycan assays that capture and detect sTRA, CA19-9, or both showed variation between the cell lines that aligned with the cell-surface classifications, consistent with the predominant production of these glycans by specific cell lines. Additional hybrid-glycan assays generally showed lower variation between the cell lines but were higher in individual cell lines that were weaker in one or both CA19-9 and sTRA, such as capture with anti-Tn and detection with VVL or capture with anti-GM2 and detection with CA19-9 (fig. S7). Thus, these data confirmed that the in vitro assays accurately capture and detect the released glycans with levels that potentially reflect the cellular levels of the corresponding glycan-defined subpopulations.

We then asked whether these hybrid glycan sandwich assays could detect glycans in PDAC patient-derived blood samples that would correctly identify glycan-defined PDAC cell subpopulations within their matched tumor samples. We analyzed 13 blood samples (plasma and serum) that had a patient-matched tumor specimen in the above analysis. As with the cell lines, the patients with CA19-9–type tumors had the highest RFU values of secreted CA19–9 (determined using the CA19–9.CA19–9 assay), the patients with sTRA type tumors had the highest relative fluorescence units (RFU) values of secreted sTRA (determined using the TR4.sTRA assay), and the patients with mixed type tumors had low RFU values levels across the four combinations of CA19-9 and sTRA (including CA19-9.sTRA and TR4.sTRA) (Fig. 5C). Using the measurements from these four assays, a decision tree with subtyping thresholds for each assay (table S9) correctly classified each blood sample by the type of the matched tumor in all but one case (Fig. 5C). Thus, using the combined measurements from the four assays, the blood sample measurements provided an indirect identification of the different types of PDAC tumors.

To determine whether our hybrid glycan sandwich assays could distinguish between plasma from patients with PDAC and plasma from control subjects, we analyzed plasma samples from 85 patients with PDAC and 60 control subjects (table S10). The PDAC subjects had resectable cancer (stage I/II) in 59 of 85 (69%) cases and stage III/IV cancer in 18 of 85 (21%) cases [8 of 85 (9.4%) not available], and the control subjects included healthy individuals and patients with benign conditions requiring differential diagnosis, such as chronic pancreatitis and benign biliary obstruction. We observed that the measurements from each of the four assays were significantly higher in the cancer cases than in the controls ($P < 0.05$, Wilcoxon-rank sum test) (Fig. 5D). No significant associations were detected between any of the assays and sex, age, or type of sample (serum versus plasma) (Fisher exact test).

Therefore, for detection, we combined the assays to produce a single test, in which a sample was classified as a case if above the high-specificity cutoff in any assay (Fig. 5E), as demonstrated previously (*23, 48*). To estimate the sensitivity (true-positive detection) of the combined assays and CA19-9 by itself, we performed 1000-fold

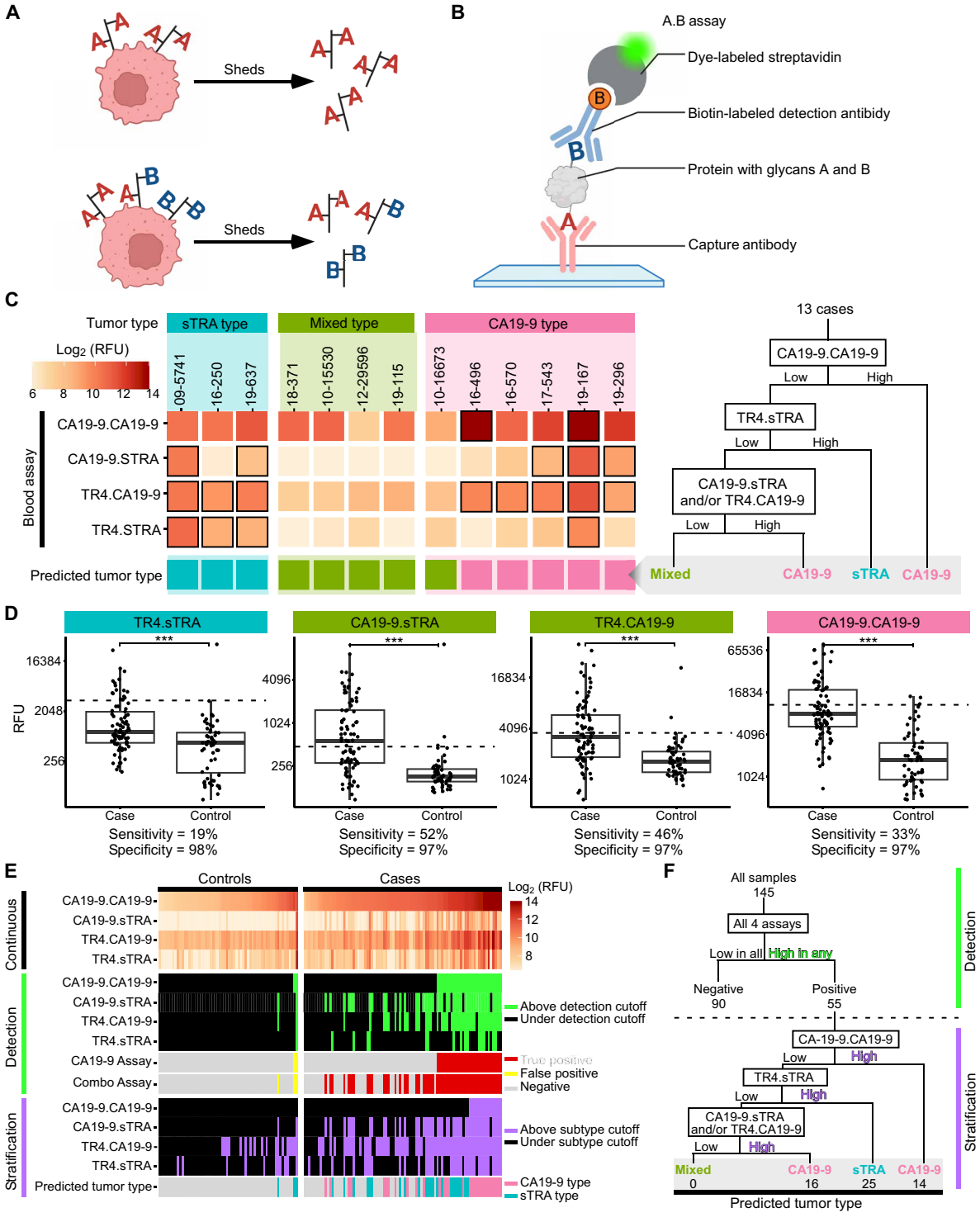

Fig. 5. Blood assays for PDAC cell–released glycans provide complementary detection of tumors with divergent PDAC subpopulations. (**A**) Model of co-secretion of glycans. (**B**) Hybrid glycan sandwich assays. (**C**) Correspondence of combined plasma assays to tumor types. The tumor types at the top were based on the classifications in Fig. 4C, and the predicted tumor types at the bottom were based on the blood assay. (**D**) Cancer-associated relative fluorescent unit (RFU) values in the individual assays. *** indicates $P < 0.001$. (**E**) Complementary detection and stratification of PDACs using the combined assays. The detection and subtype cutoffs are given in table S9. (**F**) Detection and stratification decision tree. (A) and (B) created with BioRender.com.

resampling of the data with threshold selection in each sample based on achieving 95% specificity (true-negative detection) in the controls. The estimated sensitivity was significantly greater for the combined assays (62%, 53 of 85 PDACs) relative to CA19-9 by itself (29%, 25 of 85 PDACs) ($P = 0.042$, inversion of bootstrap CI of the difference), with equivalent specificity for the combined assays (95%, 57 of 60 controls) and CA19-9 by itself (100%, 60 of 60 controls) ($P = 0.380$, inversion of bootstrap CI of the difference). Each of the additional assays other than CA19-9 showed unique elevations among the patients with PDAC, indicating complementary contributions to the improved detection performance (Fig. 5E). This significant improvement in sensitivity over CA19-9 is consistent with the ability to detect both the tumors predominant in the CA19-9 exclusive subpopulations and the tumors predominant in the sTRA exclusive subpopulations.

The ability to detect PDACs with different predominant subpopulations has value not only for earlier detection but also for the stratification of patients based on their tumor characteristics. We therefore simulated a two-stage process of detection followed by stratification, in which we used the blood measurements first to detect PDAC and then to predict the type of PDAC tumor (Fig. 5E). Among the patients that were classified as PDAC by the combined blood assays, we predicted the PDAC tumor type using the cutoffs and decision tree developed above (Fig. 5E). Of the 55 subjects classified as PDAC, 30 subjects were predicted to have the CA19-9 type tumor and 25 were predicted to have the sTRA-type tumor. This algorithm did not detect or predict tumors of the mixed type due to the generally low RFU values of the four assays for that tumor type. However, a lower CA19-9 cutoff in combination with elevation in or both of two additional assays, Tn.GM2 and Tn.VVL, identified seven PDAC samples and zero control samples with predicted mixed-type tumors (fig. S8). This result points to weaker production of CA19-9 by the mixed type tumors and concurrent production of additional cancer-associated glycans. Thus, these results indicate that the combined hybrid-glycan assays provide a means to improve PDAC detection and stratification using a blood test to identify PDACs across the spectrum of tumor types.

DISCUSSION

In the present report, we demonstrate a method to detect distinct PDAC cell subpopulation and tumor types. We used blood assays for the tumor-released glycans to indirectly determine the predominant PDAC subpopulations in the tumors, and by combining assays that complementarily detect tumors with different predominant subpopulations, we could improve both the detection and stratification of PDAC using blood assays. These developments were based on made three observations. First, PDAC cancer cell subpopulations display two to four glycan signatures of cancer that are rarely found on noncancer cells and that may be detected using our multiplexed glycan immunofluorescence pipeline. Second, using our pipeline, we have identified three broad PDAC cancer cell subpopulations: those defined by CA19-9 alone in the absence of sTRA (CA19-9 exclusive), those defined by sTRA alone in the absence of CA19-9 (sTRA exclusive), or those defined by both CA19-9 and sTRA together (dual expressing). Similarly, at the tumor level, we observed three tumor types that are defined broadly by the presence of predominantly CA19-9 exclusive cells (CA19-9 type), the presence of predominantly sTRA exclusive cells (sTRA type), or the presence of both subpopulations (mixed type). Third, using our hybrid glycan immunoassays to detect pairs of membrane-associated glycans, we found in patient-matched tumor tissue and blood samples that the relative abundances of the detected serum glycans were consistent with the same glycans at the tumor tissue levels, enabling the detection of both the CA19-9–type and sTRA-type tumors. In consequence, the detection of both the CA19-9–type and the sTRA-type tumors showed improved sensitivity and specificity when compared with CA19-9 alone. Because of these technical and conceptual developments, this methodology could enable the clinical development of a simple blood test that would improve early detection of patients with early PDAC.

Previous approaches to differentiating between PDAC cancer cell subpopulations have been based on multigene and multiprotein expression profiles in tumor tissue (*2*–*4*, *7*). This approach has been foundational for defining the classical, basal-like, and additional PDAC subpopulations, but it also is clinically limited because previously investigated RNA and protein biomarkers, such as GATA6 to identify classical PDAC or KRT17 to identify basal-like PDAC, are not detectable in the peripheral blood. Glycans have the advantage of being detectable, in many cases, on both the cell surfaces and in the blood circulation. Previous studies examined glycans in either PDAC tumor tissue or PDAC blood samples, but none has established the correspondence of the two sample types, as demonstrated here.

The glycan-defined PDAC subpopulations that we identified segregated in a similar way to the canonical, expression-based classical and basal-like PDAC subpopulations. For example, the cells with CA19-9 exclusive signatures aligned with well differentiated PDAC (fig. S3) and, in a previous study of cell lines (*26*), with epithelial morphology and the classical gene-expression signature, which is consistent with the canonical classical subpopulation. Similarly, the cells with sTRA exclusive signatures aligned with poorly differentiated PDAC (fig. S3) and previously with mixed epithelial/mesenchymal morphology and the basal-like gene expression signature (*26*), which is consistent with the canonical basal-like subpopulation. Further, the dual-expressing cells that express both CA19-9 and sTRA may correspond to the dual classical-basal cell subpopulation as they both theoretically derive from the same canonical subpopulations. These associations support a model in the differences in glycan production result from biological differences between the expression-based subtypes. Glycans are regulated by the expression of glycan processing genes, the metabolism of monosaccharides (*49*), and the routing of substrates (*19*–*21*). The distinct PDAC cancer cell subpopulations likely have differences in each of these processes, owing to distinct lineage outgrowths from progenitor populations (*43*), changes in the microenvironment (*50*), or alterations in epigenetic drivers (*51*), resulting in a remodeling of the cell-surface and cell-secreted glycans. Multiplexed glycan immunofluorescence that also incorporates protein detection could further delineate the link between glycan and protein or gene expression classification. The methodology established here readily enables such costains, which could be integrated with glycan data to further improve cancer specificity.

Our two to four glycan signatures determined from tumors were translatable to a blood test using a set of hybrid glycan assays. Because of this, we successfully identified from the blood assays the type of PDAC subpopulations that were predominant in the corresponding tumors. This was done by a straightforward assessment of positivity or negativity in four hybrid-glycan assays—one detecting only CA19-9, one detecting only sTRA, and two detecting linked CA19-9 and

sTRA. This test has the advantages of requiring only four standard immunoassays, minimal sample volume, and a straightforward decision algorithm. The ability to detect the glycans released from the different PDAC tumor types led to the detection of 62% of the PDACs, relative to 33% by CA19-9 alone using cutoffs that gave low false-positive detection (5 and 3%, respectively). Using two additional hybrid glycan assays in combination with CA19-9 at a lower cutoff increased detection sensitivity to 71% (60 of 85) with no loss in specificity. These substantial improvements upon CA19-9 using a practical test could be immediately valuable, given that CA19-9 already is being evaluated as part of algorithms to enhance PDAC surveillance among elevated-prevalence populations (*52*). The performance of the combined assays for the discrimination of early PDAC from controls will need to be determined in larger case/control studies and eventually in prospective studies to fully evaluate applicability to PDAC surveillance. In the current study, the sensitivity and specificity of CA19-9 and sTRA used in combination were consistent with previous observations (*23, 25, 26, 48*), supporting the accuracy of the findings. Thus, with the development of clinical immunoassays for each hybrid glycan pair, this methodology could be an efficient and economical approach to improving early PDAC detection.

In addition to early detection, blood assays to determine the PDAC cell subpopulations within a tumor also could have value in patient stratification. In cell-culture (*2*), organoid (*53*), and retrospective patient studies (*2–4*), tumors that are predominant in the classical and basal-like subpopulations show differential responses to treatments, but, owing to the lack of blood biomarkers, these findings are not tested in human studies to track responses in drug trials. Glycan signatures that correspond to the gene expression–defined subpopulations would therefore provide a means to translate those cell line and organoid findings to human studies. Glycan signatures could also provide functional information about cancer progression. The cell-surface glycome maintains the integrity of multicellular organization, potentially enables migration and extravasation, mediates metastatic colonization, modulates immune recognition, and regulates the promotion and suppression of inflammation (*54–57*). CA19-9 and related Lewis family glycans could promote metastatic cell seeding through interactions with selectin family receptors (*58*), and sialylated glycans such as sTRA interact with siglec receptors that lead to immunosuppression (*59*). Therefore, the distinct glycan-defined PDAC subpopulations could have differing propensities to metastasize or differing responses to immunotherapies and other treatments. Consequently, the ability to detect and distinguish the PDAC subpopulations by glycan signatures may enable improved patient stratification and targeted therapy.

This study highlights the power of multiplexed immunofluorescence technology for characterizing molecular signatures in specific cell in primary tissue samples. The current study builds on previous work and adds approaches. Multiplexed immunofluorescence is increasingly used to classify cell types and identify disease-associated cellular patterns, especially in immunological studies (*60, 61*) . Previous studies mainly quantified the average intensity of each marker in each cell and applied pattern discovery algorithms to identify groups of cells. Our work, in contrast, used automated thresholding of the raw images and calculation of the fraction positive signal per cell rather than the average intensity per cell, which enabled consistent, unbiased statistical comparisons across tissue samples. This capability, in turn, enabled a training set/test set design to reliably find and validate cancer-associated signatures. Previous work using multiplexed signatures fitted the classifications to the data, preventing the unbiased application of signatures to new data to make blinded classifications of cell types. Our platform provides the framework for broader, rigorous validation of candidate biomarker signatures.

A limitation of the current study is that the matched set of tumor and blood samples (13 samples each) may not be fully representative of the broader PDAC population. The study also used one tumor block per case, which in some samples represent a minority fraction of the tumor and therefore may not provide an accurate representation of the whole tumor. Larger patient-matched sets will be required to establish the accuracy of the blood assays for tumor-type prediction. Larger cohorts also will be needed to determine the associations between intratumoral heterogeneity in PDAC subpopulations and outcomes. The current study also revealed that the current set of hybrid glycan assays detect mainly the CA19-9–type and sTRA-type tumors but not the heterogeneous tumors. However, the present work indicates that we can develop hybrid-glycan biomarker assays for the mixed-type tumors through additional glycans that potentially improve the detection of mixed-type tumors. Further studies will be required to characterize the cells that were low in all four of the glycans studied here (code 0000). The lack of detection of these cells limited the sensitivity of the detection of PDAC and introduced uncertainty into the calculations of subpopulation heterogeneity. The current study focused on a manageable set of cancer-associated glycans to enable previously unidentified methods development, but a broader search for aberrant glycans using the methodology established here could include additional glycans previously associated with PDAC, such as sialic acids (*59*), O-linked glycans (*62, 63*), N-linked glycans examined by MALDI imaging (*64*), sialyl Lewis X (*46, 65*), and others (*66*).

Furthermore, blood biomarkers could be developed that incorporate the protein carriers of the glycans. We (*47*) and others (*67*) have used sandwich immunoassays that capture a protein and detect the glycans on the captured proteins, which can provide higher (*23, 25, 26, 48*) accuracy for cancer detection than capturing and detecting only the protein. The hybrid-glycan assay format would not detect glycans that are not coreleased on the same protein or cell fragment. Thus, assays that capture a protein and detect a glycan could have additive value if targeting proteins that are specifically secreted by PDAC cells.

Here, we show the ability to identify the presence of specific PDAC cancer cell subpopulations in tumors using a simple blood test and the resulting improvement in the overall identification of resectable PDAC cases in distinction from controls. We achieved this goal through the application of multiplexed glycan immunofluorescence to primary PDAC tumors and the detection of the tumor-released pairs of glycans in the blood plasma with a hybrid glycan sandwich assay. A blood test to detect tumors with divergent levels of distinct PDAC subpopulations could enhance PDAC identification in surveillance for PDAC. In addition, a blood test to identify tumor PDAC subpopulations could be used to translate cell line, organoid, and model-system studies of metastatic potential and therapeutic or radiation susceptibility to human studies of subpopulation-specific responses to treatment regimens. This in turn could inform how PDAC is identified, stratified, and treated in the clinic.

MATERIALS AND METHODS

Human specimens

The tissue samples were collected from surgical resections of patients with PDAC at Duke University Medical Center, University of Pittsburgh Medical Center, Trinity Health Grand Rapids, or the University Health Network, Toronto. All samples were formalin-fixed and paraffin-embedded. The tissue sections used for this study were from tissue blocks not needed for patient evaluation.

Human blood plasma or serum specimens were assembled through the clinical centers at Duke University Medical Center, University of Pittsburgh Medical Center, Trinity Health Grand Rapids, or MD Anderson Cancer Center. The samples were collected before any cancer treatment and were processed under a standard operating procedure approved by the Early Detection Research Network (*68*). All plasma samples used EDTA anticoagulant and were frozen at −80°C within 3 hours of collection. The aliquots used in the study underwent no more than three freeze/thaw cycles before use.

Multiplexed immunofluorescence

We performed immunofluorescence on 5-μm-thick sections cut from formalin-fixed, paraffin-embedded (FFPE) blocks. The FFPE blocks were deparaffinized using CitriSolv Hybrid (Decon Labs, King of Prussia, PA) that contained D-limonene and isopropyl cumene and rehydrated through an ethanol gradient of 100, 95, and 70% followed by washing with phosphate-buffered saline (PBS). Following rehydration, antigen retrieval was achieved through incubating slides in citrate buffer at 100°C for 20 min. Slides were blocked with PBS with 0.05% Tween 20 (PBST, 0.05) and 3% bovine plasma albumin (BSA) for 1 hour at room temperature (RT). Tissue staining was manual and performed in multiple rounds. The primary antibodies and lectins (table S1) were labeled for immunofluorescence with sulfo-cyanine-5 (Cy5) *N*-hydroxysuccinimide (NHS) ester or Cy3 NHS ester (Lumiprobe, MD), according to the instructions from the provider. After dialysis to remove unreacted conjugate, a Cy5-labeled antibody or lectin and a Cy3-labeled antibody or lectin were mixed into the same solution of PBST (0.05) with 3% BSA to a final concentration of 10 μg/ml. Slides were incubated overnight with this solution at 4°C in a humidified chamber.

The following day, the solutions were decanted, and the slides were washed twice in PBST (0.05) and once in 1X PBS, each time for 3 min. The slides were dried via blotting and incubated with 4′,6-diamidino-2-phenylindole (DAPI) (AnaSpec, Fremont, CA) at 10 μg/ml in 1X PBS for 15 min at RT. Two 5-min washes were performed in 1X PBS, and then slides were cover-slipped and scanned using a fluorescent microscope (AxioScan.Z1, Zeiss, Oberkochen, Germany). We next quenched the fluorescence using 6% H_2O_2 in 250 mM sodium bicarbonate (pH 9.5 to 10) and performed another round of immunofluorescence using two different antibodies or lectins using the order of VVL and GM2 in round 1, CA19-9 and TRA-1-60 in round 2, and TRA-1-60 and GSL-II in round 3. Before the second round of detection with the TRA-1-60 antibody, we treated the slides with sialidase to remove terminal sialic acids. For this step, the slides were incubated with a 1:200 dilution (from a 50,000 U/ml stock) of α2-3,6,8 neuraminidase in 5 mM $CaCl_2$ and 50 mM (pH 5.5) sodium acetate overnight at 37°C.

In each round of fluorescence imaging, the microscope collected three images at each field of view, each image corresponding to the emission maxima of DAPI, Cy3, and Cy5. The three images were saved as independent, stacked layers in a single file. Following all fluorescence imaging, the slides were stained with hematoxylin and eosin (H&E) using a standard protocol and digitally imaged using the Aperio ScanScope (Leica).

Image analysis

All fluorescence images were initially processed using SignalFinder, which automatically creates a map of the locations of pixels containing signal in each layer of the multiplexed immunofluorescence as an ome.tiff. This SignalFinder algorithm is robust to variation in backgrounds and staining intensities, as is common in data collection that spans over much time and different operators. The robustness of SignalFinder to variation is essential for accurate comparisons of signals across samples. Furthermore, dichotomized data reduce the variability associated with the data transformations (non-affine transformations) required for the alignment of the fluorescence images with the bright-field H&E image, while continuous data show pixel shearing artifacts that affect cell signal distribution. The map created by SignalFinder was aligned to the H&E image based on the signals from cell nuclei using the Warpy extension in ImageJ (*38, 39*). We used automated cell nuclei detection followed by a 10-μm expansion to estimate cell borders and then quantified the fractional area of each cell that was positive for each image layer. Each image layer corresponded to a single glycan, yielding a set of six fractional-positive values (each on a 0 to 1 scale) for each cell.

Bioinformatics methods

We trained a classification tree to distinguish the cancer from the noncancer cells using the six fractional-positive values (CA19-9, TRA, GM2, VVL, sTRA, and GSL-II) for each cell. The values from TRA and GSL-II had the lowest importance scores and did not affect classification accuracy when removed and were therefore not used in subsequent analyses (table S4). A classification tree created from the remaining four glycans was used to select thresholds in each fractional-positive value to dichotomize each cell as positive or negative for each glycan. The thresholds were selected from the first split in the tree for each glycan using trees from multiple initial selections of the training and validation sets in cross validation. We then used the dichotomized glycan data in a classification tree to identify signatures of glycans associated with cancer or noncancer. Each terminal branch of the resulting tree was assigned a four-digit binary code based on the sequence of positive or negative splits for each glycan and assigned a cancer or noncancer classification based on the majority of cells in the branch.

The analysis of apparent tumor lineage was done through network graphs of glycan signature occurrence and co-occurrence, where occurrence was defined as a minimum of 5% of cells positive for the glycan signature within the cluster or tumor. Percent co-occurrence is the fraction of observations in the child node, which also occur in the parent node. The network graph only considers stepwise accumulation of glycans.

Immunoassays

The immunoassays were based on the method presented earlier (*28, 46*). Capture antibodies were deposited in arrays on microscope slides that were functionalized to enable covalent attachment (Z Biotech, 10401-2) using a microarray printer (Dispendix I.DOT-One). The capture antibodies anti–CA19-9, anti-TR4, anti-GM2, and anti-Tn (table S1) were printed in sodium borate printing buffer (pH 8.5) with 7.5% glycerol at 50 μg/ml. Individual arrays were separated with

a 64-well gasket system (248865, Grace Bio-Labs, Bend, OR). We diluted the samples of human plasma or serum (eightfold) and cell line conditioned media (twofold) into a buffer (1X PBS with 0.1% Tween 20, 0.1% Brij-35, species-specific blocking antibodies, and protease inhibitor) and incubated each sample on an antibody array (preblocked with 1% BSA) overnight. For sTRA detection arrays, incubation with the unlabeled TRA-1-60 antibody (to block native epitopes) was performed before sialidase treatment to expose the sialylated epitopes (*28*). We prepared α2-3 neuraminidase (P0728L, New England Biolabs, Ipswich, MA) at 250 U/ml in GlycoBuffer1 (50 mM sodium acetate and 5 mM $CaCl_2$) and incubated the solution on the arrays for 1 hour or overnight at 37°C. Next, we incubated all arrays with a biotinylated detection antibody or lectin (3 μg/ml in 1X PBS with 0.1% Tween 20 and 0.1% BSA) and subsequently with Cy5-conjugated streptavidin (43-4316, Invitrogen, Carlsbad, CA) (2 μg/ml in the same buffer as the primary antibody). The slides were scanned for fluorescence at 635 nm using a microarray scanner (Innopsys InnoScan 1100 AL). The quantification of fluorescence was performed using SignalFinder-Microarray (*37*).

Serum data were collected in three independent replicates with additional replicates of assays for matched tissue samples to confirm reproducibility, and conditioned media data were collected in one to three independent replicates across the cell lines. To normalize and link the datasets that comprised the final, combined dataset, we set the minimum assay value in each dataset to a common value and multiplied each dataset by a normalization factor that gave a common median across datasets. For the conditioned media data, we normalized the separate assay data (to enable visualization on a common scale) by subtracting the baseline value across each assay and normalized the median of the logged data to a common value.

Study approval

The tissue samples were collected under protocols approved by the Institutional Review Boards at Duke University Medical Center (protocol Pro00100930), University of Pittsburgh Medical Center (protocol CR19070256-035), Trinity Health Grand Rapids (protocol 19-1127-9), or the University Health Network, Toronto (protocol 18-5116.4). The human blood plasma or serum specimens were collected by each site under the approval of the institutional review boards at Duke University Medical Center (protocol Pro00100930), University of Pittsburgh Medical Center (protocol CR19070256-035), Trinity Health Grand Rapids (protocol 19-1127-9), or MD Anderson Cancer Center (protocols PA11-0670_MODCR001 and LAB00-396_CR001). All subjects provided written, informed consent, and in accordance with an assurance filed with and approved by the US Department of Health and Human Services.

Statistical analysis

The analysis of glycan signature differences between cell types was performed using the Wilcoxon-rank sum test with stratification of ROIs by specimen identifier to correct for repeated measures. This was done using the "clusRank" R package (*69*). The *P* values for this test were corrected using the FDR correction procedure as implemented in the "p.adjust" R function.

For estimation of biomarker performance, such as sensitivity, specificity, difference in sensitivity or specificity, and OR with cancer/noncancer using ROIs from PDAC specimens, the 95% percentile bootstrap confidence intervals were estimated by 1000 bootstrap sampling of the PDAC specimens with replacement, and the *P* values were calculated by inversion of the percentile bootstrap CI using the "boot.pval" R package.

The analysis of odds of a signature belonging to different histology types or clusters of different uniformity was calculated using a binomial mixed effects regression model with the logit (log odds) link function as defined in the "glmer" function of the "lme4" package (*70*). We used the 0000 glycan code as the reference group. This allows for estimation of signature association with one histology or ROI type with respect to another while adjusting for tissue differences. The variance between clusters within a tumor was calculated as the average distance of cancer associated glycan signatures distributions of each cluster for a given tumor to the centroid (multivariate median). This was calculated using the "betadisper" function of the "vegan" R package.

Software

We developed the SignalFinder software using MATLAB and C++. Cellular and spatial signal analysis was done in MATLAB and QuPath. We used Microsoft Excel and R for analyzing numerical output and the preparation of graphs, BioRender for the preparation of graphical cartoons, and Canvas XIV for the preparation of the composite figures.

Supplementary Materials

The PDF file includes:

Supplementary Methods
Figs. S1 to S8
Legends for tables S1 to S10
Legends for data S1 to S5
References

Other Supplementary Material for this manuscript includes the following:

Tables S1 to S10
Data S1 to S5

REFERENCES AND NOTES

1. R. L. Siegel, K. D. Miller, N. S. Wagle, A. Jemal, Cancer statistics, 2023. *CA Cancer J. Clin.* **73**, 17–48 (2023).
2. E. A. Collisson, A. Sadanandam, P. Olson, W. J. Gibb, M. Truitt, S. Gu, J. Cooc, J. Weinkle, G. E. Kim, L. Jakkula, H. S. Feiler, A. H. Ko, A. B. Olshen, K. L. Danenberg, M. A. Tempero, P. T. Spellman, D. Hanahan, J. W. Gray, Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nat. Med.* **17**, 500–503 (2011).
3. R. A. Moffitt, R. Marayati, E. L. Flate, K. E. Volmar, S. G. H. Loeza, K. A. Hoadley, N. U. Rashid, L. A. Williams, S. C. Eaton, A. H. Chung, J. K. Smyla, J. M. Anderson, H. J. Kim, D. J. Bentrem, M. S. Talamonti, C. A. Iacobuzio-Donahue, M. A. Hollingsworth, J. J. Yeh, Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat. Genet.* **47**, 1168–1178 (2015).
4. P. Bailey, D. K. Chang, K. Nones, A. L. Johns, A.-M. Patch, M.-C. Gingras, D. K. Miller, A. N. Christ, T. J. C. Bruxner, M. C. Quinn, C. Nourse, L. C. Murtaugh, I. Harliwong, S. Idrisoglu, S. Manning, E. Nourbakhsh, S. Wani, L. Fink, O. Holmes, V. Chin, M. J. Anderson, S. Kazakoff, C. Leonard, F. Newell, N. Waddell, S. Wood, Q. Xu, P. J. Wilson, N. Cloonan, K. S. Kassahn, D. Taylor, K. Quek, A. Robertson, L. Pantano, L. Mincarelli, L. N. Sanchez, L. Evers, J. Wu, M. Pinese, M. J. Cowley, M. D. Jones, E. K. Colvin, A. M. Nagrial, E. S. Humphrey, L. A. Chantrill, A. Mawson, J. Humphris, A. Chou, M. Pajic, C. J. Scarlett, A. V. Pinho, M. Giry-Laterriere, I. Rooman, J. S. Samra, J. G. Kench, J. A. Lovell, N. D. Merrett, C. W. Toon, K. Epari, N. Q. Nguyen, A. Barbour, N. Zeps, K. Moran-Jones, N. B. Jamieson, J. S. Graham, F. Duthie, K. Oien, J. Hair, R. Grützmann, A. Maitra, C. A. Iacobuzio-Donahue, C. L. Wolfgang, R. A. Morgan, R. T. Lawlor, V. Corbo, C. Bassi, B. Rusev, P. Capelli, R. Salvia, G. Tortora, D. Mukhopadhyay, G. M. Petersen, D. M. Munzy, W. E. Fisher, S. A. Karim, J. R. Eshleman, R. H. Hruban, C. Pilarsky, J. P. Morton, O. J. Sansom, A. Scarpa, E. A. Musgrove, U.-M. H. Bailey, O. Hofmann, R. L. Sutherland, D. A. Wheeler, A. J. Gill, R. A. Gibbs, J. V. Pearson, N. Waddell, A. V. Biankin, S. M. Grimmond, Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* **531**, 47–52 (2016).

5. M. Chan-Seng-Yue, J. C. Kim, G. W. Wilson, K. Ng, E. F. Figueroa, G. M. O'Kane, A. A. Connor, R. E. Denroche, R. C. Grant, J. McLeod, J. M. Wilson, G. H. Jang, A. Zhang, A. Dodd, S.-B. Liang, A. Borgida, D. Chadwick, S. Kalimuthu, I. Lungu, J. M. S. Bartlett, P. M. Krzyzanowski, V. Sandhu, H. Tiriac, F. E. M. Froeling, J. M. Karasinska, J. T. Topham, D. J. Renouf, D. F. Schaeffer, S. J. M. Jones, M. A. Marra, J. Laskin, R. Chetty, L. D. Stein, G. Zogopoulos, B. Haibe-Kains, P. J. Campbell, D. A. Tuveson, J. J. Knox, S. E. Fischer, S. Gallinger, F. Notta, Transcription phenotypes of pancreatic cancer are driven by genomic events during tumor evolution. *Nat. Genet.* **52**, 231–240 (2020).
6. W. L. Hwang, K. A. Jagadeesh, J. A. Guo, H. I. Hoffman, P. Yadollahpour, J. W. Reeves, R. Mohan, E. Drokhlyansky, N. V. Wittenberghe, O. Ashenberg, S. L. Farhi, D. Schapiro, P. Divakar, E. Miller, D. R. Zollinger, G. Eng, J. M. Schenkel, J. Su, C. Shiau, P. Yu, W. A. Freed-Pastor, D. Abbondanza, A. Mehta, J. Gould, C. Lambden, C. B. M. Porter, A. Tsankov, D. Dionne, J. Waldman, M. S. Cuoco, L. Nguyen, T. Delorey, D. Phillips, J. L. Barth, M. Kem, C. Rodrigues, D. Ciprani, J. Roldan, P. Zelga, V. Jorgji, J. H. Chen, Z. Ely, D. Zhao, K. Fuhrman, R. Fropf, J. M. Beechem, J. S. Loeffler, D. P. Ryan, C. D. Weekes, C. R. Ferrone, M. Qadan, M. J. Aryee, R. K. Jain, D. S. Neuberg, J. Y. Wo, T. S. Hong, R. Xavier, A. J. Aguirre, O. Rozenblatt-Rosen, M. Mino-Kenudson, C. F. Castillo, A. S. Liss, D. T. Ting, T. Jacks, A. Regev, Single-nucleus and spatial transcriptome profiling of pancreatic cancer identifies multicellular dynamics associated with neoadjuvant treatment. *Nat. Genet.* **54**, 1178–1191 (2022).
7. H. L. Williams, A. D. Costa, J. Zhang, S. Raghavan, P. S. Winter, K. S. Kapner, S. P. Ginebaugh, S. A. Väyrynen, J. P. Väyrynen, C. Yuan, A. W. Navia, J. Wang, A. Yang, T. L. Bosse, R. L. Kalekar, K. E. Lowder, M. C. Lau, D. Elganainy, V. Morales-Oyarvide, D. A. Rubinson, H. Singh, K. Perez, J. M. Cleary, T. E. Clancy, J. Wang, J. D. Mancias, L. K. Brais, E. R. Hill, M. M. Kozak, D. C. Linehan, R. F. Dunne, D. T. Chang, A. C. Koong, A. F. Hezel, W. C. Hahn, A. K. Shalek, A. J. Aguirre, J. A. Nowak, B. M. Wolpin, Spatially resolved single-cell assessment of pancreatic cancer expression subtypes reveals co-expressor phenotypes and extensive intratumoral heterogeneity. *Cancer Res.* **83**, 441–455 (2022).
8. N. Juiz, A. Elkaoutari, M. Bigonnet, O. Gayet, J. Roques, R. Nicolle, J. Iovanna, N. Dusetti, Basal-like and classical cells coexist in pancreatic cancer revealed by single-cell analysis on biopsy-derived pancreatic cancer organoids from the classical subtype. *FASEB J.* **34**, 12214–12228 (2020).
9. E. S. Carpenter, P. Kadiyala, A. M. Elhossiny, S. B. Kemp, J. Li, N. G. Steele, R. Nicolle, Z. C. Nwosu, J. Freeman, H. Dai, D. Paglia, W. Du, K. Donahue, J. Morales, P. I. Medina-Cabrera, M. E. Bonilla, L. Harris, S. The, V. Gunchick, N. Peterson, K. Brown, M. Mattea, C. E. Espinoza, J. McGue, S. M. Kabala, R. K. Baliira, N. M. Renollet, A. G. Mooney, J. Liu, S. Bhalla, J. P. Farida, C. Ko, J. D. Machicado, R. S. Kwon, E.-J. Wamsteker, A. Schulman, M. A. Anderson, R. Law, A. Prabhu, P. A. Coulombe, A. Rao, T. L. Frankel, F. Bednar, J. Shi, V. Sahai, M. P. di Magliano, KRT17^{High}/CXCL8^{+} tumor cells display both classical and basal features and regulate myeloid infiltration in the pancreatic cancer microenvironment. *Clin. Cancer Res.* **30**, 2497–2513 (2024).
10. J. T. Topham, J. M. Karasinska, M. K. C. Lee, V. Csizmok, L. M. Williamson, G. H. Jang, R. E. Denroche, E. S. Tsang, S. E. Kalloger, H. Wong, G. M. O'Kane, R. A. Moore, A. J. Mungall, F. Notta, J. M. Loree, J. M. Wilson, O. Bathe, P. A. Tang, R. Goodwin, J. J. Knox, S. Gallinger, J. Laskin, M. A. Marra, S. J. M. Jones, D. J. Renouf, D. F. Schaeffer, Subtype-discordant pancreatic ductal adenocarcinoma tumors show intermediate clinical and molecular characteristics. *Clin. Cancer Res.* **27**, 150–157 (2021).
11. N. U. Rashid, X. L. Peng, C. Jin, R. A. Moffitt, K. E. Volmar, B. A. Belt, R. Z. Panni, T. M. Nywening, S. G. Herrera, K. J. Moore, S. G. Hennessey, A. B. Morrison, R. Kawalerski, A. Nayyar, A. E. Chang, B. Schmidt, H. J. Kim, D. C. Linehan, J. J. Yeh, Purity Independent Subtyping of Tumors (PurIST), a clinically robust, single-sample classifier for tumor subtyping in pancreatic cancer. *Clin. Cancer Res.* **26**, 82–92 (2020).
12. A. Muckenhuber, A. K. Berger, A. M. Schlitter, K. Steiger, B. Konukiewitz, A. Trumpp, R. Eils, J. Werner, H. Friess, I. Esposito, G. Klöppel, G. O. Ceyhan, M. Jesinghaus, C. Denkert, M. Bahra, A. Stenzinger, M. R. Sprick, D. Jäger, C. Springfeld, W. Weichert, Pancreatic ductal adenocarcinoma subtyping using the biomarkers hepatocyte nuclear factor-1A and cytokeratin-81 correlates with outcome and treatment response. *Clin. Cancer Res.* **24**, 351 (2018).
13. E. M. Noll, C. Eisen, A. Stenzinger, E. Espinet, A. Muckenhuber, C. Klein, V. Vogel, B. Klaus, W. Nadler, C. Rösli, C. Lutz, M. Kulke, J. Engelhardt, F. M. Zickgraf, O. Espinosa, M. Schlesner, X. Jiang, A. Kopp-Schneider, P. Neuhaus, M. Bahra, B. V. Sinn, R. Eils, N. A. Giese, T. Hackert, O. Strobel, J. Werner, M. W. Büchler, W. Weichert, A. Trumpp, M. R. Sprick, CYP3A5 mediates basal and acquired therapy resistance in different subtypes of pancreatic ductal adenocarcinoma. *Nat. Med.* **22**, 278–287 (2016).
14. K. Duan, G.-H. Jang, R. C. Grant, J. M. Wilson, F. Notta, G. M. O'Kane, J. J. Knox, S. Gallinger, S. Fischer, The value of GATA6 immunohistochemistry and computer-assisted diagnosis to predict clinical outcome in advanced pancreatic cancer. *Sci. Rep.* **11**, 14951 (2021).
15. L. Roa-Peña, C. V. Leiton, S. Babu, C.-H. Pan, E. A. Vanner, A. Akalin, J. Bandovic, R. A. Moffitt, K. R. Shroyer, L. F. Escobar-Hoyos, Keratin 17 identifies the most lethal molecular subtype of pancreatic cancer. *Sci. Rep.* **9**, 11239 (2019).
16. L. A. Delgado-Coka, L. Roa-Peña, S. Babu, M. Horowitz, E. F. Petricoin, L. M. Matrisian, E. M. Blais, N. Marchenko, F. D. Allard, A. Akalin, W. Jiang, B. K. Larson, A. E. Hendifar, V. J. Picozzi, M. Choi, K. R. Shroyer, L. F. Escobar-Hoyos, Keratin 17 is a prognostic and predictive biomarker in pancreatic ductal adenocarcinoma. *Am. J. Clin. Pathol.* **162**, 314–326 (2024).
17. G. M. O'Kane, B. T. Grünwald, G.-H. Jang, M. Masoomian, S. Picardo, R. C. Grant, R. E. Denroche, A. Zhang, Y. Wang, B. Lam, P. M. Krzyzanowski, I. M. Lungu, J. M. S. Bartlett, M. Peralta, F. Vyas, R. Khokha, J. Biagi, D. Chadwick, S. Ramotar, S. Hutchinson, A. Dodd, J. M. Wilson, F. Notta, G. Zogopoulos, S. Gallinger, J. J. Knox, S. E. Fischer, GATA6 expression distinguishes classical and basal-like subtypes in advanced pancreatic cancer. *Clin. Cancer Res.* **26**, 4901–4910 (2020).
18. K. L. Aung, S. E. Fischer, R. E. Denroche, G.-H. Jang, A. Dodd, S. Creighton, B. Southwood, S.-B. Liang, D. Chadwick, A. Zhang, G. M. O'Kane, H. A. A. Albaba, S. Moura, R. C. Grant, J. K. Miller, F. Mbabaali, D. Pasternack, I. M. Lungu, J. M. S. Bartlett, S. Ghai, M. Lemire, S. Holter, A. A. Connor, R. A. Moffitt, J. J. Yeh, L. Timms, P. M. Krzyzanowski, N. C. Dhani, D. W. Hedley, F. Notta, J. M. Wilson, M. J. Moore, S. Gallinger, J. J. Knox, Genomics-driven precision medicine for advanced pancreatic cancer - early results from the COMPASS trial. *Clin. Cancer Res.* **24**, 1344–1354 (2017).
19. L. Araujo, P. Khim, H. Mkhikian, C.-L. Mortales, M. Demetriou, Glycolysis and glutaminolysis cooperatively control T cell function by limiting metabolite supply to N-glycosylation. *eLife* **6**, e21330 (2017).
20. M. C. Lucena, P. Carvalho-Cruz, J. L. Donadio, I. A. Oliveira, R. M. de Queiroz, M. M. Marinho-Carvalho, M. Sola-Penna, I. F. de Paula, K. C. Gondim, M. E. McComb, C. E. Costello, S. A. Whelan, A. R. Todeschini, W. B. Dias, Epithelial mesenchymal transition induces aberrant glycosylation through hexosamine biosynthetic pathway activation. *J. Biol. Chem.* **291**, 12917–12929 (2016).
21. M. Rossi, P. Altea-Manzano, M. Demicco, G. Doglioni, L. Bornes, M. Fukano, A. Vandekeere, A. M. Cuadros, J. Fernández-García, C. Riera-Domingo, C. Jauset, M. Planque, H. F. Alkan, D. Nittner, D. Zuo, L. A. Broadfield, S. Parik, A. A. Pane, F. Rizzollo, G. Rinaldi, T. Zhang, S. T. Teoh, A. B. Aurora, P. Karras, I. Vermeire, D. Broekaert, J. V. Elsen, M. M. L. Knott, M. F. Orth, S. Demeyer, G. Eelen, L. E. Dobrolecki, A. Bassez, T. V. Brussel, K. Sotlar, M. T. Lewis, H. Bartsch, M. Wuhrer, P. van Veelen, P. Carmeliet, J. Cools, S. J. Morrison, J.-C. Marine, D. Lambrechts, M. Mazzone, G. J. Hannon, S. Y. Lunt, T. G. P. Grünewald, M. Park, J. van Rheenen, S.-M. Fendt, PHGDH heterogeneity potentiates cancer cell dissemination and metastasis. *Nature* **605**, 747–753 (2022).
22. C. T. McDowell, Z. Klamer, J. Hall, C. A. West, L. Wisniewski, T. W. Powers, P. M. Angel, A. S. Mehta, D. N. Lewin, B. B. Haab, R. R. Drake, Imaging mass spectrometry and lectin analysis of N-linked glycans in carbohydrate antigen–defined pancreatic cancer tissues. *Mol. Cell. Proteomics* **20**, 100012 (2021).
23. B. Staal, Y. Liu, D. Barnett, P. Hsueh, Z. He, C.-F. Gao, K. Partyka, M. W. Hurd, A. D. Singhi, R. R. Drake, Y. Huang, R. E. Brand, A. Maitra, B. B. Haab, The sTRA plasma biomarker: blinded validation of improved accuracy over CA19-9 in pancreatic cancer diagnosis. *Clin. Cancer Res.* **25**, 2745–2754 (2019).
24. U. K. Ballehaninna, R. S. Chamberlain, The clinical utility of serum CA 19-9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: An evidence based appraisal. *J. Gastrointest. Oncol.* **3**, 105–119 (2011).
25. D. Barnett, Y. Liu, K. Partyka, Y. Huang, H. Tang, G. Hostetter, R. E. Brand, A. D. Singhi, R. R. Drake, B. B. Haab, The CA19-9 and Sialyl-TRA antigens define separate subpopulations of pancreatic cancer cells. *Sci. Rep.* **7**, 4020 (2017).
26. C. Gao, L. Wisniewski, Y. Liu, B. Staal, I. Beddows, D. Plenker, M. Aldakkak, J. Hall, D. Barnett, M. K. Gouda, P. Allen, R. Drake, A. Zureikat, Y. Huang, D. Evans, A. Singhi, R. E. Brand, D. A. Tuveson, S. Tsai, B. B. Haab, Detection of chemotherapy-resistant pancreatic cancer using a glycan biomarker, sTRA. *Clin. Cancer Res.* **27**, 226–236 (2021).
27. L. Wisniewski, S. Braak, Z. Klamer, C. Gao, C. Shi, P. Allen, B. B. Haab, Heterogeneity of glycan biomarker clusters as an indicator of recurrence in pancreatic cancer. *Front. Oncol.* **13**, 1135405 (2023).
28. H. Tang, K. Partyka, P. Hsueh, J. Y. Sinha, D. Kletter, H. Zeh, Y. Huang, R. E. Brand, B. B. Haab, Glycans related to the CA19-9 antigen are increased in distinct subsets of pancreatic cancers and improve diagnostic accuracy over CA19-9. *Cell. Mol .Gastroenterol. Hepatol.* **2**, 210–221.e15 (2016).
29. Z. L. Klamer, C. M. Harris, J. M. Beirne, J. E. Kelly, J. Zhang, B. B. Haab, CarboGrove: A resource of glycan-binding specificities through analyzed glycan-array datasets from all platforms. *Glycobiology* **32**, 679–690 (2022).
30. G. F. Springer, T and Tn, general carcinoma autoantigens. *Science* **224**, 1198–1206 (1984).
31. T. Ju, G. S. Lanneau, T. Gautam, Y. Wang, B. Xia, S. R. Stowell, M. T. Willard, W. Wang, J. Y. Xia, R. E. Zuna, Z. Laszik, D. M. Benbrook, M. H. Hanigan, R. D. Cummings, Human tumor antigens Tn and Sialyl Tn arise from mutations in Cosmc. *Cancer Res.* **68**, 1636–1646 (2008).
32. J. Sinha, Z. Cao, J. Dai, H. Tang, K. Partyka, G. Hostetter, D. M. Simeone, Z. Feng, P. J. Allen, R. E. Brand, B. B. Haab, A gastric glycoform of MUC5AC is a biomarker of mucinous cysts of the pancreas. *PLOS ONE* **11**, e0167070 (2016).
33. P. W. Andrews, G. Banting, I. Damjanov, D. Arnaud, P. Avner, Three monoclonal antibodies defining distinct differentiation antigens associated with different high molecular weight

polypeptides on the surface of human embryonal carcinoma cells. *Hybridoma* **3**, 347–361 (1984).
34. L. Tan, X. Duan, P. Mutyala, T. Zhou, S. Amin, T. Zhang, B. Herbst, G. Askan, T. Itkin, Z. Xiang, F. Michelassi, M. D. Lieberman, C. A. Iacobuzio-Donahue, S. D. Leach, T. Evans, S. Chen, A targetable pathway to eliminate TRA-1-60+/TRA-1-81+ chemoresistant cancer cells. *J. Mol. Cell Biol.* **6**, mjad039 (2023).
35. N. Sasaki, K. Hirabayashi, M. Michishita, K. Takahashi, F. Hasegawa, F. Gomi, Y. Itakura, N. Nakamura, M. Toyoda, T. Ishiwata, Ganglioside GM2, highly expressed in the MIA PaCa-2 pancreatic ductal adenocarcinoma cell line, is correlated with growth, invasion, and advanced stage. *Sci. Rep.* **9**, 19369 (2019).
36. E. Ensink, J. Sinha, A. Sinha, H. Tang, H. M. Calderone, G. Hostetter, J. Winter, D. Cherba, R. E. Brand, P. J. Allen, L. F. Sempere, B. B. Haab, Segment and fit thresholding: A new method for image analysis applied to microarray and immunofluorescence data. *Anal. Chem.* **87**, 9715–9721 (2015).
37. D. Barnett, J. Hall, B. Haab, Automated identification and quantification of signals in multichannel immunofluorescence images the SignalFinder-IF platform. *Am. J. Pathology* **189**, 1402–1412 (2019).
38. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
39. N. Chiaruttini, O. Burri, P. Haub, R. Guiet, J. Sordet-Dessimoz, A. Seitz, An open-source whole slide image registration workflow at cellular precision using Fiji, QuPath and Elastix. *Front. Comput. Sci.* **3**, 780026 (2022).
40. M. Thulin, *Modern Statistics with R* (EOS Chasma Press, 2021).
41. A. Malinova, L. Veghini, F. X. Real, V. Corbo, Cell lineage infidelity in PDAC progression and therapy resistance. *Front. Cell Dev. Biol.* **9**, 795251 (2021).
42. A. M. Braxton, A. L. Kiemen, M. P. Grahn, A. Forjaz, J. Parksong, J. M. Babu, J. Lai, L. Zheng, N. Niknafs, L. Jiang, H. Cheng, Q. Song, R. Reichel, S. Graham, A. I. Damanakis, C. G. Fischer, S. Mou, C. Metz, J. Granger, X.-D. Liu, N. Bachmann, Y. Zhu, Y. Liu, C. Almagro-Pérez, A. C. Jiang, J. Yoo, B. Kim, S. Du, E. Foster, J. Y. Hsu, P. A. Rivera, L. C. Chu, F. Liu, E. K. Fishman, A. Yuille, N. J. Roberts, E. D. Thompson, R. B. Scharpf, T. C. Cornish, Y. Jiao, R. Karchin, R. H. Hruban, P.-H. Wu, D. Wirtz, L. D. Wood, 3D genomic mapping reveals multifocality of human pancreatic precancers. *Nature* **629**, 679–687 (2024).
43. T. G. Krieger, S. L. Blanc, J. Jabs, F. W. Ten, N. Ishaque, K. Jechow, O. Debnath, C.-S. Leonhardt, A. Giri, R. Eils, O. Strobel, C. Conrad, Single-cell analysis of patient-derived PDAC organoids reveals cell state heterogeneity and a conserved developmental hierarchy. *Nat. Commun.* **12**, 5826 (2021).
44. I. Dagogo-Jack, A. T. Shaw, Tumour heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol.* **15**, 81–94 (2018).
45. L. Wu, C. Gao, Comprehensive overview the role of glycosylation of extracellular vesicles in cancers. *ACS Omega* **8**, 47380–47392 (2023).
46. H. Tang, S. Singh, K. Partyka, D. Kletter, P. Hsueh, J. Yadav, E. Ensink, M. Bern, G. Hostetter, D. Hartman, Y. Huang, R. E. Brand, B. B. Haab, Glycan motif profiling reveals plasma Sialyl-Lewis X elevations in pancreatic cancers that are negative for Sialyl-Lewis A. *Mol. Cell. Proteomics* **14**, 1323–1333 (2015).
47. S. Chen, T. LaRoche, D. Hamelinck, D. Bergsma, D. Brenner, D. Simeone, R. E. Brand, B. B. Haab, Multiplexed analysis of glycan variation on native proteins captured by antibody microarrays. *Nat. Methods* **4**, 437–444 (2007).
48. B. Haab, L. Qian, B. Staal, M. Jain, J. Fahrmann, C. Worthington, D. Prosser, L. Velokokhatnaya, C. Lopez, R. Tang, M. W. Hurd, G. Natarajan, S. Kumar, L. Smith, S. Hanash, S. K. Batra, A. Maitra, A. Lokshin, Y. Huang, R. E. Brand, A rigorous multi-laboratory study of known PDAC biomarkers identifies increased sensitivity and specificity over CA19-9 alone. *Cancer Lett.* **604**, 217245 (2024).
49. A. Reyes-Oliveras, A. E. Ellis, R. D. Sheldon, B. Haab, Metabolomics and 13C labelled glucose tracing to identify carbon incorporation into aberrant cell membrane glycans in cancer. *Commun. Biol.* **7**, 1576 (2024).
50. S. Raghavan, P. S. Winter, A. W. Navia, H. L. Williams, A. DenAdel, K. E. Lowder, J. Galvez-Reyes, R. L. Kalekar, N. Mulugeta, K. S. Kapner, M. S. Raghavan, A. A. Borah, N. Liu, S. A. Väyrynen, A. D. Costa, R. W. S. Ng, J. Wang, E. K. Hill, D. Y. Ragon, L. K. Brais, A. M. Jaeger, L. F. Spurr, Y. Y. Li, A. D. Cherniack, M. A. Booker, E. F. Cohen, M. Y. Tolstorukov, I. Wakiro, A. Rotem, B. E. Johnson, J. M. McFarland, E. T. Sicinska, T. E. Jacks, R. J. Sullivan, G. I. Shapiro, T. E. Clancy, K. Perez, D. A. Rubinson, K. Ng, J. M. Cleary, L. Crawford, S. R. Manalis, J. A. Nowak, B. M. Wolpin, W. C. Hahn, A. J. Aguirre, A. K. Shalek, Microenvironment drives cell state, plasticity, and drug response in pancreatic cancer. *Cell* **184**, 6119–6137.e26 (2021).
51. J.-S. Roe, C.-I. Hwang, T. D. D. Somerville, J. P. Milazzo, E. J. Lee, B. D. Silva, L. Maiorino, H. Tiriac, C. M. Young, K. Miyabayashi, D. Filippini, B. Creighton, R. A. Burkhart, J. M. Buscaglia, E. J. Kim, J. L. Grem, A. J. Lazenby, J. A. Grunkemeyer, M. A. Hollingsworth, P. M. Grandgenett, M. Egeblad, Y. Park, D. A. Tuveson, C. R. Vakoc, Enhancer reprogramming promotes pancreatic cancer metastasis. *Cell* **170**, 875–888.e20 (2017).
52. J. F. Fahrmann, C. M. Schmidt, X. Mao, E. Irajizad, M. Loftus, J. Zhang, N. Patel, J. Vykoukal, J. B. Dennison, J. P. Long, K.-A. Do, J. Zhang, J. A. Chabot, M. D. Kluger, F. Kastrinos, L. Brais, A. Babic, K. Jajoo, L. S. Lee, T. E. Clancy, K. Ng, A. Bullock, J. Genkinger, M. T. Yip-Schneider, A. Maitra, B. M. Wolpin, S. Hanash, Lead-time trajectory of CA19-9 as an anchor marker for pancreatic cancer early detection. *Gastroenterology* **160**, 1373–1383.e6 (2021).
53. H. Tiriac, P. Belleau, D. D. Engle, D. Plenker, A. Deschênes, T. Somerville, F. E. M. Froeling, R. A. Burkhart, R. E. Denroche, G.-H. Jang, K. Miyabayashi, C. M. Young, H. Patel, M. Ma, J. F. LaComb, R. L. D. Palmaira, A. A. Javed, J. A. Huynh, M. Johnson, K. Arora, N. Robine, M. Shah, R. Sanghvi, A. B. Goetz, C. Y. Lowder, L. Martello, E. Driehuis, N. Lecomte, G. Askan, C. A. Iacobuzio-Donahue, H. Clevers, L. D. Wood, R. H. Hruban, E. D. Thompson, A. J. Aguirre, B. M. Wolpin, A. Sasson, J. Kim, M. Wu, J. C. Bucobo, P. J. Allen, D. V. Sejpal, W. Nealon, J. D. Sullivan, J. M. Winter, P. A. Gimotty, J. L. Grem, D. J. DiMaio, J. M. Buscaglia, P. M. Grandgenett, J. R. Brody, M. A. Hollingsworth, G. M. O'Kane, F. Notta, E. J. Kim, J. M. Crawford, C. E. Devoe, A. Ocean, C. L. Wolfgang, K. H. Yu, E. Li, C. R. Vakoc, B. Hubert, S. E. Fischer, J. M. Wilson, R. A. Moffitt, J. J. Knox, A. Krasnitz, S. Gallinger, D. A. Tuveson, Organoid profiling identifies common responders to chemotherapy in pancreatic cancer. *Cancer Discov.* **8**, 1112–1129 (2018).
54. T. R. McKitrick, M. E. Ackerman, R. M. Anthony, C. S. Bennett, M. Demetriou, G. A. Hudalla, K. Ribbeck, S. Ruhl, C. M. Woo, L. Yang, S. J. Zost, R. L. Schnaar, T. L. Doering, The crossroads of glycoscience, infection, and immunology. *Front. Microbiol.* **12**, 731008 (2021).
55. S. R. Stowell, C. M. Arthur, R. McBride, O. Berger, N. Razi, J. Heimburg-Molinaro, L. C. Rodrigues, J.-P. Gourdine, A. J. Noll, S. von Gunten, D. F. Smith, Y. A. Knirel, J. C. Paulson, R. D. Cummings, Microbial glycan microarrays define key features of host-microbial interactions. *Nat. Chem. Biol.* **10**, 470–476 (2014).
56. S. S. Pinho, I. Alves, J. Gaifem, G. A. Rabinovich, Immune regulatory networks coordinated by glycans and glycan-binding proteins in autoimmunity and infection. *Cell. Mol. Immunol.* **20**, 1101–1113 (2023).
57. P. Saini, O. S. Adeniji, M. Abdel-Mohsen, Inhibitory Siglec-sialic acid interactions in balancing immunological activation and tolerance during viral infections. *EBioMedicine* **86**, 104354 (2022).
58. R. Kannagi, M. Izawa, T. Koike, K. Miyazaki, N. Kimura, Carbohydrate-mediated cell adhesion in cancer metastasis and angiogenesis. *Cancer Sci.* **95**, 377–384 (2004).
59. E. Rodriguez, K. Boelaars, K. Brown, R. J. E. Li, L. Kruijssen, S. C. M. Bruijns, T. van Ee, S. T. T. Schetters, M. H. W. Crommentuijn, J. C. van der Horst, N. C. T. van Grieken, S. J. van Vliet, G. Kazemier, E. Giovannetti, J. J. Garcia-Vallejo, Y. van Kooyk, Sialic acids in pancreatic cancer cells drive tumour-associated macrophage differentiation via the Siglec receptors Siglec-7 and Siglec-9. *Nat. Commun.* **12**, 1270 (2021).
60. S. A. Väyrynen, J. Zhang, C. Yuan, J. P. Väyrynen, A. D. Costa, H. Williams, V. Morales-Oyarvide, M. C. Lau, D. A. Rubinson, R. F. Dunne, M. M. Kozak, W. Wang, D. Agostini-Vulaj, M. G. Drage, L. Brais, E. Reilly, O. Rahma, T. Clancy, J. Wang, D. C. Linehan, A. J. Aguirre, C. S. Fuchs, L. M. Coussens, D. T. Chang, A. C. Koong, A. F. Hezel, S. Ogino, J. A. Nowak, B. M. Wolpin, Composition, spatial characteristics, and prognostic significance of myeloid cell infiltration in pancreatic cancer. *Clin. Cancer Res.* **27**, 1069–1081 (2021).
61. Y.-K. Huang, M. Wang, Y. Sun, N. D. Costanzo, C. Mitchell, A. Achuthan, J. A. Hamilton, R. A. Busuttil, A. Boussioutas, Macrophage spatial heterogeneity in gastric cancer defined by multiplex immunohistochemistry. *Nat. Commun.* **10**, 3928 (2019).
62. D. Freitas, D. Campos, J. Gomes, F. Pinto, J. A. Macedo, R. Matos, S. Mereiter, M. T. Pinto, A. Polónia, F. Gartner, A. Magalhães, C. A. Reis, O-glycans truncation modulates gastric cancer cell signaling and transcription leading to a more aggressive phenotype. *EBioMedicine* **40**, 349–362 (2019).
63. S. Chugh, S. Barkeer, S. Rachagani, R. K. Nimmakayala, N. Perumal, R. Pothuraju, P. Atri, S. Mahapatra, I. Thapa, G. A. Talmon, L. M. Smith, X. Yu, S. Neelamegham, J. Fu, L. Xia, M. P. Ponnusamy, S. K. Batra, Disruption of C1galt1 gene promotes development and metastasis of pancreatic adenocarcinomas in mice. *Gastroenterology* **155**, 1608–1624 (2018).
64. C. T. McDowell, X. Lu, A. S. Mehta, P. M. Angel, R. R. Drake, Applications and continued evolution of glycan imaging mass spectrometry. *Mass Spectrom. Rev.* **42**, 674–705 (2023).
65. P. M. Pour, M. M. Tempero, H. Takasaki, E. Uchida, Y. Takiyama, D. A. Burnett, Z. Steplewski, Expression of blood group-related antigens ABH, Lewis A, Lewis B, Lewis X, Lewis Y, and CA 19-9 in pancreatic cancer cells in comparison with the patient's blood group type. *Cancer Res.* **48**, 5422–5426 (1988).
66. J. Munkley, The glycosylation landscape of pancreatic cancer. *Oncol. Lett.* **17**, 2569–2575 (2019).
67. J. Rho, P. D. Lampe, High-throughput analysis of plasma hybrid markers for early detection of cancers. *Proteomes* **2**, 1–17 (2014).
68. B. B. Haab, Y. Huang, S. Balasenthil, K. Partyka, H. Tang, M. Anderson, P. Allen, A. Sasson, H. Zeh, K. Kaul, D. Kletter, S. Ge, M. Bern, R. Kwon, I. Blasutig, S. Srivastava, M. L. Frazier, S. Sen, M. A. Hollingsworth, J. A. Rinaudo, A. M. Killary, R. E. Brand, Definitive

characterization of CA 19-9 in resectable pancreatic cancer using a reference set of serum and plasma specimens. *PLOS ONE* **10**, e0139049 (2015).
69. Y. Jiang, M.-L. T. Lee, X. He, B. Rosner, J. Yan, Wilcoxon rank-based tests for clustered data with R package clusrank. *J. Stat. Softw.* **96**, 1–26 (2020).
70. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
71. Z. Klamer, B. Haab, Combined analysis of multiple glycan-array datasets: New explorations of protein-glycan interactions. *Anal. Chem.* **93**, 10925–10933 (2021).
72. K. Partyka, K. A. Maupin, R. E. Brand, B. B. Haab, Diverse monoclonal antibodies against the CA 19-9 antigen show variation in binding specificity with consequences for clinical interpretation. *Proteomics* **12**, 2212–2220 (2012).

Acknowledgments: We thank the members of the Biorepository and Pathology Core (Van Andel Institute) for assistance with preparation and processing of the tumor and plasma/serum specimens and K. Gallik of the Optical Imaging Core (Van Andel Institute) for assistance in acquisition of the multiplexed immunofluorescence images and for QuPath training and support. We thank A. Maitra (MD Anderson Cancer Center) for providing plasma specimens. Figures 1, 4, 5, and S7 were created using BioRender.com. **Funding:** This work was supported by the National Cancer Institute grants U01CA226158 (to B.H.), U01CA152653 (to B.H.), and U01CA200466 (to R.B.). **Author contributions:** Conceptualization: B.B., B.H., Z.K., D.M.B., C.S., Y.H., R.B., P.A., A.R., and J.B.M. Data curation: G.H., A.R., Z.K., C.S., P.B., Y.H., A.S., S.G., H.T., and B.S. Formal analysis: Z.K., A.R., Y.H., and G.H. Funding acquisition: B.B.H., P.A., and S.G. Investigation: C.G., B.S., B.B., C.S., P.B., P.A., S.G., A.R., J.B.M., and G.H. Methodology: B.B.H., Z.K., B.B., Y.H., R.B., P.A., A.R., J.B.M., and G.H. Project administration: B.B.H., P.B., and P.A. Resources: J.B.M., S.G., P.A., A.S., R.B., M.B., P.B., C.S., A.R., and G.H. Software: H.-L.T., Z.K., Y.H., and A.R. Supervision: B.B.H., P.A., and J.B.M. Validation: C.S., B.B.H., Y.H., S.G., H.-L.T., A.R., B.S., B.B., and G.H. Visualization: Z.K., B.B., A.R., and B.S. Writing—original draft: B.B.H., B.B., Z.K., and D.M.B. Writing—review and editing: D.M.B., B.B.H., Z.K., C.S., M.B., Y.H., R.B., P.A., S.G., B.S., and G.H. **Competing interests:** B.B.H. and R.B. are inventors on patent 8632983 held by the Van Andel Institute for the detection glycan biomarkers in blood assays for pancreatic cancer detection, and B.H. is an inventor of patent 10753936 held by the Van Andel Institute for detecting glycan biomarkers. All other authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The SignalFinder software is available for download (https://haablab.vai.org/?tabUrl=tools). The primary image data are available at the BioStudies database (www.ebi.ac.uk/biostudies/), accession ID S-BIAD1519. The TR4 anti-glycan antibody can be provided by M.B. pending scientific review and a completed material transfer agreement. Requests for the TR4 antibody should be submitted to M.B.

Submitted 6 September 2024
Accepted 30 January 2025
Published 7 March 2025
10.1126/sciadv.adt0029