A Mouse to Human Search for Plasma Proteome Changes Associated with Pancreatic Tumor Development

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story.

Accessibility

PLoS MEDICINE

A Mouse to Human Search for Plasma Proteome Changes Associated with Pancreatic Tumor Development

Vitor M. Faca[1], Kenneth S. Song[1], Hong Wang[1], Qing Zhang[1], Alexei L. Krasnoselsky[1], Lisa F. Newcomb[1], Ruben R. Plentz[2], Sushma Gurumurthy[2], Mark S. Redston[3], Sharon J. Pitteri[1], Sandra R. Pereira-Faca[1], Renee C. Ireton[1], Hiroyuki Katayama[1], Veronika Glukhova[1], Douglas Phanstiel[1], Dean E. Brenner[4], Michelle A. Anderson[4], David Misek[4], Nathalie Scholler[1], Nicole D. Urban[1], Matt J. Barnett[1], Cim Edelstein[1], Gary E. Goodman[1], Mark D. Thornquist[1], Martin W. McIntosh[1], Ronald A. DePinho[5,6], Nabeel Bardeesy[2], Samir M. Hanash[1]*

1 Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America, 2 Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, 3 Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, 4 University of Michigan, Comprehensive Cancer Center, Ann Arbor, Michigan, United States of America, 5 Center for Applied Cancer Science of the Belfer Institute for Innovative Cancer Science, Harvard Medical School, Boston, Massachusetts, United States of America, 6 Departments of Medical Oncology, Medicine and Genetics, Dana Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts, United States of America

Abbreviations: AUC, area under the curve; CA, carbohydrate antigen; CARET, Carotene and Retinol Efficacy Trial; FDR, false discovery rate; GEM, genetically engineered mouse; ICAM1, intercellular adhesion molecule 1; IGFBP4, insulin-like growth factor binding protein 4; IHC, immunohistochemical; LC–MS/MS, liquid chromatography–tandem mass spectrometry; LCN2, neutrophil gelatinase-associated lipocalin; PanIN, pancreatic intraepithelial neoplasia; PDAC, pancreatic ductal adenocarcinoma; PTPRG, protein tyrosine phosphatase receptor type gamma; REG1A, lithostathine 1; REG3, regenerating islet-derived protein 3; ROC, receiver operating characteristic; TIMP1, tissue inhibitor of metalloproteinase 1; TNC, tenascin C; TNFRSF1A, tumor necrosis factor receptor superfamily member 1a precursor; WFDC2, WAP four-disulfide core domain 2

* To whom correspondence should be addressed. E-mail: shanash@fhcrc.org

ABSTRACT

Background

The complexity and heterogeneity of the human plasma proteome have presented significant challenges in the identification of protein changes associated with tumor development. Refined genetically engineered mouse (GEM) models of human cancer have been shown to faithfully recapitulate the molecular, biological, and clinical features of human disease. Here, we sought to exploit the merits of a well-characterized GEM model of pancreatic cancer to determine whether proteomics technologies allow identification of protein changes associated with tumor development and whether such changes are relevant to human pancreatic cancer.

Methods and Findings

Plasma was sampled from mice at early and advanced stages of tumor development and from matched controls. Using a proteomic approach based on extensive protein fractionation, we confidently identified 1,442 proteins that were distributed across seven orders of magnitude of abundance in plasma. Analysis of proteins chosen on the basis of increased levels in plasma from tumor-bearing mice and corroborating protein or RNA expression in tissue documented concordance in the blood from 30 newly diagnosed patients with pancreatic cancer relative to 30 control specimens. A panel of five proteins selected on the basis of their increased level at an early stage of tumor development in the mouse was tested in a blinded study in 26 humans from the CARET (Carotene and Retinol Efficacy Trial) cohort. The panel discriminated pancreatic cancer cases from matched controls in blood specimens obtained between 7 and 13 mo prior to the development of symptoms and clinical diagnosis of pancreatic cancer.

Conclusions

Our findings indicate that GEM models of cancer, in combination with in-depth proteomic analysis, provide a useful strategy to identify candidate markers applicable to human cancer with potential utility for early detection.

The Editors' Summary of this article follows the references.

Introduction

A major goal of the cancer biomarker field is the development of noninvasive tests that allow early cancer detection. Blood constituents, notably plasma proteins, reflect diverse physiologic or pathologic states. The ease with which this compartment can be sampled makes it a logical choice for screening applications to detect cancer at an early stage. However, the vast dynamic range of protein abundance in plasma and the likely occurrence of tumor-derived proteins in the lower range of protein abundance represent major challenges in the application of proteomic-based strategies for cancer biomarker identification [1,2]. Recent experience in comprehensive profiling of plasma proteins indicates that low-abundance proteins may be identified with high confidence following extensive plasma fractionation and with the use of high-resolution mass spectrometry [3,4].

Genomic analyses of human and mouse cancers have revealed significant concordance in chromosomal aberrations and expression profiles, establishing cross-species analyses as a highly effective filter in the identification of genes and loci embedded within complex cancer genomes [5–8]. Genetically engineered mouse (GEM) models afford defined stages of tumor development, homogenized breeding and environmental conditions, and standardized blood sampling thereby reducing biological and nonbiological heterogeneity. The concept that plasma from GEM models of cancer contains tumor-derived proteins that may be relevant as candidate markers for human cancer is attractive as suggested by SELDI (surface enhanced laser desorption/ionization) scanning technology, but it remains untested as no markers demonstrated to be applicable to human cancer have been identified using such models and methods [9].

In this study, we focused our efforts on pancreatic ductal adenocarcinoma (PDAC)—a highly lethal cancer characterized by activating mutations of the *Kras* oncogene and inactivation of the *Ink4a* and *Arf-p53* tumor suppressor pathways in the great majority of cases [10]. *Kras* activation is thought to initiate focal lesions in the pancreatic ducts, known as pancreatic intraepithelial neoplasias (PanINs), which undergo graded histological progression to PDAC in association with subsequent *Ink4a* and *Arf-p53* inactivation [11,12]. The recent generation of mice harboring these signature genetic mutations has yielded models that closely recapitulate the histopathogenesis of the human disease with $Kras^{G12D}$ initiating focal PanINs that rapidly undergo multistage progression in conjunction with *Ink4a/Arf* or *p53* mutations, resulting in invasive PDAC. Importantly, these models show broadly conserved tumor biology and molecular circuitry similar to human PDAC. The tumors exhibit a proliferative stroma (desmoplasia) and frequent metastases, express pancreatic ductal markers (CK-19) and apical mucins (e.g., Muc1, Muc5AC), show activation of developmental signaling pathways (Hedgehog, Notch, EGFR), and harbor syntenic genomic alterations to human PDAC [9,13–15].

We have applied here an intensive quantitative proteomic analysis strategy to plasmas that were sampled from this pancreatic cancer mouse model at early stage, representing PanIN, and at advanced stage of tumor development, representing PDAC, and from corresponding matched controls. With this approach, we sought to explore the merits of this well-characterized GEM model of pancreatic cancer to determine whether our proteomics technology allows identification of protein changes associated with tumor development and whether such changes are relevant to human pancreatic cancer.

Materials and Methods

Mice and Plasma Pooling

The mice for proteomics analysis were obtained by breeding Pdx1-Cre *Ink4a/Arf*$^{lox/lox}$ and *Kras*^{G12D} *Ink4a/Arf*$^{lox/lox}$ mice [13]. All mice were bred five generations onto an FVN/n genetic background. Experimental Pdx1-Cre *Kras*^{G12D} *Ink4a/Arf*$^{lox/lox}$ mice and control *Kras*^{G12D} *Ink4a/Arf*$^{lox/lox}$ and Pdx1-Cre *Ink4a/Arf*$^{lox/lox}$ mice were euthanized at age 5.5 or 7 wk (Figure 1). Lethal comas were induced by injecting mice IP with a 0.6–0.8 ml 5% Avertin (2,2,2-Tribromoethanol, Sigma-Aldrich, part number T4,840–2). Blood was obtained by cardiac puncture using a 1-ml syringe with 22-gauge needle. Blood was placed in K$_3$EDTA coated tubes (Fisher) and centrifuged at 4 °C for 5 min at 3,000 rpm. The supernatant (plasma) was removed and frozen in 100 µl aliquots on dry ice and stored at −80 °C. In all cases, the mice were subjected to autopsy, and the pancreas was fixed for histological analysis. Mice were excluded from the study if they exhibited extra-pancreatic pathology as is observed in a subset of Pdx1-Cre *Kras*^{G12D} mice [9]. Pooling of samples was based on age as well as the extent of the disease based on histological examination. For early stage PanIN pool (PanIN-1 to PanIN-3 lesions) plasma analysis, median age of the mice was 5.5 wk, while for PDAC plasma analysis median age was 7 wk. Approximately one-third of the *Kras Ink4a/Arf* mice present with the most common pathology observed in human cases—glandular. Thus, in our selection of mice with PDAC, we only used the corresponding plasma if the tumor areas were almost exclusively glandular (i.e., less than ~5% nonglandular pathologies). Age matched controls were used for both PanIN and PDAC. All mice were male.

Sample Preparation

PDAC, PanIN, and respective control plasma pools obtained from seven to eight individual mice (1 ml of each pool) were individually immunodepleted of the top three most abundant proteins (albumin, IgG, and transferrin) using a Ms-3 column (4.6 × 250 mm; Agilent). Briefly, columns were equilibrated with buffer A at 0.5 ml/min for 13 min, and aliquots of 75 µl of the pooled sera were injected after filtration through a 0.22-µm syringe filter. The flow-through fractions were collected for 10 min at a flow rate of buffer A of 0.5 ml/min, combined and stored at −80 °C until use. The column bound material was recovered by elution for 8 min with buffer B at 1 ml/min. Subsequently, immunodepleted samples were concentrated using Centricon YM-3 devices (Millipore) and rediluted in 8 M urea, 30 mM Tris (pH 8.5), 0.5% OG (octyl-beta-d-glucopyranoside, Roche). Samples were reduced with DTT in 50 µL of 2 M Tris-HCl (pH 8.5) (0.66 mg DTT/mg protein), and isotopic labeling of intact proteins in cysteine residues were performed with acrylamide. Normal control samples received the light acrylamide isotope (D0 acrylamide) (>99.5% purity, Fluka), and PDAC and PanIN cancer samples received the heavy 2,3,3′-D3-acrylamide isotope (D3 acrylamide) (>98% purity, Cambridge Isotope Laboratories). Alkylation with acrylamide was performed for 1 h at room temperature by adding to the
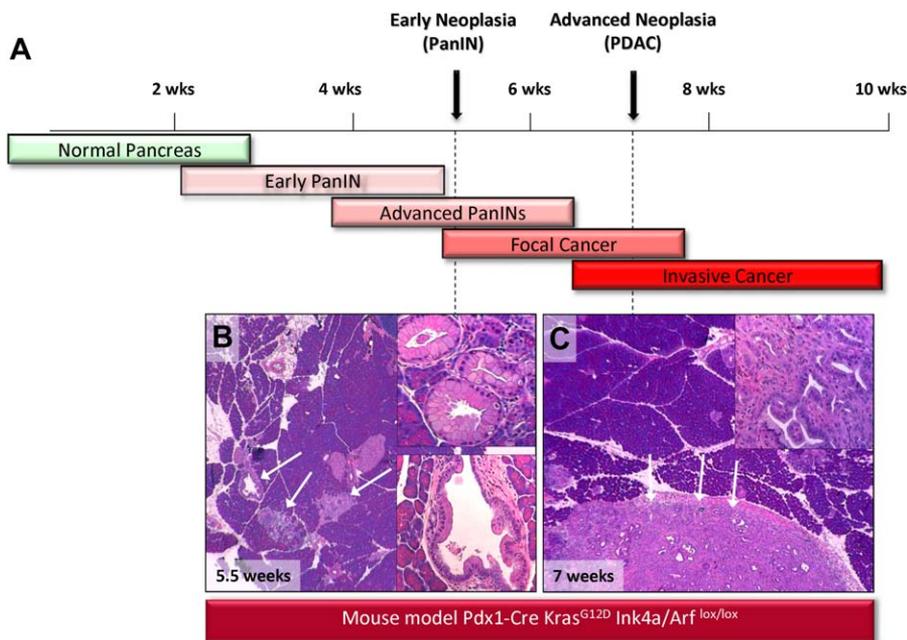
Figure 1. PanIN and PDAC Mice for Plasma Proteomic Analyses

(A) *Pdx1-Cre Ink4a/Arf* $^{lox/lox}$ and *Kras* ^{G12D} *Ink4a/Arf* $^{lox/lox}$ mice at average of 5.5 and 7 wk of age representing PanIN and PDAC lesions, respectively, were selected on the basis of histological analysis of tumor for this study.

(B and C) show representative histology for PanIN and PDAC stages. Plasma was pooled from each disease group and each corresponding age and sex-matched controls to yield 1 ml per phenotype for proteomic analysis.

doi:10.1371/journal.pmed.0050123.g001

protein solution 7.1 mg D0-acrylamide or 7.4 mg D3-acrylamide per milligram protein, diluted in a small volume of 2 M Tris-HCl (pH 8.5) [16].

Protein Fractionation

The two sets of samples (PDAC × control and PanIN × control) were processed in the same identical way. The 2-D protein fractionation has been performed on the basis of the Intact-Protein Analysis System (IPAS) approach [3,17,18], with some modifications. The workflow is summarized in Figure 2. Briefly, after isotopic labeling, the cancer plasma pool and normal pool were mixed, diluted to 10 ml with 20 mM Tris in 6% isopropanol, 4 M urea (pH 8.5), and immediately injected in a Mono-Q 10/100 column (Amersham Biosciences) for the anion-exchange chromatography, the first dimension of the protein fractionation. The buffer system consisted of solvent A (20 mM Tris in 6% isopropanol, 4 M urea [pH 8.5]) and solvent B (20 mM Tris in 6% isopropanol, 4 M urea, 1 M NaCl [pH 8.5]). The separation was performed at 4.0 ml/min in a gradient of 0% to 35% solvent B in 44 min; 35% to 50% solvent B in 3 min; 50% to 100% solvent B in 5 min; and 100% solvent B for an additional 5 min. A total of 12 pools were collected and run individually in reversed-phase chromatography, the second dimension of the process. The reversed-phase fractionation was carried out in a Poros R2 column (4.6 × 50 mm, Applied Biosystems) using TFA/Acetonitrile as buffer system (solvent A: 95% H2O, 5% Acetonitrile, 0.1% TFA and solvent B: 90% Acetonitrile, 10% H2O, 0.1% TFA) at 2.7 ml/min. The gradient used was 5% solvent A until absorbance reached base line (desalting step) and then 5%–50% solvent B in 18 min; 50%–80% solvent B in 7 min; and 80%–95% solvent B in 2 min. Sixty fractions of 900 μl were collected during the run, corresponding to a total

of 720 fractions. Aliquots of 200 μl of each fraction, correspondent approximately of 20 μg of protein, were separated for mass-spectrometry shotgun analysis.

Mass Spectrometry Analysis

For protein identification we performed in-solution trypsin digestion with the lyophilized aliquots of the 720 individual fractions. Individual digested fractions 4 to 60 from each reversed-phase run were pooled in 13 pools, corresponding to a total of 156 fractions for analysis from each PDAC and PanIN experiments. Digests were analyzed in a LTQ-FT mass spectrometer (Thermo-Finnigan) coupled to a nano-Aquity nanoflow chromatography system (Waters). The liquid chromatography separation was performed in a 25-cm column (Picofrit 75 μm ID, New Objectives, in-house-packed with MagicC18 resin) using a 90-min linear gradient from 5% to 40% of acetonitrile in 0.1% formic acid at 250 nl/min. The spectra were acquired in a data-dependent mode in m/z range of 400 to 1,800, with selection of the five most abundant +2 or +3 ions of each MS spectrum for MS/MS analysis. Mass spectrometer parameters were: capillary voltage of 2.1 KV, capillary temperature of 200 °C, resolution of 100,000, and FT target value of 2,000,000.

Protein Identification

The acquired data were automatically processed by the Computational Proteomics Analysis System (CPAS) [19]. Searches were performed considering cysteine alkylation with the light form of acrylamide as a fixed modification and heavy form of acrylamide (+3.01884) as a variable modification. For the identification of proteins with false discovery rate (FDR) < 1%, LC/MS/MS spectra of PDAC and PanIN samples were subjected to tryptic and semi-tryptic searches
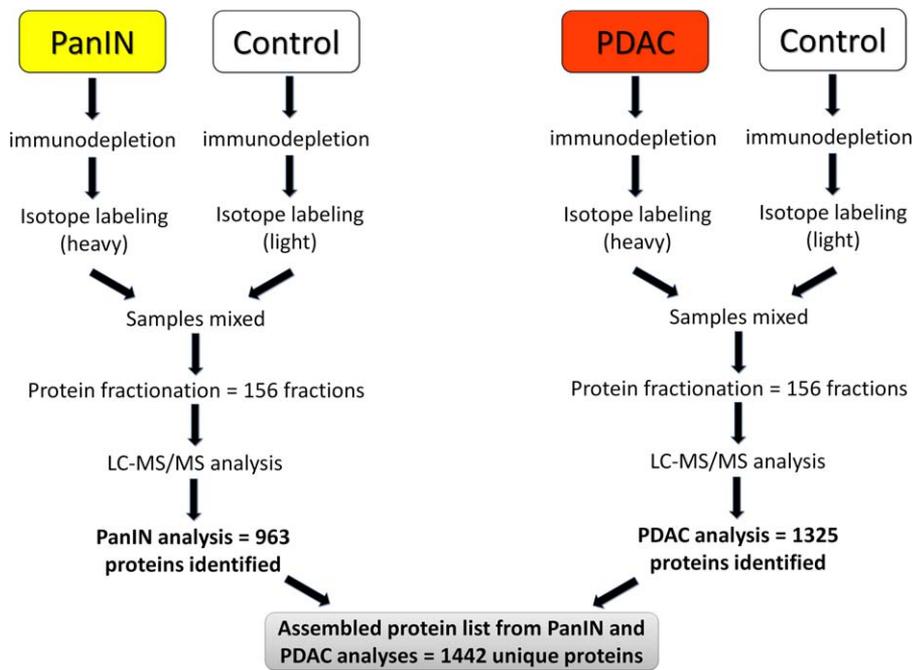
Figure 2. Schematic of Mouse Plasma Proteomic Analysis

Pools of plasma from PanIN and PDAC mice along with corresponding controls were similarly processed and combined after differential isotopic labeling. Subsequent protein fractionation involved anion-exchange and reversed-phase chromatography. Individual fractions were analyzed by LC–MS/MS after in-solution digestion. Data were processed using Computational Proteomics Analysis System (CPAS).

doi:10.1371/journal.pmed.0050123.g002

against a database consisting of forward and reversed mouse IPI databases released in 01/2006 (v.3.12) using X!Tandem [20]. The database search results were then analyzed by PeptideProphet [21] and ProteinProphet [22] programs. Our high confidence list of identifications retained proteins with ProteinProphet scores $\geq$ 0.95 and two or more peptides per protein. For PDAC, 18,409 unique peptides corresponding to 1,040 proteins were identified in forward sequence, whereas only eight peptides corresponding to four proteins were identified in reversed sequence, resulting in a false positive identification rate for peptides of 8/18,409 or 0.04%, and proteins of 4/1,040 or 0.4%. For PanIN, 16,319 unique peptides, corresponding to 559 proteins were identified in forward sequence, whereas only five peptides, corresponding to two proteins were identified in reversed sequence, and this resulted in a false positive identification rate for peptides of 5/16,319 or 0.03%, and proteins of 2/559 or 0.4%. A secondary list of protein identifications with less than 5% FDR consisted of tryptic searches using the same algorithm and databank, but only proteins with ProteinProphet score > 0.7 and PeptideProphet score > 0.2 were retained. The result with <1% FDR searches were later appended with data from the 5% FDR searches on the basis of external cross-correlated biological information from different sources, such as tissue specificity or mRNA expression in pancreatic cancer. The number of MS events (spectral counts) was obtained for all the proteins with less than 5% FDR (including less than 1% FDR) from tryptic searches only.

Quantitative Analysis of Acrylamide Isotopes

The quantitative approach consisted of differential labeling of peptides containing cysteine with acrylamide isotopes (heavy or light) [16]. Quantitative information was extracted using a script designated "Q3" that was developed in-house to obtain the relative quantification for each pair of peptides identified by MS/MS that contains cysteine residues [16]. Only peptides with a minimum of 0.75 PeptideProphet score and mass deviation inferior to 20 ppm were considered. Peptide isotopic ratios were plotted in logarithmic scale in a histogram and the median of the distribution was centered at zero. This normalization approach was chosen since the great majority of proteins were not expected to be dysregulated in cases compared to controls (Figure S1). All normalized peptide ratios for a specific protein were averaged to compute an overall protein ratio. Proteins with quantitative information presented as "cancer only," only had detected peptides labeled with the heavy form of acrylamide. All peptide and protein ratios were calculated in logarithmic scale, but reported in linear scale. Statistical significance of the protein quantitative information was obtained via two procedures: (i) for those proteins with multiple peptides quantified, a p-value for the mean log-ratio, which has mean zero under the null hypothesis, was calculated using one-sample t-test; (ii) for proteins with a single paired MS event, the probability for the ratio was extrapolated from the distribution of ratios in a control-control experiment whereby the same sample was labeled with heavy and light acrylamide (Figure S1).

mRNA Analysis of Pancreatic Tissue

Total pancreas RNA was isolated from wild-type FVB/n mice using the Trizol reagent protocol (Invitrogen) with the slight modifications; in brief, freshly harvested pancreas was homogenized in 15 ml Trizol, centrifuged, and the aqueous layer was extracted with chloroform, and finally isopropanol precipitation was performed by adding 0.5 volumes high salt buffer (0.8 M NaCitrate/1.2 M NaCl) and 0.5 volume

isopropanol. A second round of purification was performed using the RNAeasy kit (Qiagen). Total RNA from PDAC arising in Pdx1-Cre LSL-KrasG12D Ink4a/Arflox/lox mice was extracted using the Trizol Reagent and then by RNAeasy using the standard protocols. Expression profiling of normal pancreas ($n = 2$ specimens) and PDAC RNA ($n = 4$ specimens) was performed on Affymetrix 430 A2.0 microarrays.

Histology and Immunohistochemistry

Tissue sections were fixed in 10% formalin. Paraffin-embedded 6–8-µm sections of the pancreas were used for histological and immunohistochemical (IHC) studies. The following primary antibodies were used: tumor necrosis factor receptor superfamily member 1a precursor (TNFRSF1A) (Abcam); anti-human TNF RI/TNFRSF1A antibody (R&D Systems); TIMP-1 Ab-2 mouse mAb (Lab Vision Corporation); monoclonal anti-human TIMP-1 antibody (R&D Systems); protein tyrosine phosphatase receptor type gamma (PTPRG) polyclonal antibody (Orbigen); Human/Mouse/Rat Tenascin C Mab (R&D Systems); Mouse ALCAM Biotinylated Affinity Purified Pab (R&D Systems); and CD166 (ALCAM) mouse monoclonal antibody (Novocastra). Sections for immunohistochemistry were deparaffinized with xylol and rehydrated. After the washing steps with PBS, antigen retrieval was performed by microwave heating the slides for 10 min in Antigen Unmasking Solution (Vector). The slides were then washed in PBS and incubated for 10 min in 1% H_2O_2, rinsed with PBS, and incubated 1 h in blocking solution (5% normal serum + 0.3% Triton X-100). Hybridization with the primary antibody was carried out overnight at 4 °C. After PBS rinse, secondary antibody (Jackson Immunoresearch) was incubated for 1 h. The manufacturer's protocols were used for ABC and DAB substrates (Vector); slides were counterstained with hematoxylin (Vector) and dehydrated in 40%, 70%, 90%, and 100% ethanol.

ELISA

Mouse ALCAM, tissue inhibitor of metalloproteinase 1 (TIMP1), and intercellular adhesion molecule 1 (ICAM1) (R&D Systems) measurements were performed according to manufacturer's protocol in aliquots of the same mice used for the plasma proteomic analysis. For ALCAM, mouse plasma dilution was 1:8. Human ALCAM, insulin-like growth factor binding protein 4 (IGFBP4), sTNFRSF1, TIMP1 (R&D Systems), ICAM1 (Biosource), neutrophil gelatinase-associated lipocalin (LCN2) (Antibody Shop), lithostathine 1 (REG1A) (Biovendor), and regenerating islet-derived protein 3 (REG3) (Pancrepap) measurements were performed on sera from PDAC, matched controls, and pancreatitis using commercially available ELISAs according to the manufacturer's protocol. HE4 measurements were performed according to Scholler et al. [23]. Additionally, PDAC and matched controls sera were also assayed for carbohydrate antigen 19.9 (CA 19–9) (Alpha Diagnostic International). All sera samples and standards were run in duplicate with absorbance measured on the SpectraMax Plus 384 and results calculated with SoftMax Pro v4.7.1 (Molecular Devices).

Statistical Analysis of ELISA Data

Prior to statistical analysis, all candidate markers had their protein concentration standardized on the basis of the control group concentration mean. In that way all candidate marker concentrations have mean 0 and variance 1 in the control group. In short, if mu0 and sd0 are the mean and standard deviation of a candidate marker, their standardized concentration (Y′) will be Y′ = (y − mu0)/sd0. This method facilitates cross-candidate marker comparison and places all markers on the same scale [24,25]. p-Values for individual markers were computed using the nonparametric Wilcoxon rank-sum test. To avoid over-fitting issues, composite markers summarizing a panel were generated using a predefined combination rule that considers the panel positive if any individual marker is positive (e.g., exceeds a threshold on the standardized scale). p-Values that measure whether the AUC of the composite markers are statistically different from CA19.9 were computed using a method described by DeLong et al. [26].

Human Samples

Newly diagnosed serum samples from patients were obtained at the time of diagnosis following informed consent using IRB-approved guidelines from the University of Michigan. A total of 30 serum samples were obtained from patients with a confirmed diagnosis of pancreatic adenocarcinoma who were seen in the Multidisciplinary Pancreatic Tumor Clinic at the University of Michigan Comprehensive Cancer Center. Anonymous serum samples from the pancreatic cancer patients were randomly selected from a clinic population that consists of 15% of individuals presenting with early stage (i.e., stage 1/2) disease and 85% presenting with advanced stage (i.e., stage 3/4). The information on individual characteristics is presented in Table S1. Inclusion criteria for the study consisted of confirmed diagnosis of pancreatic cancer, the ability to provide written informed consent, and the ability to provide 40 ml of blood. Exclusion criteria included chemotherapy or radiation therapy prior to blood draw and a diagnosis of other malignancies within 5 y from the time of blood draw. Sera were also obtained from 15 patients with chronic pancreatitis who were seen in the Gastroenterology Clinic at University of Michigan Medical Center and from 20 control healthy individuals collected at the University of Michigan under the auspices of the Early Detection Research Network (EDRN). The mean age of the tumor group was 65 y and of the chronic pancreatitis group was 54 y. Individuals from whom control sera were obtained were age and sex matched to the tumor group. All of chronic pancreatitis sera were collected in an elective setting in the clinic in the absence of an acute flare. All blood and sera were collected and processed using the same standardized protocol. Blood samples were maintained at room temperature for 30–60 min to allow the clot to form and then centrifuged at 1,300 × g at 4 °C for 20 min. The serum was removed, transferred to a polypropylene capped tube in 1 ml aliquots, and frozen. The frozen samples were stored at −70 °C until assayed. All serum samples were labeled with a unique identifier. None of the samples were thawed more than twice before analysis.

To address the relevance of proteins observed up-regulated in the PanIN stage mouse model plasma, we submitted a proposal to the Carotene and Retinol Efficacy Trial (CARET), a cohort study that involved 18,314 individuals with increased cancer risk, to do a blinded validation study of our relevant proteins. CARET identified all individuals (13) in this cohort from whom blood was collected approximately a year prior to

the diagnosis of pancreatic cancer (actual mean = 10 mo), at a time when they were completely asymptomatic, as well as matched controls that were not diagnosed with cancer over a 4-y follow-up period, irrespective of their state of general health otherwise. The information on individual characteristics is presented in Table S2.

Results

Proteomic Analysis of Mouse Plasma

Plasma obtained from PDAC-prone mice engineered with activated Kras and Ink4a/Arf deficiency [13] was subjected to proteomic analysis. The study was designed to test directly whether current proteomics technologies allow for quantitative analysis and identification of protein changes associated with tumor development in the mouse and whether such changes have relevance to human tumors.

Mice harboring *Pdx1-Cre Kras*^{G12D} *Ink4a/Arf*$^{lox/lox}$ mutations exhibit stereotypical neoplastic progression from pancreatic cancer precursor lesions (PanINs) present at ~2 wk of age to advanced PDAC by 6 to 10 wk of age [13]. A plasma pooling strategy was applied for in-depth proteomic analysis. Blood was obtained from mice at the PanIN stage and at the PDAC stage (at 5.5 and 7 wk, respectively) and from age and sex matched controls, thus constituting four pools of plasma (Figure 1). To guarantee a good homogeneity among pooled plasma samples, the tumor stage was confirmed for individual mice by histopathology prior to pooling. For quantitative proteome analysis, we applied differential isotopic labeling to each tumor pool and its matched control [16], followed by extensive fractionation of intact proteins [3]. The experimental workflow is presented in Figure 2.

Each experiment generated 156 plasma fractions on the basis of anion-exchange and reversed-phase chromatography, which were analyzed separately by liquid chromatography–tandem mass spectrometry (LC–MS/MS) following tryptic digestion. Some 2,800,000 mass spectra were produced and analyzed in this study. Collectively, the PanIN and PDAC experiments resulted in a primary list of 1,095 unique high confidence proteins with <1% FDR on the basis of reverse-database searches (Table S3 presents the full list of protein identifications). To this primary list, we appended 347 additional proteins with <5% FDR (Table S4). The latter proteins had corresponding mRNA expression in pancreas tissue >2-fold compared to the mean of 61 mouse tissue expression surveys from published data [27] and/or mRNA expression in pancreatic cancer >2-fold compared to normal tissue, in mouse (this study) or human (prior study [28]).

On the basis of UniProt keywords, 25% of identified proteins in the list of 1,442 proteins contained a signal peptide for secretion, and 20% were annotated as glycoproteins. Of note, the list contained a relatively large percentage (9%) of membrane proteins on the basis of Gene Ontology cellular component annotation. Peptides for several membrane proteins identified were derived exclusively from the extracellular domain. Epidermal growth factor receptor, for example, was detected in several fractions with peptides spanning amino acids 25 to 647 representing the extracellular N-terminal domain. These results are consistent with shedding of extracellular domains into the circulation [29].

To estimate the concentration range of mouse plasma proteins identified, we correlated spectral counting data

(number of MS2 events/protein) [30] to known concentrations of proteins in plasma (http://www.rulesbasedmedicine.com/). We observed a significant correlation between spectral counts for a given protein and its plasma protein concentration (R^2 = 0.84) (Figure 3A). From this analysis, we estimated that our proteomic approach allowed for identification of plasma proteins across seven orders of magnitude and detection of some proteins in mouse plasma at concentrations as low as 1 ng/ml. In addition, the number of proteins identified was greater at lower predicted plasma concentrations on the basis of spectral counts (Figure 3B), indicating substantial depth of analysis achieved with extensive protein fractionation.

The majority of medium to high abundance proteins were detected in both PanIN and PDAC experiments, while most differences in protein identifications between the two experiments represented lower abundance proteins (Figure S2A). Likewise, in duplicate LC–MS/MS analysis of the same fractions, most differences in protein identifications observed represented lower abundance proteins (Figure S2B). Similar experiments in which independent replicates of samples were analyzed resulted in 60% of protein sampling/identification in both experiments [3]. These differences in protein identifications between the two experiments are largely attributed to mass spectrometry limitations in dynamic range and speed, specifically when analyzing complex samples such as plasma. In addition to mass spectrometry limitations, some of the differences observed between the two experiments may result from occurrence of some proteins at a higher level of abundance at the PDAC stage compared to PanIN. Importantly, since in each experiment (PDAC and PanIN) cancer and respective control samples were analyzed together after isotopic labeling followed by mixing, methodological variations related to fractionation and sample processing were minimized.

Tumor Related Changes in Mouse Plasma

We used acrylamide isotopic labeling of cysteine residues to obtain relative quantitative information between disease and control samples. This labeling approach is chemically very efficient as evidenced by lack of unlabeled cysteines in searching mass spectra [16]. Additionally, this labeling chemistry is fully compatible with the intact protein approach, without significantly affecting protein physical-chemical characteristics. In duplicate experiments performed with independent replicates of samples, there were no proteins that showed quantitative inconsistencies (up-regulated in one experiment and down-regulated in the other) (unpublished data). Among the 621 quantified proteins, 165 were found to be up-regulated (≥1.5, p < 0.05) in cancer samples (PDAC or PanIN or both) compared to controls (Table S5).

A significant proportion of plasma proteins is synthesized in the liver and may be affected as part of the host response. To distinguish between such classical plasma proteins from proteins that may be derived from the pancreas in our dataset, we cross-referenced the 1,442 proteins identified in our analyses with published proteome profiles of mouse liver tissue [31,32]. Approximately 38% of the 1,442 proteins were identified in mouse liver tissue, consisting mostly of relatively abundant plasma proteins. Sixty-seven of these proteins showed increased levels with tumor development in the mouse (Table S5). In contrast, proteins estimated to be of low
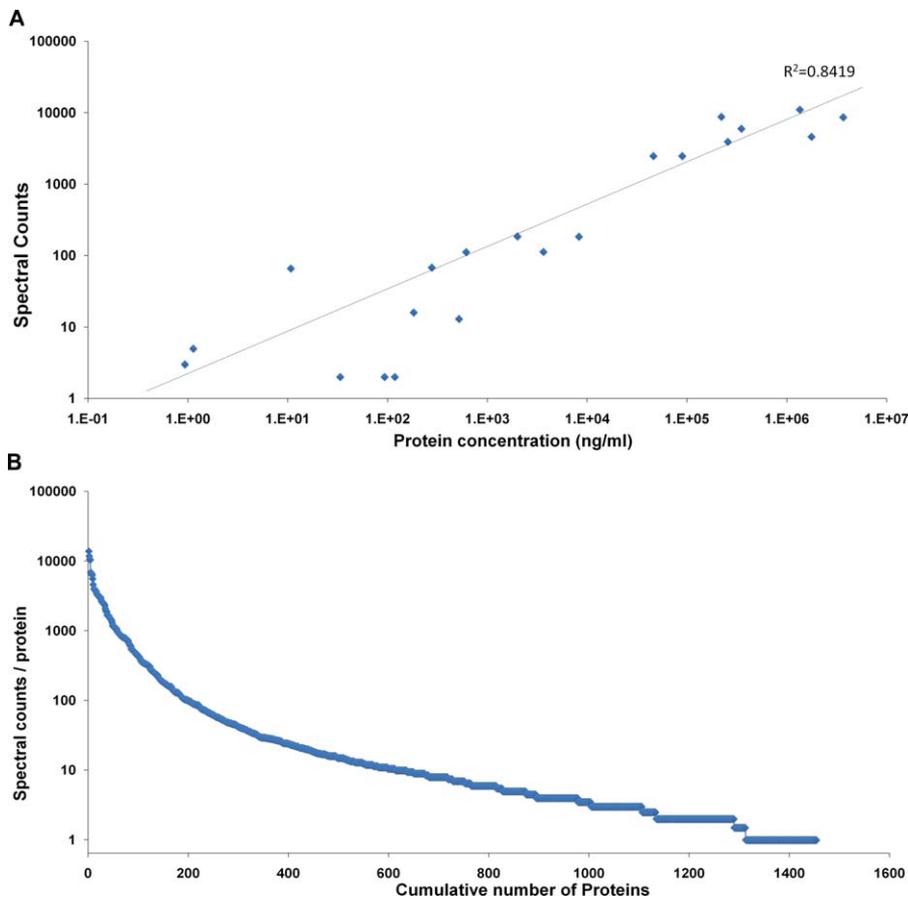
Figure 3. Identification of Low Abundance Proteins in Mouse Plasma

(A) Spectral counts (number of MS2-events acquired per protein) in the experiment performed for early stage pancreatic cancer mouse plasma protein (PanIN) were correlated with protein plasma concentration reported by Rules-Based Medicine (http://www.rulesbasedmedicine.com/case3/Table3.htm). The 21 proteins used for this estimation were: Serpina1b, Adipoq, A2m, Apoa1, Apoc3, Apoh, B2m, C3, Ceacam1, Crp, Fabp1, F7, Ftl1, Fgb, Hp, Icam1, Igf1, Mb, Serbp1, Timp1, Vcam1. As an approximation, we estimated the protein concentration with the correlation (log spectral counts = [0.623 × log protein concentration] + 0.0625).

(B) Taking into consideration the correlation of spectral counts and protein concentration, we observed an inverse relationship between the total number of proteins identified and their abundance (number of MS2/proteins).

doi:10.1371/journal.pmed.0050123.g003

abundance in the protein list had a much greater representation of pancreatic proteins relative to liver proteins on the basis of tissue protein and/or mRNA data (Figure S3).

The following criteria were applied to select a subset of proteins potentially relevant to pancreatic cancer: (i) mean protein ratio in neoplasm/normal plasma ≥ 1.5 ($p < 0.05$) in PDAC and PanIN on the basis of isotopic labeling ratios, and/or occurrence of isotope-labeled peptides in cancer samples but not in controls; (ii) not known to represent acute-phase reactants, complement or coagulation proteins according to Ingenuity Pathway Analysis annotation (Ingenuity Systems) (Table S5); and (iii) mouse protein has a corresponding ortholog gene in human. Also included in this list were proteins that were similarly elevated in either PDAC or PanIN and that had evidence of increased expression of corresponding genes in pancreatic cancer for mouse (data obtained in this study) and for human [28]. These criteria resulted in subset of 45 proteins of potential interest from the set of 165 up-regulated proteins (Table 1).

To further support our findings we measured protein levels in mouse pancreatic tissue and in mouse plasma for a subset of up-regulated proteins. These proteins were selected on the

basis of their potential relationship with pancreatic cancer and the availability of antibodies and ELISA kits with the requisite specificity. IHC analysis was done for CD166 antigen precursor (ALCAM), receptor-type tyrosine-protein PTPRG, TIMP1, and tenascin C (TNC). All tested proteins demonstrated strong IHC staining in mouse PanIN and pancreatic cancer tissue sections (Figure 4). Circulating protein levels of ALCAM, ICAM1, and TIMP1 in the same mouse plasma used in the proteomic approach were measured by ELISA (Figure 5). ALCAM, ICAM1, and TIMP1 had significantly higher levels in PDAC mice plasmas. TIMP1 was significantly elevated in PanIN plasma samples as well.

Relevance of Mouse Findings to Human Pancreatic Cancer

The relevance to human pancreatic cancer of proteins upregulated in mouse plasma with tumor development was investigated using human tissue and/or blood samples. Immunohistochemistry was performed for PTPRG, TNFRSF1A, and TNC, all of which showed positive IHC staining in human pancreatic cancer (Figure 3). Data for ALCAM, ICAM1, LCN2, TNFRSF1A, TIMP1, REG1A, REG3, WFDC2 (whey-acidic protein [WAP] four-disulfide core

Table 1. Proteins Potentially Relevant to Pancreatic Cancer Revealed by the Proteomic Analysis of Mouse Model Plasma

Criteria for Candidate Selection	Name	Protein Quantification*		mRNA Expression			Previously Analyzed				
		PDAC Cancer/ Normal	PanIN Cancer/ Normal	Human		Mouse Cancer/ Normal	Tissue	Plasma	Proteomics[b]		
				Cancer/ Normal	Pancreatitis/ Normal				Tissue	Pancreatic Juice	Plasma
Increased levels in PDAC and PanIN	**CD5L**	2.7	2.1	0.6	0.7	0.3					
	CTRB1	6.1	9.1	0.3	0.3	0.3			↓ [52]	↑ [53], ■ [54,55]	
	IL1RAP	Cancer only	Cancer only	6.4	4.9	0.4					
	LCN2[a]	6.9	Cancer only	13.4	3.8	39.8	IHC [46]			■ [54]	
	LRG1	2.9	2.0	—	—	1.3				■ [54]	↑ [56]
	PRG4	10.0	2.5	1.5	1.6	0.7					
	REG1A[a]	2.3	3.9	0.3	0.3	0.4		ELISA [57,58]	↓ [52], ↑[59]	↑ [53], ■ [55]]	
	REG3[a]	5.9	2.8	—	—	0.7		ELISA [58]			
	SLPI	1.9	2.3	3.9	1.5	1.9			↓ [60]		
	SYCN	4.2	5.0	—	—	0.1			↓ [61]		
	TIMP1[a]	Cancer only	(1.7)[c]	4.1	3.4	22.0	IHC [62]	ELISA [51]			
Increased levels in PDAC and increased expression of corresponding genes	**ALCAM**[a]	2.8	(0.9)[c]	1.1	1.1	2.8			↑ [52]		
	COL18A1	Cancer only	—	1.2	1.1	13.7					
	COL15A1	1.8	—	1.2	1.4	5.1					
	CTGF	2.0	0.8	1.2	3.9	89.6			↓ [60]		
	CTSS	2.4	—	3.4	2.8	16.8					
	CXCL16	2.8	—	—	—	3.2					
	FBLN2	2.6	—	4.1	8.2	5.7					
	FSTL1	6.5	1.1	2.3	3.4	12.7					
	HTRA1	Cancer only	—	—	—	2.9					
	ICAM1[a]	1.6	0.9 (1.0)[c]	2.2	1.5	3.9	IHC [63]	ELISA [40]			
	LIMS1	2.2	—	0.9	0.8	16.3					
	LTBP4	Cancer only	—	1.7	1.9	2.2					
	LTF	2.3	—	5.7	0.9	17.0	IHC [64]		↓ [60]	■ [54]	
	LYZ	3.2	1.1	3.6	2.6	33.3	IHC [65]		↓ [60]	■ [54]	
	MSH6	2.3	—	1.0	0.9	3.2					
	PTPRG[a]	Cancer only	—	2.8	2.4	3.8					
	SOD3	2.9	0.3	0.5	0.6	4.7			↑ [59]		
	SPARCL1	1.6	0.8	1.0	1.9	5.7					
	TNC[a]	Cancer only	—	6.6	6.9	180.0	IHC [39]	ELISA [39]			
	TNFRSF1A[a]	2.5	—	2.4	2.6	1.5		ELISA [41]			
	VASP	2.3	—	3.8	2.0	4.2					
	WFDC2[a]	2.6	—	2.1	0.9	25.0	IHC [48]				
	ZDHHC20	Cancer only	—	—	—	3.4					
Increased levels in PanIN and increased expression of corresponding genes	**CD248**	1.3	1.8	—	—	6.4					
	CD97	1.2	2.6	4.0	3.7	9.9	IHC [66]	ELISA [66]			
	CDH1	—	2.0	1.9	0.8	2.6					
	EFEMP2	1.2	3.1	2.1	3.4	30.0			↑ [52]		
	EFNA1	1.0	Cancer only	2.3	1.9	2.6					
	GKN1	1.6	Cancer only	—	—	12.9					
	IGFBP4[a]	1.3	1.5	2.7	5.2	37.5					
	TFF2	—	3.3	4.0	4.6	0.1	IHC [67]				
	TGFBI	—	2.4	3.2	4.7	15.9			↓ [60], ↑[52]		
	THBS1	1.2	3.1	2.5	5.9	2.6	IHC [68]		↓ [60], ↑[52]		

[a]Proteins further validated by immunoassays.
[b]↑, up-regulated; ↓, down-regulated; ■, detected.
[c]Measurements of protein ratio performed by ELISA for the same mouse plasma samples used in the proteomic analysis.
*p-Value for all ratios <0.05.
doi:10.1371/journal.pmed.0050123.t001

domain 2), and IGFBP4 were obtained by ELISA. These proteins that were up-regulated in plasma from tumor-bearing mice were assayed in human sera from 30 patients with PDAC to assess their significance individually and as a panel, together with CA19–9, a marker that is currently in clinical use as a pancreatic cancer marker (Table 2) [33]. As a control group, we analyzed sera from 20 matched healthy individuals and ten to 15 individuals with chronic pancreatitis, obtained using the same protocol and storage conditions. Information regarding patient characteristics and
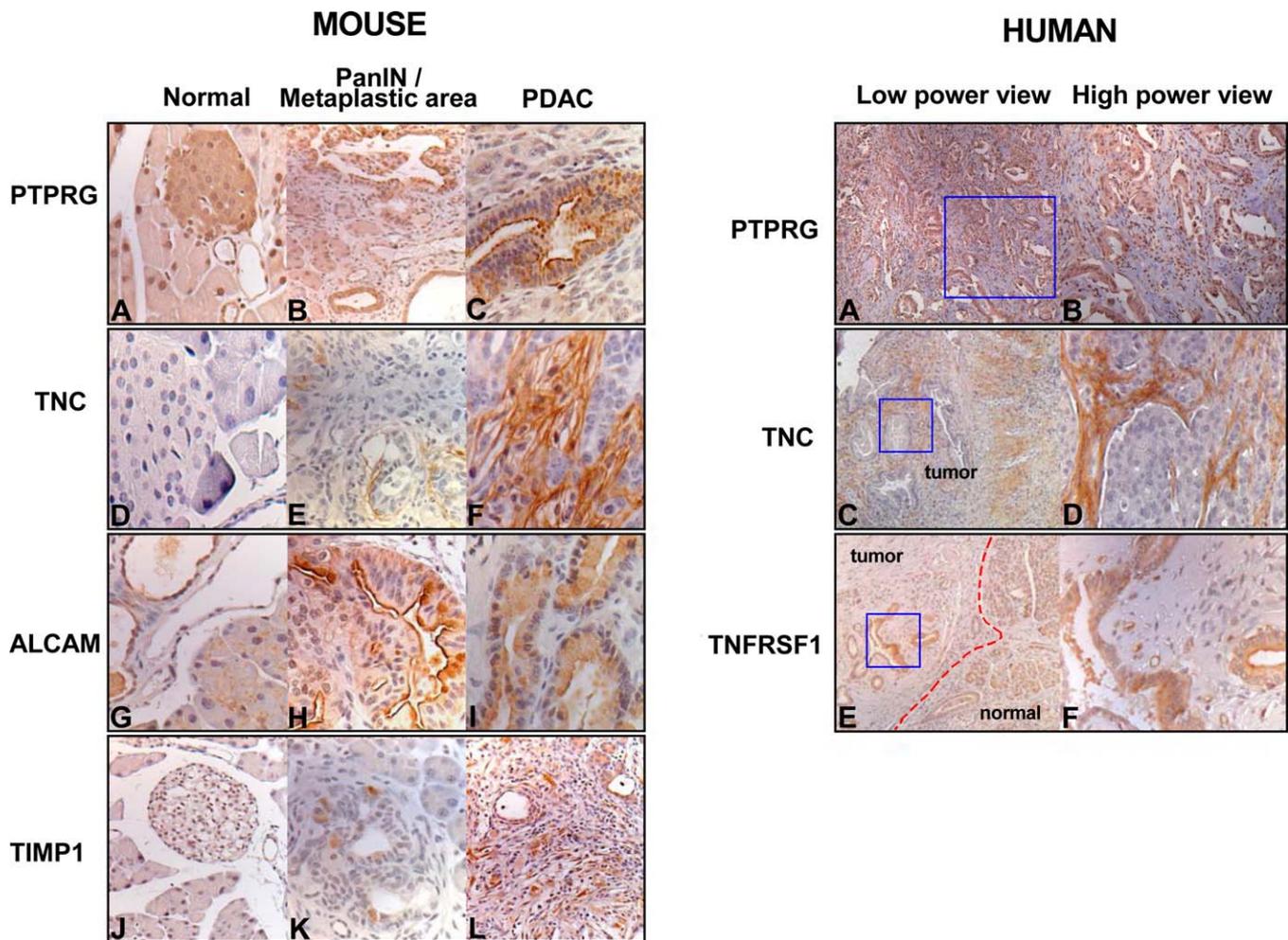
Figure 4. IHC Analysis of Candidate PDAC Biomarkers in Mouse and Human Tissue

Mouse, left photomicrographs: PTPRG expression (A–C). Note islet staining in normal pancreas (A); membranous staining is seen in PanIN and PDAC epithelium (B and C). TNC expression (D–F). Note lack of staining in normal pancreatic tissue (D); strong expression is present in stroma of PanIN (E) and PDAC (F). ALCAM expression (G–I). Note membranous staining of the normal pancreatic acinar and ductal cells (G); increased staining is present in the PanIN epithelium (H) and PDAC cells (I). TIMP1 expression (J–L). Note lack of staining in normal pancreatic tissue (J); staining is observed in association with acinar-ductal metaplasia (K) and both PDAC stromal and tumor cells (L). (A–C, J–L, magnification 200×; D–I, magnification 400×).

Human, right photomicrographs: PTPRG expression (A, B). Note membranous staining in PDAC epithelium and absence of staining in normal pancreas. TNC expression (C, D). Note expression in PDAC stroma. TNFRSF1 expression (E, F). Note membranous staining in PDAC epithelium; normal pancreatic tissue is negative. Dashed red lines subdivide different histology of the tissue analysed, and blue boxes indicate the adjacent magnified region. Six independent tumor specimens were stained per antibody for mouse IHC and three for human. PTTGF, TNFRSF1, and ALCAM each showed positive staining in at least 15% of tumor cells. TNC and TIMP1 showed positive staining in at least 50% of stromal cells. A, C, and E, magnification 100×; B, magnification 200×; D and F, magnification 400×.

doi:10.1371/journal.pmed.0050123.g004

tumor stage is provided in Table S1. Statistical analysis was performed for individual proteins and for the entire panel as a group. All but one of the proteins were significantly elevated in cancer compared to one or both control groups ($p < 0.03$). Seven proteins were compared between cancer and both control groups, and five of seven were significant in both comparisons ($p < 0.03$) (Figure S4 for box plots, Table 2). Only one protein (LCN2) did not achieve statistical significance. For proteins that yielded statistically significant differences between cancer and healthy individuals, the areas under the curve (AUCs) ranged between 0.75 and 0.89 (Figure S5), and between cancer and pancreatitis the AUCs ranged between 0.74 and 0.92. Of note, a panel of all the proteins tested, inclusive of those that did not achieve statistical significance individually so as to avoid any overfitting, yielded

an AUC of 0.96 in contrast to CA 19–9, which yielded an AUC of 0.79 (Figure 6A and 6B).

Plasma analysis of mice at the PanIN stage allows us to test whether protein changes in plasma observed at an early tumor stage in the mouse may be up-regulated in individuals with pancreatic cancer before actual clinical diagnosis. To that effect, a blinded analysis was conducted using sera collected as part of CARET, which included 18,314 participants [34]. The CARET study was intended to test the effect of daily beta-carotene and retinyl palmitate on cancer incidence and death in individuals with a history of smoking or asbestos exposure. All participants (13) in the cohort diagnosed with pancreatic cancer between 7–13 mo following a blood draw (mean = 10 mo) and an equal number of controls that were matched for age, sex, year of CARET
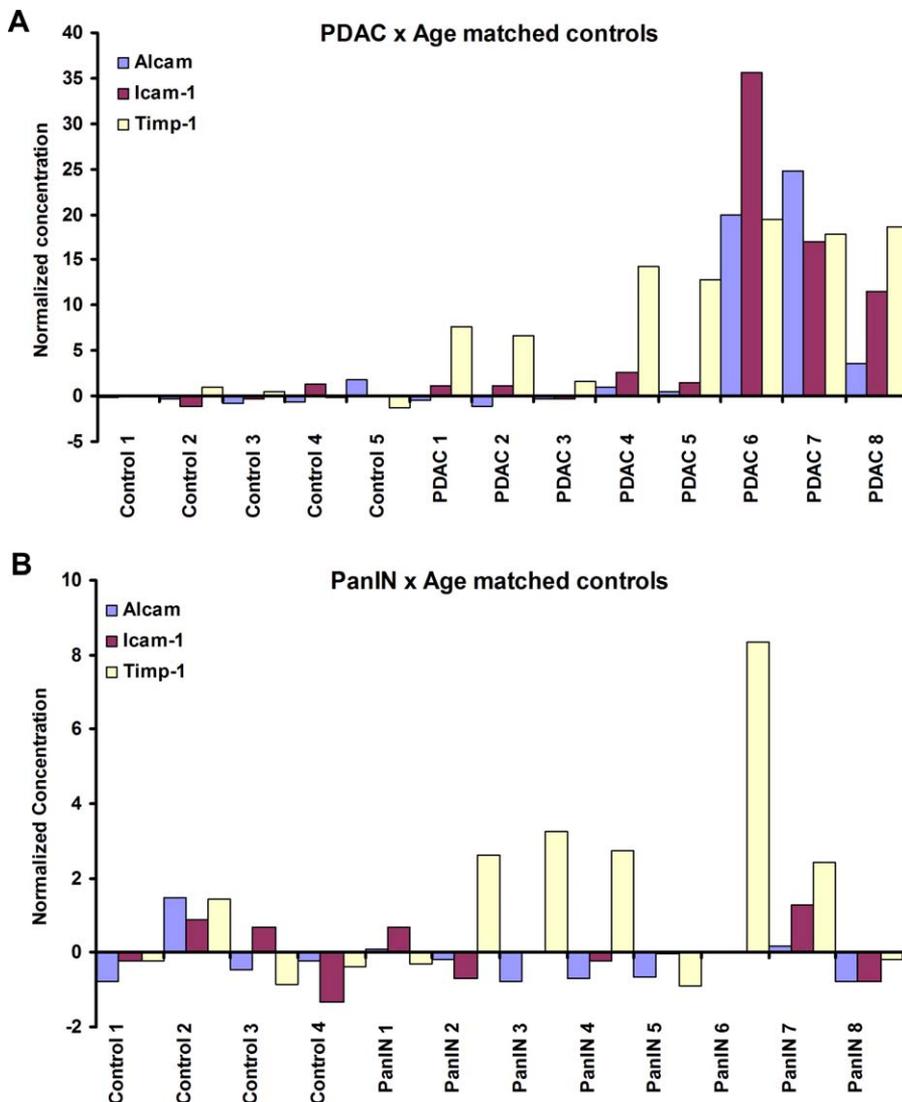
Figure 5. Validation Study of ALCAM, ICAM1, and TIMP1 in Mouse Plasma by ELISA

(A) Plasma from the same individual mice used for proteomics discovery were utilized for validation by means of ELISA. ALCAM, TIMP1, and ICAM1 were all elevated in plasma of the PDAC mice.

(B) TIMP1 was also elevated in plasma of PanIN mice. ALCAM overall (cancer plus controls) concentration in mouse plasma was 19 ng/ml; ICAM1 was 163 ng/ml; and TIMP1was 6.2 ng/ml. The low ng/ml concentrations of these proteins support the substantial depth of analysis achieved with our discovery platform. Normalization of concentration was performed as described in the Materials and Methods.

doi:10.1371/journal.pmed.0050123.g005

enrollment, and time of blood draw in relation to enrollment and who were not diagnosed with pancreatic cancer on the basis of information in the CARET database, were identified by CARET for the blinded pancreatic cancer validation study. The pancreatic cancer and control groups were also matched for CARET intervention. Information regarding CARET patient characteristics and tumor stage is provided in Table S2. We tested five proteins that were up-regulated in mouse plasma at the PanIN stage (LCN2, REG1A, REG3, TIMP1, and IGFBP4) together with CA19.9, without knowledge of which individuals developed pancreatic cancer subsequent to the blood draw and which individuals were matched controls (Table 2). When tested individually, two of the five proteins (IGFBP4 and TIMP1) showed significance at 0.05 and 0.04, respectively. CA19.9 was significant at 0.04. As a panel, the five proteins achieved an AUC of 0.817 ($p = 0.005$), inclusive

of the three proteins that did not achieve statistical significance individually to avoid any overfitting. When the panel of five proteins was combined with CA19.9, an AUC of 0.911 was achieved (Figure 6C).

Discussion

Our findings here indicate that plasma proteomic analysis of GEM models of cancer provide a useful strategy to identify candidate markers applicable to human cancer with potential utility for early detection. This is very relevant, since there is a compelling need to develop blood-based markers that allow early cancer detection, classify tumors to direct therapy, and monitor disease progression, regression, or recurrence. Early detection is particularly relevant to pancreatic adenocarcinoma, which is the fourth leading cause of cancer death in the United States and with a 5-y survival rate of only 3%. Because

Table 2. ELISA Analysis of Up-Regulated Proteins in Pancreatic Cancer Sera and Prediagnostic Sera

Status of Diagnosis	Protein[a]	Cancer		Controls		Pancreatitis		AUC		Wilcoxon Rank-Sum Test (p-Value)	
		Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Cancer Versus Controls	Cancer Versus Pancreatitis	Cancer Versus Controls	Cancer Versus Pancreatitis
Confirmed Diagnosis[b]	CA 19–9 (U/ml)	147.7	75.6	14.7	17.0	65.0	82.0	0.98	0.79	<0.001	0.007
	ALCAM	139.9	65.9	84.2	26.0	96.4	30.6	0.85	0.74	<0.001	0.008
	TIMP1	322.7	196.0	182.3	33.7	178.3	39.3	0.89	0.88	<0.001	<0.001
	ICAM1	457.4	384.8	270.7	67.8	121.4	64.9	0.75	0.92	0.029	<0.001
	LCN2	162.6	83.8	138.6	49.2	120.3	26.8	0.65	0.61	0.452	0.318
	REG1A	1,108.4	624.9	691.3	342.2	1,277.6	1,219.5	0.79	0.57	0.002	0.528
	REG3	22.6	15.1	8.6	5.1	11.0	5.9	0.88	0.77	<0.001	0.011
	IGFBP4	153.5	125.6	116.1	74.5	89.6	29.6	0.69	0.74	0.106	0.022
	TNFRSF1A	2.8	1.5	1.8	0.9	—	—	0.82	—	0.008	—
	WFDC2 (HE4)	10.9	18.4	2.5	8.2	—	—	0.89	—	0.001	—
Prediagnosis	CA19–9 (U/ml)	56.1	74.2	10.6	6.7	—	—	0.74	—	0.040	—
	TIMP1	212.7	62.4	162.1	36	—	—	0.74	—	0.040	—
	IGFBP4	61	33.1	45.6	50.5	—	—	0.72	—	0.050	—
	LCN2	123.5	42.6	104.4	66.1	—	—	0.68	—	0.110	—
	REG3	25.2	21.7	17.2	14.3	—	—	0.65	—	0.220	—
	REG1A	2,111.4	1,151.9	2,059.1	1325.1	—	—	0.52	—	0.980	—

[a]All protein concentrations in human serum, except CA 19–9, are reported in ng/ml.
[b]Thirty human PDAC sera (20 sera for TNFRSF1A and HE4), 20 matched controls, and ten chronic pancreatitis (15 sera for TIMP1 and ALCAM) were assayed for the listed biomarker candidates.
doi:10.1371/journal.pmed.0050123.t002

of limitations in diagnostic methods and a lack of specific symptoms at an early stage, the disease is often diagnosed at late stages. In contrast, early stage disease is associated with prolonged survival following surgical resection of the tumor [35]. Therefore, improvement in means to detect pancreatic cancer early would be expected to impact outcome.

While published studies have pointed to the merits of proteomics for cancer marker identification, the challenge of discovering markers applicable to early detection has been substantial. Mass spectrometry has evolved from a tool to identify and characterize isolated proteins or for mass peak profiling to a platform for interrogating complex proteomes. However, even with recent improvements in sensitivity and mass accuracy, the complexity of the plasma proteome far exceeds the current capabilities of mass spectrometry to fully resolve their individual protein and peptide constituents in a single analysis. Current strategies to achieve in-depth coverage require sample fractionation followed by separate analyses of individual fractions or capture of protein or peptide subsets [2]. The depth of proteomic analysis achieved in this study through extensive fractionation of intact proteins and reliance on high-resolution mass spectrometry has allowed identification of low abundance proteins [3]. In addition, reliance on acrylamide isotope labeling of cysteines has allowed quantitative measures to be derived from mass spectrometric analysis of the plasma proteome. More importantly, the identification of proteins is not restricted to peptides containing cysteine residues, since in our work-flow there is no capture step of isotopically labeled peptides as in the isotope-coded affinity (ICAT) method tags [36], thus providing a comprehensive list of peptides in the digests and consequently better protein coverage and confidence in protein identification. Such depth of analysis is necessary to identify potential tumor specific biomarkers and to extend discovery beyond abundant protein changes resulting from

inflammation or acute-phase reaction. As a result, changes in plasma proteins relevant to pancreatic cancer could be identified across a wide dynamic range of protein abundance.

Some of the proteins identified in this study as potentially relevant to pancreatic cancer have already been associated with cancer, as evidenced from Ingenuity Pathway Analysis. In total, 13 proteins were previously investigated in pancreatic cancer tissue or for a smaller number in human blood by immunoassay (Table 1) and found to be elevated. Among those, MMP2 and its inhibitor TIMP1 are known to be involved in tumor progression and extracellular matrix degradation [37]. REG1A and REG3 are proteins highly secreted by pancreatic islet cells and have also been described as potential markers for pancreatic diseases [38]. ICAM1 and TNC are involved, respectively, in cellular attachment and inhibition of adhesion of cells to the extracellular matrix [39,40]. TNFRSF1A, has been associated with the acute-phase process [41].

Elevated levels of ALCAM, IGFBP4, LCN2, and WFDC2 in circulation in pancreatic cancer are novel findings. ALCAM is a cell adhesion molecule critical to tumor development and progression [42]. The form of ALCAM detected in circulation, corresponds to the shed extracellular domain of this integral membrane protein. The process of shedding is promoted by metalloproteases [43]. Overexpression of IGFBP4, the smallest protein from the IGF binding protein family, has been related to tumor growth [44]. LCN2 has been shown to play a role in regulating cellular growth and metastasis in colon cancer [45] and to be overexpressed in pancreatic cancer at the mRNA levels [46], concordant with gene expression analysis of tumors from our mouse model (Table 1). Interestingly, WFDC2 (or HE4), a promising biomarker for ovarian cancer [47], was found in our study to be up-regulated in mouse PDAC plasma, with concordant mRNA expression. Additionally, WFDC2 was listed as up-regulated at
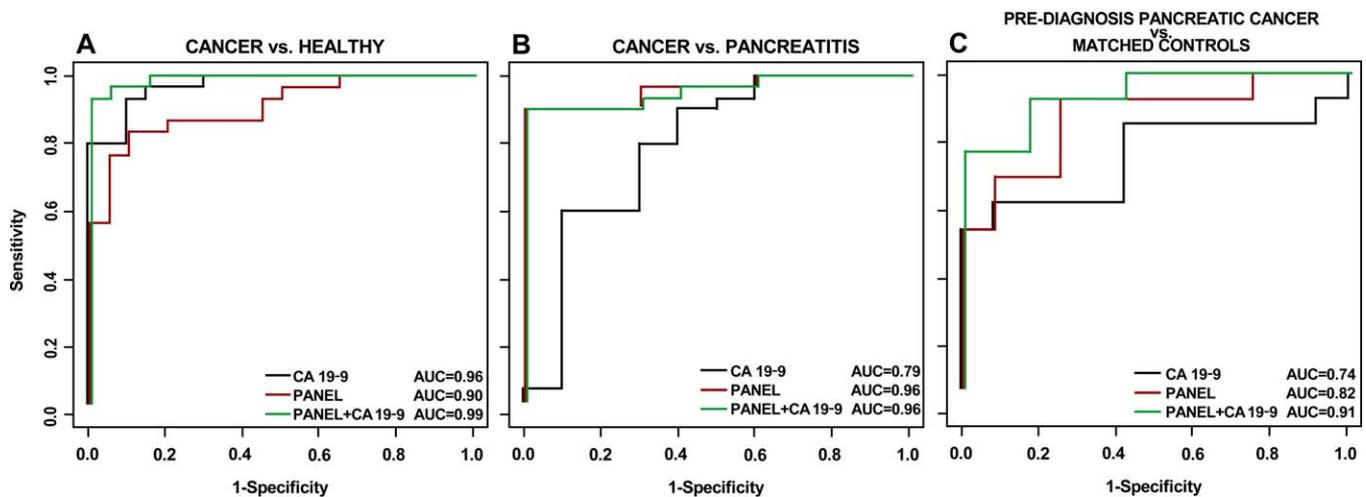
Figure 6. ROC in Assays of Human Samples for Two Panels of Proteins Identified in Proteomic Analysis of Plasmas from Tumor-Bearing Mice

ROC curves based on ELISA measurements of ALCAM, ICAM1, LCN2, TIMP1, REG1A, REG3, and IGFBP4 as a panel with or without CA19–9 comparing pancreatic cancer versus healthy controls (A) and pancreatic cancer versus pancreatitis (B). This panel of candidates was chosen on the basis of up-regulation in tumor-bearing mice. As expected, CA19–9 performed well in comparisons with healthy individuals as controls; however, the chosen panel was significantly better than CA19–9 alone when pancreatitis patients were used as controls.

(C) The panel tested with prediagnostic sera (LCN2, TIMP1, REG1A, REG3, and IGFBP4) was chosen on the basis of up-regulation at the PanIN stage. This panel performed slightly better in comparison to CA19–9 but a combination of CA19–9 with the panel of candidates significantly improved discrimination between early stage (prediagnostic) sera and matched controls. Standardization procedures and composite marker ROCs were generated without fitting, by inclusion of all tested candidate markers. Specimens from controls and pancreatitis patients were obtained from the same institution and with the same protocol for blood collection. For details see Materials and Methods and Tables S1 and S2.

doi:10.1371/journal.pmed.0050123.g006

the gene and protein levels in human PDAC tissue in a recent study [48], suggesting that this protein may also have relevance to pancreatic cancer. The whey-acidic protein (WAP) family has been described to be involved in tumor progression through the regulation of the NFkB signaling pathway [49], and a second member of this family, secretory leucocyte proteinase inhibitor (SLPI), is among our list of candidates up-regulated in both PDAC and PanIN mouse samples. PTPRG, a tyrosine phosphatase receptor that was validated in this study in both mouse and human by immunohistochemistry, has been recently described in gastric cancer as a potential tumor suppressor gene that is methylated in metastatic cells [50].

All together, the prior association of proteins identified in this study with cancer and for some with demonstrated function in pancreatic cancer is indicative of the utility of mouse models for deciphering protein changes relevant to pancreatic and other cancers in humans. Also, it should be emphasized that previously, these proteins were studied independently of each other and not identified through a systematic profiling study as presented here.

Because mice can be sampled at defined stages of tumor development and under controlled breeding conditions, greater standardization is possible using mouse models compared to human studies. Mouse models also allowed in this study investigations at an early stage of tumor development (PanIN), allowing identification of proteins associated with early events in tumorigenesis. The strong concordance between mouse and human pancreatic cancer in both tissue and circulating markers is striking. From the list of nine candidate markers found elevated by proteomics and validated in human samples, only LCN2 was not significantly elevated.

Our analysis of candidate protein markers in newly

diagnosed patient samples confirmed that CA19–9 discriminates pancreatic cancer at the time of diagnosis well from healthy controls (see Table S1). CA19–9 levels were elevated in more than 80% of patients compared with healthy controls. However the sensitivity and specificity of CA19–9 in other settings relevant to pancreatic cancer, namely in discriminating between pancreatitis and pancreatic cancer and for detecting cancer at an early stage, are much reduced compared with its power to discriminate newly diagnosed pancreatic cancer and healthy individuals [51], hence the need for additional markers to constitute a panel with improved sensitivity and specificity for discriminating pancreatic cancer from pancreatitis and for detecting the disease at an early stage prior to onset of symptoms. In this respect, TIMP1 and ICAM1 had superior performance when cancer samples were compared to samples from pancreatitis patients. The panel of candidate markers that we tested, together with CA19–9, significantly improved sensitivity and specificity in preclinical samples.

The next steps in building on our findings include developing high throughput assays for additional candidate markers identified, for which such assays are not currently available, and to expand validation studies to address specific applications, notably for implementing a panel-based test to distinguish between pancreatitis and pancreatic cancer and to further assess the utility of a panel approach for detecting pancreatic cancer early among individuals at increased risk of developing the disease.

Supporting Information

Figure S1. Distribution of Quantitative Events

(A) Equal amounts of total immunodepleted nonfractionated human plasma were labeled with heavy and light acrylamide and analyzed

with LC–MS/MS. The histogram represents the distribution of 4,371 quantitative events. From this control-control events distribution, the number of events that exceeds a given ratio was determined. For instance, there were 125 up-regulated events (ratio $\geq$ 2.0), which corresponds to 2.8% ($p = 0.028$).
(B) Distribution of quantitative events for PDAC experiment. Total, $n = 65,640$; up-regulated, $n = 14,420$ (22%).
(C) Distribution of quantitative events for PanIN experiment. Total, $n = 58,063$; up-regulated, $n = 9,396$ (16%).
Found at doi:10.1371/journal.pmed.0050123.sg001 (675 KB TIF).

Figure S2. Concordance in Protein Identification between PanIN and PDAC Experiments

(A) High abundant proteins were detected consistently in both PanIN and PDAC experiments, while low abundant proteins were more susceptible to sampling limitations in data acquisition for LC–MS/MS.
(B) Protein identification concordance between duplicate runs of fractions. Reversed-phase fractions from anion exchange fraction 6 of both PanIN and PDAC were run in duplicate, and the same behavior was observed. Protein concentration was estimated on the basis of spectral counts (Figure 3). The majority of medium to high abundance proteins (>1 μg/ml or >100 ms events) were detected in both PanIN and PDAC experiments, while most differences in protein identifications between the two experiments represented lower abundance proteins.
Found at doi:10.1371/journal.pmed.0050123.sg002 (608 KB TIF).

Figure S3. Protein Tissue Distribution Versus Relative Protein Abundance

The 1,442 proteins identified were correlated to tissue specificity using published datasets from mouse liver proteomic profiling studies [31,32] and one human tissue mRNA expression study [27]. Proteins estimated to be of low abundance (<100 ng/ml) had a much greater representation of pancreatic proteins relative to liver proteins based on tissue protein and/or mRNA. Protein concentration was estimated on the basis of MS2 events (Figure 2).
Found at doi:10.1371/journal.pmed.0050123.sg003 (270 KB TIF).

Figure S4. Box Plot for ELISA Measurements of Proteins Relevant to Pancreatic Cancer

A total of 30 human PDAC sera (20 sera for TNFRSF1A and WFDC2), 20 healthy, and ten chronic pancreatitis (15 sera for TIMP1 and ALCAM) were assayed for the plotted proteins. More statistical detail about this data is presented in Table 2. The horizontal axis legend represents: 1, cancer; 2, normal; 3, pancreatitis.
Found at doi:10.1371/journal.pmed.0050123.sg004 (1.4 MB TIF).

Figure S5. Receiver Operating Characteristic Curves for ELISA Measurements of Proteins Relevant to Pancreatic Cancer

All data points for each individual protein were used in the Receiver Operating Characteristic (ROC) plots, without applying any cut-off value. The total number of samples used in this analysis was: 30 human PDAC sera (20 sera for TNFRSF1A and WFDC2), 20 healthy, and ten chronic pancreatitis (15 sera for TIMP1 and ALCAM). Detailed statistical information for these proteins is presented in Table 2.
Found at doi:10.1371/journal.pmed.0050123.sg005 (1.5 MB TIF).

Table S1. Clinical Characteristics and Protein Assay Values for Newly Diagnosed Pancreatic Cancer Patients
Found at doi:10.1371/journal.pmed.0050123.st001 (25 KB XLS).

Table S2. Clinical Characteristics and Protein Assay Values for Pancreatic Cancer Patients from the CARET Cohort Study
Found at doi:10.1371/journal.pmed.0050123.st002 (21 KB XLS).

Table S3. Proteins Identified with <1% FDR
Found at doi:10.1371/journal.pmed.0050123.st003 (1.2 MB XLS).

Table S4. Proteins Identified with <5% FDR and Relevant to Pancreatic Cancer
Found at doi:10.1371/journal.pmed.0050123.st004 (149 KB XLS).

Table S5. Proteins Up-Regulated in PDAC and/or PanIN Mouse Plasma
Found at doi:10.1371/journal.pmed.0050123.st005 (79 KB XLS).

References

1. Hanash S (2003) Disease proteomics. Nature 422: 226–232.
2. Hanash SM, Pitteri SJ, Faca VM (2008) Mining the plasma proteome for cancer biomarkers. Nature 452: 571–579.
3. Faca V, Pitteri SJ, Newcomb L, Glukhova V, Phanstiel D, et al. (2007) Contribution of protein fractionation to depth of analysis of the serum and plasma proteomes. J Proteome Res 6: 3558–3565.
4. States DJ, Omenn GS, Blackwell TW, Fermin D, Eng J, et al. (2006) Challenges in deriving high-confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study. Nat Biotechnol 24: 333–338.
5. Sweet-Cordero A, Mukherjee S, Subramanian A, You H, Roix JJ, et al. (2005) An oncogenic KRAS2 expression signature identified by cross-species gene-expression analysis. Nat Genet 37: 48–55.
6. Zender L, Spector MS, Xue W, Flemming P, Cordon-Cardo C, et al. (2006) Identification and validation of oncogenes in liver cancer using an integrative oncogenomic approach. Cell 125: 1253–1267.
7. Kim M, Gans JD, Nogueira C, Wang A, Paik JH, et al. (2006) Comparative oncogenomics identifies NEDD9 as a melanoma metastasis gene. Cell 125: 1269–1281.
8. Maser RS, Choudhury B, Campbell PJ, Feng B, Wong KK, et al. (2007) Chromosomally unstable mouse tumours have genomic alterations similar to diverse human cancers. Nature 447: 966–971.
9. Hingorani SR, Petricoin EF, Maitra A, Rajapakse V, King C, et al. (2003) Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. Cancer Cell 4: 437–450.
10. Hezel AF, Kimmelman AC, Stanger BZ, Bardeesy N, Depinho RA (2006) Genetics and biology of pancreatic ductal adenocarcinoma. Genes Dev 20: 1218–1249.
11. Moskaluk CA, Hruban RH, Kern SE (1997) p16 and K-ras gene mutations in the intraductal precursors of human pancreatic adenocarcinoma. Cancer Res 57: 2140–2143.
12. DiGiuseppe JA, Hruban RH, Goodman SN, Polak M, van den Berg FM, et al. (1994) Overexpression of p53 protein in adenocarcinoma of the pancreas. Am J Clin Pathol 101: 684–688.
13. Aguirre AJ, Bardeesy N, Sinha M, Lopez L, Tuveson DA, et al. (2003) Activated Kras and Ink4a/Arf deficiency cooperate to produce metastatic pancreatic ductal adenocarcinoma. Genes Dev 17: 3112–3126.
14. Bardeesy N, Aguirre AJ, Chu GC, Cheng KH, Lopez LV, et al. (2006) Both p16(Ink4a) and the p19(Arf)-p53 pathway constrain progression of pancreatic adenocarcinoma in the mouse. Proc Natl Acad Sci U S A 103: 5947–5952.
15. Hingorani SR, Wang L, Multani AS, Combs C, Deramaudt TB, et al. (2005) Trp53R172H and KrasG12D cooperate to promote chromosomal instability and widely metastatic pancreatic ductal adenocarcinoma in mice. Cancer Cell 7: 469–483.
16. Faca V, Coram M, Phanstiel D, Glukhova V, Zhang Q, et al. (2006) Quantitative analysis of acrylamide labeled serum proteins by LC-MS/MS. J Proteome Res 5: 2009–2018.
17. Misek DE, Kuick R, Wang H, Galchev V, Deng B, et al. (2005) A wide range of protein isoforms in serum and plasma uncovered by a quantitative intact protein analysis system. Proteomics 5: 3343–3352.

18. Wang H, Clouthier SG, Galchev V, Misek DE, Duffner U, et al. (2005) Intact-protein-based high-resolution three-dimensional quantitative analysis system for proteome profiling of biological fluids. Mol Cell Proteomics 4: 618–625.

19. Rauch A, Bellew M, Eng J, Fitzgibbon M, Holzman T, et al. (2006) Computational Proteomics Analysis System (CPAS): an extensible, open-source analytic system for evaluating and publishing proteomic data and high throughput biological experiments. J Proteome Res 5: 112–121.

20. MacLean B, Eng JK, Beavis RC, McIntosh M (2006) General framework for developing and evaluating database scoring algorithms using the TANDEM search engine. Bioinformatics 22: 2830–2832.

21. Keller A, Nesvizhskii AI, Kolker E, Aebersold R (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem 74: 5383–5392.

22. Nesvizhskii AI, Keller A, Kolker E, Aebersold R (2003) A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem 75: 4646–4658.

23. Scholler N, Crawford M, Sato A, Drescher CW, O'Briant KC, et al. (2006) Bead-based ELISA for validation of ovarian cancer early detection markers. Clin Cancer Res 12: 2117–2124.

24. McIntosh MW, Drescher C, Karlan B, Scholler N, Urban N, et al. (2004) Combining CA 125 and SMR serum markers for diagnosis and early detection of ovarian carcinoma. Gynecol Oncol 95: 9–15.

25. Pepe MS, Longton G (2005) Standardizing diagnostic markers to evaluate and compare their performance. Epidemiology 16: 598–603.

26. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44: 837–845.

27. Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, et al. (2004) A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci U S A 101: 6062–6067.

28. Logsdon CD, Simeone DM, Binkley C, Arumugam T, Greenson JK, et al. (2003) Molecular profiling of pancreatic adenocarcinoma and chronic pancreatitis identifies multiple genes differentially regulated in pancreatic cancer. Cancer Res 63: 2649–2657.

29. Hood BL, Zhou M, Chan KC, Lucas DA, Kim GJ, et al. (2005) Investigation of the mouse serum proteome. J Proteome Res 4: 1561–1568.

30. Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendoza A, et al. (2005) Comparison of label-free methods for quantifying human proteins by shotgun proteomics. Mol Cell Proteomics 4: 1487–1502.

31. Foster LJ, de Hoog CL, Zhang Y, Zhang Y, Xie X, et al. (2006) A mammalian organelle map by protein correlation profiling. Cell 125: 187–199.

32. Kislinger T, Cox B, Kannan A, Chung C, Hu P, et al. (2006) Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling. Cell 125: 173–186.

33. Goonetilleke KS, Siriwardena AK (2007) Systematic review of carbohydrate antigen (CA 19–9) as a biochemical marker in the diagnosis of pancreatic cancer. Eur J Surg Oncol 33: 266–270.

34. Goodman GE, Thornquist MD, Balmes J, Cullen MR, Meyskens FL Jr., et al. (2004) The Beta-Carotene and Retinol Efficacy Trial: incidence of lung cancer and cardiovascular disease mortality during 6-year follow-up after stopping beta-carotene and retinol supplements. J Natl Cancer Inst 96: 1743–1750.

35. Bardeesy N, DePinho RA (2002) Pancreatic cancer biology and genetics. Nat Rev Cancer 2: 897–909.

36. Shiio Y, Aebersold R (2006) Quantitative proteome analysis using isotope-coded affinity tags and mass spectrometry. Nat Protoc 1: 139–145.

37. Gong YL, Xu GM, Huang WD, Chen LB (2000) Expression of matrix metalloproteinases and the tissue inhibitors of metalloproteinases and their local invasiveness and metastasis in Chinese human pancreatic cancer. J Surg Oncol 73: 95–99.

38. Okamoto H (1999) The Reg gene family and Reg proteins: with special attention to the regeneration of pancreatic beta-cells. J Hepatobiliary Pancreat Surg 6: 254–262.

39. Esposito I, Penzel R, Chaib-Harrireche M, Barcena U, Bergmann F, et al. (2006) Tenascin C and annexin II expression in the process of pancreatic carcinogenesis. J Pathol 208: 673–685.

40. Markocka-Maczka K (2003) [Concentration of serum soluble forms of ICAM-1 (sVCAM-1) and VCAM-1 (sVCAM-1) in patients with chronic pancreatitis and in patients with pancreatic carcinoma]. Wiad Lek 56: 147–151.

41. Barber MD, Fearon KC, Ross JA (1999) Relationship of serum levels of interleukin-6, soluble interleukin-6 receptor and tumour necrosis factor receptors to the acute-phase protein response in advanced pancreatic cancer. Clin Sci (Lond) 96: 83–87.

42. Ofori-Acquah SF, King JA (2008) Activated leukocyte cell adhesion molecule: a new paradox in cancer. Transl Res 151: 122–128.

43. Rosso O, Piazza T, Bongarzone I, Rossello A, Mezzanzanica D, et al. (2007) The ALCAM shedding by the metalloprotease ADAM17/TACE is involved in motility of ovarian carcinoma cells. Mol Cancer Res 5: 1246–1253.

44. Durai R, Davies M, Yang W, Yang SY, Seifalian A, et al. (2006) Biology of insulin-like growth factor binding protein-4 and its role in cancer (review). Int J Oncol 28: 1317–1325.

45. Lee HJ, Lee EK, Lee KJ, Hong SW, Yoon Y, et al. (2006) Ectopic expression of neutrophil gelatinase-associated lipocalin suppresses the invasion and liver metastasis of colon cancer cells. Int J Cancer 118: 2490–2497.

46. Laurell H, Bouisson M, Berthelemy P, Rochaix P, Dejean S, et al. (2006) Identification of biomarkers of human pancreatic adenocarcinomas by expression profiling and validation with gene expression analysis in endoscopic ultrasound-guided fine needle aspiration samples. World J Gastroenterol 12: 3344–3351.

47. Hellstrom I, Raycraft J, Hayden-Ledbetter M, Ledbetter JA, Schummer M, et al. (2003) The HE4 (WFDC2) protein is a biomarker for ovarian carcinoma. Cancer Res 63: 3695–3700.

48. Galgano MT, Hampton GM, Frierson HF Jr. (2006) Comprehensive analysis of HE4 expression in normal and malignant human tissues. Mod Pathol 19: 847–853.

49. Bouchard D, Morisset D, Bourbonnais Y, Tremblay GM (2006) Proteins with whey-acidic-protein motifs and cancer. Lancet Oncol 7: 167–174.

50. Wang JF, Dai DQ (2007) Metastatic suppressor genes inactivated by aberrant methylation in gastric cancer. World J Gastroenterol 13: 5692–5698.

51. Koopmann J, Rosenzweig CN, Zhang Z, Canto MI, Brown DA, et al. (2006) Serum markers in patients with resectable pancreatic adenocarcinoma: macrophage inhibitory cytokine 1 versus CA19–9. Clin Cancer Res 12: 442–446.

52. Chen R, Yi EC, Donohoe S, Pan S, Eng J, et al. (2005) Pancreatic cancer proteome: the proteins that underlie invasion, metastasis, and immuno-logic escape. Gastroenterology 129: 1187–1197.

53. Chen R, Pan S, Cooke K, Moyes KW, Bronner MP, et al. (2007) Comparison of pancreas juice proteins from cancer versus pancreatitis using quantitative proteomic analysis. Pancreas 34: 70–79.

54. Gronborg M, Bunkenborg J, Kristiansen TZ, Jensen ON, Yeo CJ, et al. (2004) Comprehensive proteomic analysis of human pancreatic juice. J Proteome Res 3: 1042–1055.

55. Zhou L, Lu Z, Yang A, Deng R, Mai C, et al. (2007) Comparative proteomic analysis of human pancreatic juice: methodological study. Proteomics 7: 1345–1355.

56. Kakisaka T, Kondo T, Okano T, Fujii K, Honda K, et al. (2007) Plasma proteomics of pancreatic cancer patients by multi-dimensional liquid chromatography and two-dimensional difference gel electrophoresis (2D-DIGE): up-regulation of leucine-rich alpha-2-glycoprotein in pancreatic cancer. J Chromatogr B Analyt Technol Biomed Life Sci 852: 257–267.

57. Motoo Y, Satomura Y, Mouri I, Mouri H, Ohtsubo K, et al. (1999) Serum levels of pancreatitis-associated protein in digestive diseases with special reference to gastrointestinal cancers. Dig Dis Sci 44: 1142–1147.

58. Satomura Y, Sawabu N, Mouri I, Yamakawa O, Watanabe H, et al. (1995) Measurement of serum PSP/reg-protein concentration in various diseases with a newly developed enzyme-linked immunosorbent assay. J Gastroenterol 30: 643–650.

59. Chen R, Brentnall TA, Pan S, Cooke K, Moyes KW, et al. (2007) Quantitative proteomics analysis reveals that proteins differentially expressed in chronic pancreatitis are also frequently involved in pancreatic cancer. Mol Cell Proteomics 6: 1331–1342.

60. Gronborg M, Kristiansen TZ, Iwahori A, Chang R, Reddy R, et al. (2006) Biomarker discovery from pancreatic cancer secretome using a differential proteomic approach. Mol Cell Proteomics 5: 157–171.

61. Crnogorac-Jurcevic T, Gangeswaran R, Bhakta V, Capurso G, Lattimore S, et al. (2005) Proteomic analysis of chronic pancreatitis and pancreatic adenocarcinoma. Gastroenterology 129: 1454–1463.

62. Bloomston M, Shafii A, Zervos EE, Rosemurgy AS (2002) TIMP-1 over-expression in pancreatic cancer attenuates tumor growth, decreases implantation and metastasis, and inhibits angiogenesis. J Surg Res 102: 39–44.

63. van Grevenstein WM, Hofland LJ, Jeekel J, van Eijck CH (2006) The expression of adhesion molecules and the influence of inflammatory cytokines on the adhesion of human pancreatic carcinoma cells to mesothelial monolayers. Pancreas 32: 396–402.

64. Jonas L, Kruger B, Tessenow W (1993) Immunohistochemical detection of lactoferrin in different human glandular tissues with special reference to the exocrine pancreas. Acta Histochem 95: 53–59.

65. Kim JH, Ho SB, Montgomery CK, Kim YS (1990) Cell lineage markers in human pancreatic cancer. Cancer 66: 2134–2143.

66. Aust G, Steinert M, Schutz A, Boltze C, Wahlbuhl M, et al. (2002) CD97, but not its closely related EGF-TM7 family member EMR2, is expressed on gastric, pancreatic, and esophageal carcinomas. Am J Clin Pathol 118: 699–707.

67. Shi SQ, Cai JT, Yang JM (2006) Expression of trefoil factors 1 and 2 in precancerous condition and gastric cancer. World J Gastroenterol 12: 3119–3122.

68. Tobita K, Kijima H, Dowaki S, Oida Y, Kashiwagi H, et al. (2002) Thrombospondin-1 expression as a prognostic predictor of pancreatic ductal carcinoma. Int J Oncol 21: 1189–1195.

Editors' Summary

Background. Cancers are life-threatening, disorganized masses of cells that can occur anywhere in the human body. They develop when cells acquire genetic changes that allow them to grow uncontrollably and to spread around the body (metastasize). If a cancer is detected when it is still small and has not metastasized, surgery can often provide a cure. Unfortunately, many cancers are detected only when they are large enough to press against surrounding tissues and cause pain or other symptoms. By this time, surgical removal of the original (primary) tumor may be impossible and there may be secondary cancers scattered around the body. In such cases, radiotherapy and chemotherapy can sometimes help, but the outlook for patients whose cancers are detected late is often poor. One cancer type for which late detection is a particular problem is pancreatic adenocarcinoma. This cancer rarely causes any symptoms in its early stages. Furthermore, the symptoms it eventually causes—jaundice, abdominal and back pain, and weight loss—are seen in many other illnesses. Consequently, pancreatic cancer has usually spread before it is diagnosed, and most patients die within a year of their diagnosis.

Why Was This Study Done? If a test could be developed to detect pancreatic cancer in its early stages, the lives of many patients might be extended. Tumors often release specific proteins—"cancer bio-markers"—into the blood, a bodily fluid that can be easily sampled. If a protein released into the blood by pancreatic cancer cells could be identified, it might be possible to develop a noninvasive screening test for this deadly cancer. In this study, the researchers use a "proteomic" approach to identify potential biomarkers for early pancreatic cancer. Proteomics is the study of the patterns of proteins made by an organism, tissue, or cell and of the changes in these patterns that are associated with various diseases.

What Did the Researchers Do and Find? The researchers started their search for pancreatic cancer biomarkers by studying the plasma proteome (the proteins in the fluid portion of blood) of mice genetically engineered to develop cancers that closely resemble human pancreatic tumors. Through the use of two techniques called high-resolution mass spectrometry and acrylamide isotopic labeling, the researchers identified 165 proteins that were present in larger amounts in plasma collected from mice with early and/or advanced pancreatic cancer than in plasma from control mice. Then, to test whether any of these protein changes were relevant to human pancreatic cancer, the researchers analyzed blood samples collected from patients with pancreatic cancer. These

samples, they report, contained larger amounts of some of these proteins than blood collected from patients with chronic pancreatitis, a condition that has similar symptoms to pancreatic cancer. Finally, using blood samples collected during a clinical trial, the Carotene and Retinol Efficacy Trial (a cancer-prevention study), the researchers showed that the measurement of five of the proteins present in increased amounts at an early stage of tumor development in the mouse model discriminated between people with pancreatic cancer and matched controls up to 13 months before cancer diagnosis.

What Do These Findings Mean? These findings suggest that in-depth proteomic analysis of genetically engineered mouse models of human cancer might be an effective way to identify biomarkers suitable for the early detection of human cancers. Previous attempts to identify such biomarkers using human samples have been hampered by the many noncancer-related differences in plasma proteins that exist between individuals and by problems in obtaining samples from patients with early cancer. The use of a mouse model of human cancer, these findings indicate, can circumvent both of these problems. More specifically, these findings identify a panel of proteins that might allow earlier detection of pancreatic cancer and that might, therefore, extend the life of some patients who develop this cancer. However, before a routine screening test becomes available, additional markers will need to be identified and extensive validation studies in larger groups of patients will have to be completed.

Additional Information. Please access these Web sites via the online version of this summary at http://dx.doi.org/10.1371/journal.pmed. 0050123.

- The MedlinePlus Encyclopedia has a page on pancreatic cancer (in English and Spanish). Links to further information are provided by MedlinePlus
- The US National Cancer Institute has information about pancreatic cancer for patients and health professionals (in English and Spanish)
- The UK charity Cancerbackup also provides information for patients about pancreatic cancer
- The Clinical Proteomic Technologies for Cancer Initiative (a US National Cancer Institute initiative) provides a tutorial about proteomics and cancer and information on the Mouse Proteomic Technologies Initiative