A Migration Signature and Plasma Biomarker Panel for Pancreatic Adenocarcinoma

Seetharaman Balasenthil1*, Nanyue Chen1*, Steven T. Lott1*, Jinyun Chen2, Jennifer Carter3, William E. Grizzle4, Marsha L. Frazier2, Subrata Sen3, and Ann McNeill Killary1

1 Department of Genetics, The University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030, USA. 2 Department of Epidemiology, The University of Texas M.D. Anderson Cancer Center, Houston, Texas 77030, USA. 3 Department of Molecular Pathology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030, USA, 4 The Department of Pathology and The Comprehensive Cancer Center, The University of Alabama, Birmingham, Alabama.

*denotes equal contribution

Running title:

Pathways Approach Identifies Pancreatic Cancer Biomarkers

Key words: Pancreatic cancer, biomarkers.

Financial support:

This research was supported by grant U01CA111302 to AMK from the National Cancer Institute Early Detection Research Network.

Correspondence should be addressed to:

Ann McNeill Killary, Ph.D.
Department of Genetics
The University of Texas M. D. Anderson Cancer Center
Houston, Texas 77030
713-834-6395-telephone
713-792-1474-fax
(akillary@mdanderson.org).

Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Copyright © 2010 American Association for Cancer Research
ABSTRACT

Pancreatic ductal adenocarcinoma is a disease of extremely poor prognosis for which there are no reliable markers of asymptomatic disease. To identify pancreatic cancer biomarkers, we focused on a genomic interval proximal to the most common fragile site in the human genome, chromosome 3p12, which undergoes smoking-related breakage, loss of heterozygosity and homozygous deletion as an early event in many epithelial tumors, including pancreatic cancers. Using a functional genomics approach, we identified a seven gene panel (TNC, TFPI, TGFB1, SEL-1L, L1CAM, WWTR1 and CDC42BPA) that was differentially expressed across three different expression platforms, including pancreatic tumor/normal samples. In addition, IPA network analysis and literature searches indicated that this seven gene panel functions in one network associated with cellular movement/morphology/development, indicative of a ‘migration signature’ of the 3p pathway. We tested whether two secreted proteins from this panel, Tenascin C (TNC) and Tissue Factor Pathway Inhibitor (TFPI), could serve as plasma biomarkers. Plasma ELISA assays for TFPI/TNC resulted in a combined AUC of 0.88 and, with addition of CA 19-9, a combined AUC for the three gene panel (TNC/TFPI/CA19-9) was 0.99 with 100% specificity at 90% sensitivity and 97.22% sensitivity at 90% specificity. Validation studies using TFPI only in a blinded sample set increased the performance of CA19-9 alone from an AUC of 0.84 to 0.94 with the two gene panel. Results identify a novel 3p pathway associated migration signature and plasma biomarker panel that has utility for discrimination of pancreatic cancer from normal controls and promise for clinical application.
INTRODUCTION

Pancreatic cancer is the fourth leading cause of cancer-related mortality in both men and women in the United States. Estimates suggest that virtually 83% of the over 42,470 cases in the United States diagnosed with the disease will also die from it, making pancreatic cancer the most deadly of cancers when grouped by organ site (1). Biomarkers for the early detection of pancreatic cancer are urgently needed. However, individual molecular biomarkers with high sensitivity and specificity needed for population-based screening have not been discovered. CA-19-9 has been studied extensively and yet has failed to demonstrate the desired predictive value necessary for early detection and diagnosis (2). Although many high throughput discovery platforms, including proteomic, genomic and transcriptomic approaches have been utilized and biomarker candidates identified, no one platform or molecule has been successfully validated in large population screens. As a part of the National Cancer Institute Early Detection Research Network, our goal is to assemble a panel of blood-based biomarker candidates; given that no one biomarker has yet shown promise for early detection. We hypothesized that a panel of early detection biomarkers for pancreatic cancer could be discovered by the identification of the earliest genetic pathways aberrant in smoking related cancers, such as lung, renal, and pancreatic cancers. Evidence that the driver events in smoking related cancers remain to be discovered comes from the sequencing of the cancer genomes of 24 cases of pancreatic adenocarcinoma (3). Results indicated that although smokers have a significantly higher number of genetic alterations than do nonsmokers, including mutations, homozygous deletions and amplifications, smoking-related genetic alterations did not seem to correlate with known driver genes mutated in pancreatic cancer, including KRAS, p53, CDKN2A/p16, etc suggesting that the genetic determinants of smoking
related tumors are not driven by these major genes (3). Thus, we furthermore hypothesized that, in order to identify the earliest genetic alterations associated with pancreatic cancer, we must target the most common intervals of cytogenetic deletion associated with tobacco exposure given that these alterations might be shared by both smoking-related as well as non-smoking-related pancreatic cancer. Loss of chromosome 3p has been documented as an initiating event in a cytogenetic pathway involved in smoking-related cancers, thus, we chose a functional genomic approach to target the genetic pathway deregulated by the deletion of the chromosome 3p12 region. This genomic interval is known to undergo loss of heterozygosity and homozygous deletion in smoking related tumors and is proximal to the most common fragile site in the human genome, FRA3B, which has been shown to be expressed in active smokers and thought to facilitate chromosome breakage following carcinogenic exposure from tobacco (4). High frequency LOH has been observed in the FRA3B region encompassing 3p13-3p21 in a variety of epithelial cancers with homozygous deletions found proximal to the fragile site in lung, kidney, breast and pancreatic tumors, suggestive that deletion of this genetic interval could be an early event in the genesis of smoking related cancers (5). Our previous physical and functional mapping experiments demonstrated that the introduction of a normal copy of chromosome 3p into renal cell carcinoma cell lines via microcell fusion suppressed tumorigenicity in nude mice in both orthotropic and subcutaneous injection (6-8). Fine mapping of suppressed and unsuppressed hybrids localized the NRC-1 tumor suppressor locus to a 4.75 Mb interval within chromosome 3p12 (9). In addition, we have previously shown high frequency LOH in distinct intervals along 3p in malignant pancreatic islet cell tumors but not in precursor cystic lesions in a kindred with von Hippel Lindau disease, an autosomal dominant cancer syndrome characterized by the development of multiple tumor types, two of which progress to malignancy (renal and
pancreas), suggestive that loss of the 3p12 is requisite for malignant conversion (7). Experiments reported herein document the utility of a functional genomic approach not only to identify 3p12 pathway genes differentially expressed in pancreatic tumor/normal samples but also to determine their relevance as blood-based pancreatic cancer biomarkers.

**MATERIALS AND METHODS**

**Patients and clinical samples**

EDTA plasma samples from pancreatic adenocarcinoma cases were obtained from the TEXGEN repository, a Texas Medical Center consortium that houses sera and plasma from MDACC, Baylor and Texas Children’s hospital. Control plasma was obtained from individuals who were screened for different cancer types at MDACC and were free of malignancy or any benign condition. For validation purposes a blinded set of EDTA plasma samples from controls and pancreatic adenocarcinoma patients were also obtained from The University of Alabama, Birmingham (W.E.G).

**RNA extraction, microarray and Quantitative Realtime PCR analyses**

Frozen pancreatic tumors and adjacent macroscopically and microscopically normal appearing pancreas from the same patients (matched) were obtained from untreated, retrospective pancreatic adenocarcinoma samples available from the M. D. Anderson Cancer Center tumor bank and our collaborator (MLF). Total RNA was extracted from these samples using a miRNeasy Mini Kit (Qiagen, Valencia, CA). Microarray hybridization and scanning was performed according to Affymetrix protocols). Quantitative RT-PCR analysis was carried out as per manufacturer’s protocol (Applied Biosystems) using specific primers. Data were analyzed according to the
comparative Ct method and were normalized by glyceraldehyde-3-phosphate dehydrogenase expression.

**Suppression Subtractive Hybridization (SSH) Library**

The SSH library was previously constructed using as starting materials for library construction microcell hybrids formed by the introduction of defined fragments of a normal chromosome 3p into a renal cell carcinoma cell line and subsequent assay of those microcell hybrids for tumor formation in athymic nude mice (10). Methods utilized for library construction have been previously published and are shown schematically in Fig. 1A. We hypothesized that the resultant differentially expressed cDNAs obtained from the SSH library should represent genes up or downregulated by the introduction of the *NRC-1* tumor suppressor locus and could represent genes in a functional chromosome 3p12 tumor suppression pathway and conversely, identification of this pathway could elucidate how loss of this genomic region and deregulation of this pathway could be involved in the early stages of pancreatic cancer. We furthermore hypothesized that characterization of this library could define genetic networks in pancreatic cancer that could serve as a source for biomarkers for early detection.

**Bioinformatic analyses**

To generate the highest quality expression data, the PDNN (positional-dependent-nearest-neighbor) model was chosen to account for existing probe variation in specific binding with the labeled target material (11). Existing algorithms, such as MAS 5.0, do not take into account probe specific variation in binding efficiency and can result in variation in probe signal that vary over two orders of magnitude within a single probeset. Probe normalization and summarization was performed using the PerfectMatch software suite utilizing the PDNN algorithm.
For each probeset, the software outputs both the natural logarithm transformed expression level and correlation coefficient between the observed and modeled data. Data was filtered to identify genes exhibiting greater than or equal to 2 fold changes in gene expression.

**Network and gene ontology analysis**

Selected genes were investigated for network and gene functional interrelation by Ingenuity Pathways Analysis (IPA) software (Ingenuity Systems, [www.ingenuity.com](http://www.ingenuity.com)) (12). IPA scans the set of input genes to identify networks by using Ingenuity Pathways Knowledge Base for interactions between identified ‘Focus Genes’, in our case, the common genes identified from our pathways approach and known and hypothetical interacting genes stored in the knowledge base in IPA software, to generate a set of networks with a maximum network size of 35 genes/proteins. Networks are displayed graphically as genes/gene products (‘nodes’) and the biological relationships between the nodes (‘edges’). All edges are from canonical information stored in the Ingenuity Pathways Knowledge Base. In addition, IPA computes a score for each network according to the fit of the user’s set of significant genes. The score indicates the likelihood of the Focus Genes in a network from Ingenuity’s knowledge base being found together due to random chance. A score of 3, as the cutoff for identifying gene networks, indicates that there is only a 1/1000 chance that the focus genes shown in a network are due to random chance. Therefore, a score of 3 or higher indicates a 99.9% confidence level to exclude random chance.

**Large Tenascin-C ELISA**

Plasma levels of the predominant isoform of TNC, TNC-large variant (TNC-L) were determined using a Human Tenascin-C Large (HMV) (FNIII-B) ELISA kit (IBL-America, Minneapolis, MN), which detects human TNC high molecular weight variant by sandwich ELISA.
samples were diluted 100-fold and then incubated in 96-well ELISA plates precoated with anti-TNC (4C8MS) Ab for 1 hour at 37°C. After washing the wells 7 times with wash buffer, a horseradish peroxidase-conjugated anti-TNC (4F10TT) antibody was added and incubated for 30 minutes at 4°C. After washing the wells 9 times with wash buffer, chromogen was added and incubated at room temperature in dark for 30 min. The reaction was stopped by the addition of stop solution and the absorbance at 450 nm was determined using an ELISA plate reader (Spectramax Plus Microplate Reader, Molecular Devices, Sunnyvale, CA) within 30 minutes of addition of stop solution, with the correction wavelength set at 540nm. Results were mean absorbance of duplicate wells.

**TFPI ELISA**

Plasma TFPI levels were determined using a commercially available Quantikine Human TFPI ELISA kit (DTFP10) (R&D Systems, Inc. Minneapolis, MN), which detects predominantly free TFPI and a very small percentage of LDL and HDL-bound TFPI by sandwich ELISA. The samples were diluted 200-fold and then incubated along with an assay diluent in 96-well ELISA plates precoated with anti-human TFPI monoclonal antibody for 2h at room temperature. After washing the wells 4 times with wash buffer, a horseradish peroxidase-conjugated anti-human TFPI was added and incubated for 1h at room temperature. After washing the wells 4 times with wash buffer, chromogen with substrate was added and incubated at room temperature in dark for 30 min. The reaction was stopped by the addition of stop solution and the absorbance at 450 nm was determined using an ELISA plate reader (Spectramax Plus Microplate Reader, Molecular Devices, Sunnyvale, CA) within 30 minutes of addition of stop solution, with the correction wavelength set at 540nm. Results were mean absorbance of duplicate wells.
CA 19-9 ELISA

CA 19-9 levels were measured in plasma samples (10μl) using a commercially available ELISA kit (DRG International Inc, Mountainside, NJ) according to manufacturer’s instructions. 10 μl of samples were incubated along with an assay buffer in 96-well ELISA plates precoated with murine monoclonal anti-CA 19-9 antibody for 90 minutes at 37°C. After washing the wells 5 times with wash buffer, a horseradish peroxidase-conjugated anti-CA 19-9 was added and incubated for 90 minutes at 37°C. After washing the wells 5 times with wash buffer, chromogen with substrate was added and incubated at room temperature in dark for 20 min. The reaction was stopped by the addition of stop solution and the absorbance at 450 nm was determined using an ELISA plate reader (Spectramax Plus Microplate Reader, Molecular Devices, Sunnyvale, CA) within 15 minutes of addition of stop solution. Results were mean absorbance of duplicate wells. For these experiments, we did not assign predetermined cut-off values to assess the specificity and sensitivity.

Statistics

Differences in plasma levels between normal and pancreatic adenocarcinoma were analyzed using the Student’s t test. To provide additional statistical rigor, the Mann-Whitney U test was also used to analyze the difference between normal and pancreatic adenocarcinoma samples. Two-sided P-values less than 0.05 were considered statistically significant. We constructed receiver operating characteristic curves (ROC) and calculated the area under the curve (AUC) to evaluate the specificity and sensitivity of predicting cases and controls by each protein and by the combination of these proteins. All statistical analyses were performed using the Stata 10.1 (Stata Corporation, College Station, TX).
RESULTS

Suppression Subtractive Hybridization (SSH) Library Identifies Candidate Chromosome 3p12 Pathway Genes. Three different expression platforms were utilized to identify genes in the chromosome 3p12 pathway to tumorigenesis in pancreatic cancer (schematically illustrated in Fig. 1A). We previously constructed an SSH library using as starting materials for library construction, microcell hybrid clones containing defined fragments of chromosome 3p12 which were either suppressed or unsuppressed for tumorigenicity following injection of microcell hybrid clones into athymic nude mice (10). cDNAs differentially expressed from this SSH library should represent genes up or downregulated by the 3p12 tumor suppressor locus and therefore represent a 3p12 downstream pathway for biomarker discovery. From the SSH library, 880 partial cDNAs were obtained that were differentially expressed between suppressed and unsuppressed hybrids. PCR products were obtained from 763 clones (87% of the library) and used as templates for sequencing. 569/763 clones (75%) were identified as subtracted due to the presence of appropriate adaptors and BLAST searches against the RefSeq database indicated 117 clones had no matches, short or poor quality sequence or chimerism. The remainder of the 452/763 clones produced 2297 sequences; however, a number of duplications were present in the
library. After filtering for redundancy and for only those sequences with Entrez Gene matches, 507 Entrez Gene matches were obtained (Table S1, Supplementary Appendix).

**Expression Profiling Validates SSH Library.** In order to validate differential expression observed in the sequenced cDNA clones obtained from the SSH library, we utilized a second expression platform. Following analysis using the PerfectMatch software, probesets overlapping the 507 Entrez Gene sequences from the SSH library were identified on a GeneChip U133 plus 2.0 array (Affymetrix, Inc.) and examined for differential expression by interrogating the array with the starting materials for construction of the SSH library i.e., microcell hybrids containing fragments of 3p used to construct the SSH library. By screening the same cDNAs identified from the SSH library on a commercial array, we were able to directly compare expression profiles across platforms to identify differentially expressed sequences which we subsequently filtered by bioinformatic analyses to identify those genes with expression differences of ≥ 2 fold. Supplementary Table S2 illustrates the 82 gene list of chromosome 3p12 pathways genes identified by this approach.

**Expression Analyses of Pancreatic Tumor/Normal Samples Provides a Third Platform to Stratify Data.** In order to identify those genes of the 82 gene list that were relevant to pancreatic cancer, we identified probesets on a U133A plus 2.0 array corresponding to the 82 gene set, and then interrogated the array with pancreatic tumor/normal samples. Frozen tumor and adjacent macroscopically and microscopically normal appearing pancreatic tissues from the same patient (matched) were obtained from untreated, retrospective pancreatic adenocarcinoma samples available from the M. D. Anderson Cancer Center tumor bank and our collaborator (MLF). The characteristics of the patients used for this study are presented in Supplementary Table S3. Total
RNA from matched tumor/adjacent normal samples (8 paired samples) was utilized to interrogate the array and resultant bioinformatic analysis performed to identify differentially expressed sequences across all three platforms. Bioinformatic analysis involved the stratification of the data set by selecting only those genes from the 82 gene list that were significantly differentially expressed across all three expression platforms resulting in a 7 gene set, p<0.05 (Table 1). The seven gene panel [WWTR1 (13,14), TGFBI (15,16), TFPI (17,18), CDC42BPA (19,20), LICAM, (21,22), TNC (23,24) and SEL1L (25,26)] demonstrated at least greater than or equal to 2 fold difference in gene expression and consistently remained as top candidates that were differentially expressed across three platforms. One gene, SEL1L, was significantly downregulated in 8/8 tumors as compared with normal adjacent tissue, while the other 6 genes were up-regulated. Importantly, 5 out of 7 genes had also been previously published as being differentially expressed by immunohistochemistry in pancreatic tumor samples. In addition, we screened SEL1L and TNC by quantitative RT-PCR in 26 matched pancreatic tumor/normal samples to confirm differential expression (p values of 0.002 and 0.038 respectively. (Supplementary Table 4).

As another validation of our 7 gene biomarker panel, we performed in silico analyses of publicly available microarray data sets from Oncomine (http://www.oncomine.org) to determine if the 7 genes identified by our functional genomic screen were also found as differentially expressed in published pancreatic tumor/normal expression screens. Representative results of an individual data set (27), in which all our 7 gene biomarker gene panel are expressed are presented in (Supplementary Figure S1). Analysis of the biomarker panel demonstrated that the expression of six genes [TNC, TFPI, TGFBI, LICAM, CDC42BPA and WWTR1] were found to be significantly upregulated in pancreatic cancer when compared with normal pancreatic tissue.
In contrast, SEL-1L was found to be significantly downregulated in pancreatic adenocarcinomas compared to the normal pancreas (Supplementary Figure S1). Furthermore, we also analyzed the expression of our panel in three different expression datasets obtained from microdissected pancreatic tumor/normal samples (28-30). Importantly, although none of the 3 databases used listed all 7 genes as differentially expressed, subsets of each of the 7 gene panel were represented in these datasets as well (Supplementary Figures S2 and S3). Thus, insilico analyses has confirmed out studies using cross platform functional approaches although none of the previously published compendiums of expression profiles identified these genes as a panel or studied their potential as blood-based pancreatic cancer biomarkers. Thus, our functional approach identified a novel panel differentially expressed in multiple datasets which hitherto had not been studied for blood-based biomarker development.

**Ingenuity Pathway Analysis Identifies a Single Network and Migration Signature for 3p Pathway Genes.** To determine the functional relationships among the seven genes confirmed by our functional genomic pathway studies, Ingenuity Pathway Analysis (IPA) was queried for known or hypothetical interactions among the seven genes in the panel as well as all other genes in the Ingenuity database. With the exception of WWTR1/TAZ which was not present in the IPA database, all the other 6 genes were used as focus genes for IPA. Unsupervised IPA network analysis identified a network of 35 genes that included all six focus genes and 29 additional genes (Score =16) (Table 2). The interactive relationship between the genes in the network are shown in Fig.1B. Importantly, all 6 genes were classified into a single network related to cellular movement, cell morphology and cellular development (Table 2). Of the 6 genes, 5 also were a part of a network involving cell signaling and cell interaction (p=7.07E-06~3.07E-02) and cell movement (p=1.38E-04~3.51E-02). The extremely low probability of obtaining this number of
differentially expressed molecules in one network by chance alone is reflected by the p value for the network (p=1.0E-16), indicating that this network is deregulated in a highly significant, non-random manner in pancreatic cancer cells. In addition, WWTR1/TAZ has also been reported to function in the regulation of cell migration in breast cancer (13). Therefore, we conclude that our functional genomic pathway approach has identified a gene signature related to cell movement, morphology and organization, suggestive that the loss of the 3p12 locus in pancreatic cancer could be related to loss of polarity and aberrant migration associated with early events in malignant transformation of pancreatic ductal epithelial cells (Table 2).

TNC and TFPI are Candidate Plasma Biomarkers that Distinguish Pancreatic Cancer from Normal Screening Controls. Three of the seven gene panel (TGFBI, TFPI, and TNC) were also secreted proteins. Sandwich ELISA assays were then performed on two of the three secreted proteins, TFPI and TNC, for which commercial ELISAs were available to determine their ability to function as plasma biomarkers. The patient population characteristics used in the present study with respect to age, sex, alcohol intake history, smoking history, diabetic history, site of the disease, staging and survival data are presented in Table 3. Results indicated that individual plasma TNC-L levels, presented in the form of a scatter plot (Fig. 2A), were significantly different between pancreatic cancer patients versus normal screening controls in that the median plasma TNC-L levels was 342.6 pg/ml in patients with pancreatic adenocarcinoma (n=36) as compared to levels in normal subjects (243.3) (n=19) (Student’s t test, P=0.0006; Mann-Whitney’s U test, P=0.0004). Fig. 2B illustrates the receiver operator characteristic (ROC) curve for TNC, and the area under curve (AUC) was 0.79, with a specificity of 47% at 90% sensitivity and sensitivity of 25% at 90% specificity.
We next extended our study to further validate the significance of TFPI as a potential plasma biomarker by ELISA. Fig. 2C depicts the individual plasma TFPI levels of normal and pancreatic cancer patients. Plasma TFPI levels of patients with pancreatic adenocarcinoma were significantly higher compared with normal subjects (Student’s t test, \( P=0.0004 \); Mann-Whitney’s U test, \( P<0.0001 \)), with the median plasma TFPI level of 27.0 ng/ml in pancreatic adenocarcinoma patients (n=36) compared to normal subjects (15.3) (n=19). Fig. 2D illustrates the ROC curve to compare the ability of plasma TFPI in distinguishing patients with pancreatic cancer versus normal subjects. The AUC for TFPI was 0.87 with a specificity was 63% given 90% sensitivity and a sensitivity of 64% given 90% specificity.

We next tested whether the combination of the two markers, TNC and TFPI, could increase sensitivity and specificity for discrimination between cancer and normal plasma samples. The combined AUC for both markers was 0.88 (Fig. 2E). The combination of markers resulted in a specificity of 63% given 90% sensitivity and a sensitivity of 67% given 90% specificity. These combined results, then, suggest that the two gene panel identified through our functional genomic studies has high sensitivity and specificity to discriminate tumor and normal samples in the plasma and indicate that these genes are candidate blood-based biomarkers for pancreatic cancer. The diagnostic potential of candidate biomarkers TNC and TFPI, relative to and in combination with CA19-9, the standard serum biomarker for pancreatic cancer, was determined. Results indicated that individual plasma CA 19-9 levels, presented in the form of a scatter plot (Fig. 2F), were significantly different between pancreatic cancer patients versus normal screening controls in that the median plasma CA 19-9 levels was 173 U/ml in patients with pancreatic adenocarcinoma (n=36) as compared to levels in normal subjects (11.9) (n=19) (Student’s t test,
P<0.00001; Mann-Whitney’s U test, P <0.00001). Fig. 2G illustrates the ROC curve for CA 19-9, and the AUC was 0.93, with a specificity of 94.74% at 90% sensitivity and sensitivity of 91.67% at 90% specificity. Fig. 2H illustrates the ROC curve for CA 19-9 and TFPI combined, and the AUC was 0.99, with a specificity of 100% at 90% sensitivity and sensitivity of 97.22% at 90% specificity.

We next tested whether the combination of the two markers TNC and TFPI with CA 19-9 could further increase sensitivity and specificity for discrimination between cancer and normal plasma samples. The combined AUC for all the markers was 0.99 (Supplementary Figure S4). The combination of markers resulted in a specificity of 100% given 90% sensitivity and a sensitivity of 97.22% given 90% specificity.

To further strengthen our findings and to validate our results in a different sample set, we tested TFPI levels by ELISA in plasma samples from normal and pancreatic adenocarcinoma patients collected at the University of Alabama (UAB), Birmingham. A description of the UAB sample set is given in Table 3). Since the addition of TNC did not add significantly to the sensitivity and specificity (Supplementary Figure S4), we analyzed only TFPI levels in these samples. Fig. 3A depicts the individual plasma TFPI levels of normal and pancreatic cancer patients. Plasma TFPI levels of patients with pancreatic adenocarcinoma were significantly higher compared with normal subjects (Student’s t test, P= 0.000014; Mann-Whitney’s U test, P=0.0001), with the median plasma TFPI level of 45.7 ng/ml in pancreatic adenocarcinoma patients (n=37) compared to normal subjects (25.6) (n=15). Fig. 3B illustrates the ROC curve to compare the ability of plasma TFPI in distinguishing patients with pancreatic cancer versus normal subjects. The AUC for TFPI was 0.87 with a specificity of 46.67% at 90% sensitivity and 70.27 % sensitivity at 90% specificity.
Next we analyzed CA 19-9 by ELISA in these samples and results indicated that individual plasma CA 19-9 levels, presented in the form of a scatter plot (Fig. 3C), were significantly different between pancreatic cancer patients versus normal screening controls in that the median plasma CA 19-9 levels was 171.2 U/ml in patients with pancreatic adenocarcinoma (n=37) as compared to levels in normal subjects (15.7) (n=15) (Student’s t test, $P=0.0000000400227$; Mann-Whitney’s U test, $P=0.0001$). Fig. 3D illustrates the receiver operator characteristic (ROC) curve for CA 19-9, and the area under curve (AUC) was 0.84, with a specificity of 13.33% at 90% sensitivity and 75.68% sensitivity at 90% specificity. The combined AUC for TFPI and CA 19-9 was 0.94 (Fig. 3E). The combination of markers resulted in a specificity of 86.67% at 90% sensitivity and 83.78% sensitivity at 90% specificity.

**DISCUSSION**

Attempts to identify pancreatic cancer biomarkers have failed to produce a single marker with the sensitivity and specificity necessary for population screening. We reasoned that a targeted strategy to identify differentially expressed genes related to the earliest cytogenetic aberrations might be more successful in developing such biomarkers because we would be able to focus on those aberrantly expressed genes, which may be involved in initiating the pathways that ultimately lead to tumorigenesis, invasion and metastasis. Using three different expression platforms, we identified a 7 gene set as being differentially expressed between pancreatic cancer and normal samples and from which we have validated a subset of markers for differential expression by ELISA assays. Our results indicate that we have been able to identify two relevant
blood-based biomarker candidates for pancreatic cancer. Several published reports document the potential of using IHC staining of TNC, an extra-cellular matrix protein, as a potential marker of early disease as well as a predictor of poor prognosis in several tumor types, including colon, bladder and pancreas (31-34). TNC has been shown to be overexpressed in the stroma by IHC in a variety of different cancers, including pancreatic cancer (31-33, 35). In addition to stromal expression, TNC expression increases from low grade PanIN-1A and 1B intraductal precursor lesions to high grade PanIN -2 and 3 lesions to invasive lesions, suggestive that TNC could be an early IHC marker of disease (31). However, reports of the utility of TNC as a blood-based biomarker have been limited to colorectal cancer (36), where TNC spliced variant overexpression in plasma was observed and is considered a potential biomarker for colorectal cancer. Our data implicate TNC as a potential plasma marker for pancreatic cancer which based on its expression in precursor lesions and upregulation with increasing stage suggests that it might be a marker of early disease.

Our results also identify TFPI as a novel pancreatic plasma biomarker candidate. TFPI is a major inhibitor of the tissue factor pathway of blood coagulation in vivo. Plasma TFPI levels have been shown to increase significantly at the time of diagnosis compared to controls and reach near normal levels following surgical removal of pancreatic tumor (37). Plasma TFPI has also been reported to be increased in acute pancreatitis compared to normal subjects (38). Plasma levels of total TFPI has been found to be upregulated in a number of solid tumors involving colon, pancreas and stomach (39). Recently, increased TFPI expression levels have also been reported in colon and breast tumor tissues (40). Our study suggests that plasma TFPI levels may be a potential biomarker for pancreatic cancer and can also serve as a prognostic marker owing to its modulation following surgery. Our results indicate that the combined analysis of TFPI and CA
19-9 in plasma can discriminate pancreatic adenocarcinoma patients from normal screening controls with better sensitivity and specificity than CA19-9 alone, at least in our pilot studies. These results could have significance for the early screening of pancreatic cancer given that CA19-9 alone fails to have adequate predictive value. Since CA19-9 is typically a serum marker, these results would furthermore suggest that future analyses of the value of CA19-9 screening in plasma as well as its use in combination with TFPI are warranted to determine if the combined panel could result in a viable test for general population screening.

Furthermore, using a functional genomic approach, we have identified cancer-associated network associated with the differential expression of the chromosome 3p12 locus implicated in smoking related malignancies. Ingenuity pathway analyses identified cell movement/cell morphology/cell differentiation as the single critical network associated with the 3p12 pathway with 7/7 genes in the panel functionally implicated previously in the regulation of cell movement and migration. Thus, although the importance of migration has been thoroughly characterized in relationship to metastasis, the role of cell movement and migration in early stages of cancer initiation is not well understood. Since the chromosome 3p12 region has been shown to undergo LOH, homozygous deletion as an early event in smoking related cancers, potentially, loss of this region could play a role in the loss of polarity, cell migration and movement in the earliest stages of malignant transformation and epithelial-mesenchymal-transition related to invasive disease. We have earlier characterized one novel gene \textit{DEAR1} (annotated as TRIM62) in detail from our SSH library. Genetic complementation of \textit{DEAR1} in a breast cancer cell line carrying a \textit{DEAR1} mutation, resulted in restoration of acinar morphogenesis while knockdown resulted in loss of polarity and tissue architecture in 3D culture (10). Thus, DEAR1 falls within this 3p pathway and functions in the regulation of cellular morphology and differentiation as it relates to changes in 3D acinar
morphogenesis. In addition, other members of the 7 gene panel have also been documented to play a role in cell migration and early stages of pancreatic cancer, suggestive that this network may be critically deregulated in early pancreatic cancer. TNC has been shown to mediate proliferation and migration of astrocytes in a wound assay (41). Tumor associated isoforms of TNC has been shown to promote breast cancer cell invasion and growth (42). TNC signaling has been reported to play an important role in mammary tumor growth and metastasis and knockdown of TNC exhibited significant impairment in cell migration and anchorage-dependent cell proliferation in breast cancer cell line (43). In addition TNC has also been shown to stimulate glioma cell migration (44). TFPI has been shown to control migration of endothelial cells (45). Extracellular matrix bound TFPI through an interaction with tissue factor/VIIa complex localized on cancer cells has also been shown to facilitate cancer cell migration and adhesion (18).

SEL-1L, is the only member of the 7 gene panel shown to be downregulated in pancreatic tumors versus normal samples this study. SEL-1L is expressed abundantly only in the normal pancreas and is downregulated in pancreatic cancers. SEL-1L was first reported as a pancreas specific transcript but later found to be highly expressed in normal pancreas and present at very low levels in several other adult tissues (25,46,47). SEL-1L loss of expression has been observed in 17% of pancreatic adenocarcinomas. Thus, SEL-1L represents a gene that shows a tissue-restricted pattern of expression with the only tissue demonstrating abundant expression being the pancreas. In addition, SEL-1L maps into a genomic interval for the insulin-dependent diabetes mellitus locus at chromosome 14q24.3-q31 but was later excluded as a candidate gene for diabetes (48). Induced expression of SEL-1L in pancreatic cancer cells has been shown to decrease the clonogenity and anchorage-independent growth and it also delayed tumor growth in immunodeficient mice (49). In addition SEL-1L has been reported to affect pancreatic cancer
cell cycle and invasiveness by modulating the expression of PTEN and genes involved in cell-matrix interactions (26). SEL-1L has been shown to be a negative regulator of Notch signaling. The Notch pathway has been extensively studied and regulates cell fate decisions in a large number of adult and embryonic tissues. Components of the Notch signaling pathway have been shown to be overexpressed in pancreatic adenocarcinomas (50), with activation of Notch signaling observed in PanIN lesions. Thus, SEL1L loss of expression could represent a very early marker of pancreatic cancer. L1CAM has been reported to be overexpressed in a number of different tumors types including colon, breast, ovarian, melanoma, gliomas, neuroblastomas and pancreatic neuroendocrine tumors (21). Immunohistochemical staining of L1CAM was observed in chronic pancreatitis tissues and was absent in normal pancreatic tissues (51). Importantly, L1CAM has been shown to play a role in migration and in the malignant transformation of pancreatic adenocarcinoma (51). Importantly, upregulation of L1CAM expression by IHC has been observed in later stage, high grade PanIN lesions as compared with Pan IN 1A/B lesions which are not thought to have a high risk for progression to pancreatic cancer, suggestive of the role of L1CAM early in transition to pancreatic adenocarcinoma (52). Interestingly WWTR1/TAZ, a transcription cofactor, was also found to regulate cell migration and invasion (13, 14). WWTR1 has also been reported to be amplified in pancreatic cancer cell lines and in pancreatic cancer (53). WWTR1 was found to play a role in the migration, invasion and tumorigenesis of breast cancer cells (13). TGFBI has been reported to be overexpressed in colon and pancreatic cancer (16, 54). TGFBI is an excreted extracellular matrix protein reported to play a role in cell-matrix regulation as well as cell migration in bone (55). TGFBI has recently been found as one of a gene panel upregulated during hematopoietic stem cell lineages as they differentiated and became migratory, suggesting a role for TGFBI in stem cell migration between
niches (56). CDC42BPA, a protein kinase has also been implicated in tumor cell invasion (20). CDC42BPA in complex with a leucine-rich adaptor protein LRAP35a, and MYO18A has been shown to play a crucial role in cell protrusion and migration (57). Therefore, all 7 genes in our panel have been closed linked functionally to the control of cell migration in cancer and potentially in the early stages of pancreatic tumorigenesis.

In addition, three of the seven genes identified as most differentially expressed, L1CAM, TGFBI and CDC42BPA are mutated in the germline in genetic disorders including CRASH syndrome, Thiel Behnke corneal dystrophy and Crohn’s disease respectively (58,59). Given that germline mutations underlying genetic disorders are very rare, the finding that our study identified 3/7 genes as being mutated in hereditary diseases indicates the functional significance of this migration pathway in early development, the deregulation of which could be of critical importance in pancreatic cancer initiation and progression.

In conclusion, we have taken a pathways approach to biomarker discovery by utilizing three different expression based platforms to identify chromosome 3p12 pathway genes differentially expressed between pancreatic tumor/normal samples which could serve as candidate biomarkers for the early detection of pancreatic cancer. Biomarker panels described herein will be further validated in larger case-control studies with the Early Detection Research Network (EDRN) of the National Cancer Institute. Additional candidates from the 7 gene list and associated IPA network members should also be investigated for their ability to improve performance of current panels as well. Future studies are also warranted to investigate the role of cell polarity and migration in the initiation of pancreatic cancer and the potential for biomarker discovery by a targeted pathway approach.
ACKNOWLEDGMENTS

We gratefully acknowledge the NCI Early Detection Research Network for support and guidance in our ongoing biomarker discovery projects. We are also grateful to Dr. David Gold and Dr. David Stivers who provided bioinformatics support for this project.

REFERENCES


22. Rathjen FG, Schachner M. Immunocytological and biochemical characterization of a new neuronal cell surface component (L1 antigen) which is involved in cell adhesion. EMBO J 1984;3:1-10.


<table>
<thead>
<tr>
<th>Symbol</th>
<th>Title</th>
<th>Major Biological Functions</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>WWTR1/TAZ</td>
<td>WW domain containing transcription regulator 1</td>
<td>Cofactor of transcription, cell migration, EMT.</td>
<td>(13,14)</td>
</tr>
<tr>
<td>TGFBI</td>
<td>Transforming growth factor, beta-induced, 68kDa</td>
<td>Cell adhesion, migration, cell matrix interaction</td>
<td>(15;16)</td>
</tr>
<tr>
<td>TFPI</td>
<td>Tissue factor pathway inhibitor</td>
<td>Cell adhesion, migration and proliferation</td>
<td>(17;18)</td>
</tr>
<tr>
<td>CDC42BPA</td>
<td>CDC42 binding protein kinase alpha (DMPK-like)</td>
<td>Cell morphogenesis, Cell signaling</td>
<td>(19;20)</td>
</tr>
<tr>
<td>L1CAM</td>
<td>L1 cell adhesion molecule</td>
<td>Cell morphogenesis, migration and cell survival</td>
<td>(21;22)</td>
</tr>
<tr>
<td>TNC</td>
<td>Tenascin C (hexabrachion)</td>
<td>Cell adhesion, migration, proliferation, and angiogenesis</td>
<td>(23;24)</td>
</tr>
<tr>
<td>SEL1L</td>
<td>Sel-1 suppressor of lin-12-like</td>
<td>Negative regulation of colony formation, growth and invasion; Cell-matrix interaction</td>
<td>(25;26)</td>
</tr>
</tbody>
</table>
### Table 2. Networks Identified from Ingenuity Pathway Analyses

<table>
<thead>
<tr>
<th>Network</th>
<th>Genes in Ingenuity Network*</th>
<th>Functions</th>
<th>Score**</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ACAN, ALCAM, CAT, CD47, CDC42, <strong>CDC42BPA</strong>, CDC42BPB, CHI3L1, CR1, DKK1, FDXR, GCH1, <strong>LICAM</strong>, LIMK2, MMP19, NCAN, NFkB (complex), PRRX1, PTHLH, PTK2, PTPRZ1, SAA@, SDC4, <strong>SEL1L</strong>, SNAP91, SQLE, SYVN1, <strong>TFPI</strong>, TGFβ3, <strong>TGFBI</strong>, TNC, TNF, TP53, TPM1, VCAN</td>
<td>Cellular Movement, Cell Morphology, Cellular Development</td>
<td>16</td>
</tr>
</tbody>
</table>

* Genes in boldface were identified from our functional genomic pathway approach as differentially expressed across three platforms as focus genes; additional genes listed were identified by IPA. ** A score > 3 is considered significant.
FIGURE LEGENDS

Fig. 1. A, Schematic flow diagram of functional genomic approach to identify chromosome 3p12 pathway genes. Three expression platforms are denoted that were screened to identify candidates for validation in plasma as potential pancreatic cancer biomarkers. B, Putative gene network derived from Ingenuity Pathway analysis software. Edges are displayed with labels that describe the nature of the relationship between the nodes. The lines between genes represent known interactions, with solid lines representing direct interactions and dashed lines representing indirect or hypothetical interactions. Color highlighting (red-upregulation; green-downregulation) indicates pathway associated genes discovered by this study and non-highlighted genes are those identified by IPA.

Fig. 2. Candidate biomarkers from functional genomic approach validated in plasma of pancreatic cancer patients versus controls. A, Plasma L-Tn-C concentrations of pancreatic carcinoma patients and normal subjects. Line, median plasma Tn-C level. The difference between normal and pancreatic adenocarcinoma samples are statistically significant (Mann-Whitney’s U test, \( P=0.0004 \)). B, Receiver operating characteristic curve for differentiating between normal and pancreatic carcinoma patients based on the plasma Tn-C ELISA assay. The area under the curve (AUC) was 0.79. The specificity was 47% given 90% sensitivity and the sensitivity was 25% given 90% specificity; C, Plasma TFPI concentrations in pancreatic carcinoma patients and normal subjects. Line, median plasma TFPI level. The difference between normal and pancreatic adenocarcinoma samples are statistically significant (Mann-Whitney’s U test, \( P<0.0001 \)). D, Receiver operating characteristic curve for differentiating between normal and pancreatic
carcinoma patients based on the plasma TFPI ELISA assay. The area under the curve (AUC) was 0.87. The specificity was 63% given 90% sensitivity and the sensitivity was 64% given 90% specificity; E, Receiver operating characteristic curve for differentiating between normal and pancreatic carcinoma patients based on the combinations of two markers, plasma TNC and TFPI ELISA. The combined AUC is 0.88. F, Plasma CA 19-9 concentrations of pancreatic carcinoma patients and normal subjects. Line, median plasma CA19-9 level. The difference between normal and pancreatic adenocarcinoma samples are statistically significant (Mann-Whitney’s U test, $P=0.0001$). G, Receiver operating characteristic curve for differentiating between normal and pancreatic carcinoma patients based on the plasma CA 19-9 ELISA assay. The area under curve (AUC) was 0.93, with a specificity of 94.74% at 90% sensitivity and sensitivity of 91.67% at 90% specificity; H, Receiver operating characteristic curve for differentiating between normal and pancreatic carcinoma patients based on the combinations of two markers, plasma CA 19-9 and TFPI ELISA. The combined AUC is 0.99.

**Fig. 3.** A, Plasma TFPI concentrations of pancreatic carcinoma patients and normal subjects. Line, median plasma TFPI level. The difference between normal and pancreatic adenocarcinoma samples are statistically significant (Mann-Whitney’s U test, $P=0.0001$). B, Receiver operating characteristic curve for differentiating between normal and pancreatic carcinoma patients based on the plasma TFPI ELISA assay. The area under curve (AUC) was 0.87 with a specificity of 46.67% at 90% sensitivity and 70.27% sensitivity at 90% specificity. C, Plasma CA 19-9 concentrations of pancreatic carcinoma patients and normal subjects. Line, median plasma CA19-9 level. The difference between normal and pancreatic adenocarcinoma samples are statistically significant (Mann-Whitney’s U test, $P=0.0001$). D, Receiver operating characteristic.
curve for differentiating between normal and pancreatic carcinoma patients based on the plasma CA 19-9 ELISA assay. The area under curve (AUC) was 0.84, with a specificity of 13.33% at 90% sensitivity and 75.68% sensitivity at 90% specificity. E, Receiver operating characteristic curve for differentiating between normal and pancreatic carcinoma patients based on the combinations of two markers, plasma CA 19-9 and TFPI ELISA. The combined AUC was 0.94 and resulted in a specificity of 86.67% at 90% sensitivity and 83.78% sensitivity at 90% specificity.
Table 3. Characteristics of samples used in the study

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Test set (TEXGEN)</th>
<th>Validation set (UAB)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of patients</td>
<td>No. of controls</td>
</tr>
<tr>
<td></td>
<td>(n = 36)</td>
<td>(n = 19)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>20</td>
<td>11</td>
</tr>
<tr>
<td>Female</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>Age (Y)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 50</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>50-60</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>61-70</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>71-80</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>81-90</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>36</td>
<td>19</td>
</tr>
<tr>
<td>Black</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Histology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>36</td>
<td>0</td>
</tr>
<tr>
<td>Alcohol History</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>10</td>
<td>N/A</td>
</tr>
<tr>
<td>Former</td>
<td>14</td>
<td>N/A</td>
</tr>
<tr>
<td>Never</td>
<td>12</td>
<td>N/A</td>
</tr>
<tr>
<td>Smoking history</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>10</td>
<td>N/A</td>
</tr>
<tr>
<td>Former</td>
<td>13</td>
<td>N/A</td>
</tr>
<tr>
<td>Never</td>
<td>13</td>
<td>N/A</td>
</tr>
<tr>
<td>Diabetes history</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>7</td>
<td>N/A</td>
</tr>
<tr>
<td>No</td>
<td>29</td>
<td>N/A</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>II</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>III</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>IV</td>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>Not available (N/A)</td>
<td>17</td>
<td>-</td>
</tr>
<tr>
<td>Site</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td>Head</td>
<td>21</td>
<td>-</td>
</tr>
<tr>
<td>Pancreas overlapping lesion</td>
<td>7</td>
<td>-</td>
</tr>
<tr>
<td>Tail</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct extension</td>
<td>9</td>
<td>-</td>
</tr>
<tr>
<td>Direct extension + lymph node</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>Distant</td>
<td>17</td>
<td>-</td>
</tr>
<tr>
<td>Localized</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Regional lymphnode involvement</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Not available</td>
<td>4</td>
<td>-</td>
</tr>
</tbody>
</table>
A Migration Signature and Plasma Biomarker Panel for Pancreatic Adenocarcinoma


Cancer Prev Res Published OnlineFirst November 11, 2010.

Updated version
Access the most recent version of this article at:
doi:10.1158/1940-6207.CAPR-10-0025

Supplementary Material
Access the most recent supplemental material at:
http://cancerpreventionresearch.aacrjournals.org/content/suppl/2010/11/15/1940-6207.CAPR-10-0025.DC1

Author Manuscript
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.