



(43) International Publication Date
27 September 2012 (27.09.2012)

(10) International Publication Number
WO 2012/129192 A2

(51) International Patent Classification:
C12Q 1/37 (2006.01)

(21) International Application Number:
PCT/US2012/029723

(22) International Filing Date:
19 March 2012 (19.03.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/454,455 18 March 2011 (18.03.2011) US

(71) Applicant (for all designated States except US): **FOX CHASE CANCER CENTER** [US/US]; 333 Cottman Avenue, Philadelphia, PA (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **YEUNG, Anthony T.** [US/US]; 1445 Londonberry Lane, Huntingdon Valley, PA 19006 (US).

(74) Agents: **RIGAUT, Kathleen D.** et al.; Dann, Dorfman, Herrell & Skillman, 1601 Market Street, Suite 2400, Philadelphia, PA 19103-2307 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report (Rule 48.2(g))

(54) Title: MUCIN 5B AS A PANCREATIC CYST FLUID SPECIFIC BIOMARKER FOR ACCURATE DIAGNOSIS OF MUCINOUS CYSTS AND OTHER MARKERS USEFUL FOR DETECTION OF PANCREATIC MALIGNANCY

(57) Abstract: Compositions and methods which indicate an increased risk for pancreatic carcinoma in a test subject are disclosed.



WO 2012/129192 A2

**MUCIN 5B AS A PANCREATIC CYST FLUID SPECIFIC BIOMARKER FOR
ACCURATE DIAGNOSIS OF MUCINOUS CYSTS AND OTHER MARKERS
USEFUL FOR DETECTION OF PANCREATIC MALIGNANCY**

5

By

Anthony T. Yeung

This application claims priority to US Provisional Application 61/454,455 filed
10 March 18, 2011 the entire contents being incorporated herein by reference as though set
forth in full.

Pursuant to 35 U.S.C. §202(c) it is acknowledged that the U.S. Government has
certain rights in the invention described, which was made in part with funds from the
15 National Institutes of Health, Grant Number, CA119242.

FIELD OF THE INVENTION

This invention relates to the fields of oncology and proteomic analysis. More
20 specifically, the invention discloses biomarkers that are present in pancreatic cyst or
ductal fluid which are indicative of an increased risk for the development of pancreatic
cancer and methods of use thereof in diagnostic and prognostic assays. Also disclosed
are screening assays utilizing the biomarkers of the invention to identify agents useful for
the treatment of pancreatic cancer. The invention also relates to a biomarker and method
25 of use thereof for differentiating mucinous cysts from non-mucinous cysts.

BACKGROUND OF THE INVENTION

Several publications and patent documents are cited throughout the specification
in order to describe the state of the art to which this invention pertains. Each of these
30 citations is incorporated herein by reference as though set forth in full.

Increasing use of high resolution computerized tomography and magnetic
resonance imaging in clinical practice has resulted in detection of a growing number of
pancreatic cysts (1). As a result, clinicians are frequently asked to determine the
biological nature of these cystic lesions, and to make treatment recommendations
35 accordingly. However, there are currently no diagnostic indicators that are
consistently reliable, obtainable, and conclusive for diagnosing and risk-stratifying
pancreatic cysts. The sensitivity of pancreatic cyst fluid cytology has been reported as

only 27-64% in most series. Several studies have suggested that a variety of tumor markers (e.g., CEA(2), CA 19-9, CA 15-3) may distinguish mucinous from non-mucinous cystic lesions, and also may predict whether a cyst harbors areas of malignant transformation (3-5).

The biologic nature and histopathologic features of pancreatic cysts are varied (3, 6). Ten to twenty percent of pancreatic cysts are neoplastic, including neoplasms which grow as cystic structures (i.e., primary cystic neoplasms of the pancreas), and solid neoplasms that have undergone cystic degeneration. Serous cystadenomas (microcystic adenomas) account for approximately 32-39% of the primary cystic neoplasms and have very low malignant potential. Mucinous cystic neoplasms, which include mucinous cystadenomas and intraductal papillary mucinous neoplasms, are a subgroup of primary cystic neoplasms that have malignant potential. Nomenclature describing the evolution of these lesions, from benign to malignant, is provided elsewhere(7, 8). Mucinous cystic neoplasms and intraductal papillary mucinous neoplasms (IPMNs) account for approximately 10-45% and 21-33% of primary cystic neoplasms, respectively (6, 9-11). Two subtypes of IPMN have been described (1, 12), a main duct variant and a branch duct variant; the latter may have a more indolent course. There are other less common forms of primary cystic neoplasms of the pancreas, such as solid pseudopapillary tumors.

In the absence of reliable methods of quantifying the malignant potential of a suspected pre-malignant cystic neoplasm of the pancreas, if existing clinical parameters suggest the presence of one such lesion in a person that is otherwise an acceptable surgical risk, partial or total pancreatectomy may be recommended but can result in significant morbidity and mortality (13). Alternatively, a conservative "watch-and-wait" approach (i.e., serial imaging over time) is advocated for some patients, but this strategy may be suboptimal due to incremental costs accrued during surveillance, and the possibility that malignant transformation may occur between surveillance time points.

In light of all the foregoing, it is clear that a more reliable method for identifying those patients at increased risk for developing pancreatic cancer is urgently needed.

30 SUMMARY OF THE INVENTION

In accordance with the present invention, a method of diagnosing an increased risk for the development of pancreatic cancer in a human test subject is provided. An exemplary method entails isolating a pancreatic fluid specimen (e.g., cyst or ductal fluid) from the subject; analyzing the fluid specimen for the presence of 1, 2, 3, 4, 5, 6, 7, 8, 9

or 10 or more or all of the biomarkers associated with increased risk of pancreatic carcinoma, wherein the presence of said biomarkers is indicative of an increased risk for pancreatic cancer. In one embodiment, said biomarkers are selected from the group consisting of mucin 1, mucin 2, mucin 5AC, mucin 5B, mucin 6, CEA CAM 1,
 5 CEACAM 6, CEACAM 7, CEACAM 8, S100-A6, S100-A8, S100 A9 and S100 A-11.

The risk of pancreatic cancer increases when the proteomic analysis shows an increase in several combinations of biomarkers. These are as follows:

- a) the presence of several isoforms of mucins e.g., mucin 1, 2, 5AC, 5B, and 6;
- b) the presence of both mucins and certain isoforms of CEA, including CEACAM 1, 6,
 10 7, and 8; and c) mucins and CEA are variable but CEACAM8 is present and at least two of S100-A6, A8, A9, or A11 are present.

Also provided is a solid support comprising antibodies which are immunospecific for the biomarkers described above. Such supports can include, without limitation, filters, biacore chips, ELISA plates and the like.

15 In yet another embodiment, a method for differentiating a pancreatic mucinous cyst from a non-mucinous cyst in a human test subject is provided. An exemplary method entails providing a pancreatic cyst fluid specimen from the subject and analyzing the fluid specimen for the presence or absence of mucin 5B, presence of this mucin being indicative of a mucinous cyst. If a mucinous cyst is present the method
 20 further comprises assessing the patient over time for development of pancreatic cancer. The cyst fluid can be analyzed by a variety methods, however, Thermo Quantum Access TSQ triple-quadrupole mass spectrometry and qTOF mass spectrometry are preferred.

Finally, methods for identifying pancreatic biomarker proteins from ductal fluids are also disclosed. The presence or absence of such biomarker proteins can also be used
 25 to advantage to diagnose an increased or decreased risk for the development of pancreatic cancer. The markers of the invention are also useful for stratifying malignant disease of the pancreas.

BRIEF DESCRIPTION OF THE DRAWINGS

30 Figure 1 is a table providing clinical information for the pancreatic cyst fluid samples tested herein. CEA (ng/mL) and Amylase (units/mL) are results from clinical lab immuno assays. MCA=mucinous cystadenoma. Cytology categories: A- Benign: No evidence of benign mucinous epithelium, atypical cells or carcinoma. C- Atypical/suspicious cytology.

Figures 2A – 2O provide tables showing representatives of 137 plasma proteins distributed among the pancreatic cyst samples. The numbers presented in Figure 2A are emPAI scores which are roughly proportional to protein abundance. Dark (>1), medium (0.1 to 1), and light (<0.1) shading denotes relative protein abundance. No proteins were detected for the empty boxes. CEA = ng/mL.

Figure 3 is a table listing the pancreatic enzyme proteins in the pancreatic cyst samples. Legends are the same as for Figure 2.

10

Figure 4 is a table listing proteomic biomarkers for pancreatic cancer e.g., mucins, CEACAMs, and S100s found in the pancreatic cyst samples. Legends are the same as for Figure 2.

15 Figure 5 is a micrograph of an EUS image of a pancreatic cyst that may be mucinous. Note needle in cyst.

Figure 6 provides an overview of cystic lesion of the pancreas. Other cyst types include: congenital cysts, cystic degeneration of primary solid pancreatic tumors, solid pseudopapillary neoplasms, retention cysts and gastrointestinal duplication cysts.

20

Figure 7 is a schematic diagram of the assay method of the invention.

Figure 8 is a table (Table 6) showing clinical information for the pancreatic duct fluid samples.

25

Figure 9 is a table (Table 7) showing blood proteins observed in the pancreatic duct fluid samples.

30 Figures 10A and 10B are a pair of tables (Tables 8A and 8B) showing relevant proteomic biomarkers for neoplasms in pancreatic ductal fluids.

Figures 11A and 11B (Tables 9A and 9B) are a pair of tables showing relevant proteomic biomarkers found in the pancreatic cyst samples in the study presented in Example 4.

5

DETAILED DESCRIPTION OF THE INVENTION

There are currently no diagnostic indicators that are consistently reliable, obtainable, and conclusive for diagnosing and risk-stratifying pancreatic cysts. To establish more effective diagnostic biomarkers and to provide deeper understanding about the molecular profile within these cysts, we identified and quantified about 500 cyst fluid proteins and correlated the findings to clinical parameters, when available. Pancreatic cyst fluids were collected by endoscopic ultrasound-guided fine needle aspiration (EUS-FNA) from 20 patients. The proteins in the cyst fluids were ascertained by LC/MS/MS analysis of every gel slice from 1 D gel fractionation of each sample, using partial peptide sequencing on a highly accurate and stable mass spectrometer. Measurements of traditional markers of mucins, amylase, and CEA were obtained simultaneously from 15 micrograms of protein from less than 40 microliters of cyst fluids per sample. The proteomic techniques utilized provided comprehensive information about pancreatic enzymes, plasma infiltrate proteins, and proteins that may have been produced by the pancreas epithelium. Our data suggest that diagnosis based upon the proteome of pancreatic cyst fluid may include the expression of two homologs of amylase, five mucins, five CEA-related cell adhesion molecules (CEACAMs), and four S100 homologs. Furthermore, our study indicates that proteomic profiling using small amounts of cyst fluids can be a valuable tool for diagnosing and risk-stratifying cystic lesions of the pancreas.

Bodily fluids aspirated from pancreatic cysts contain hundreds of different proteins. Some of the proteins are natural pancreatic enzyme secretions; others are plasma derived; yet others may be released by the cyst epithelium either normally or as a result of cellular transformation. Proteomics by mass spectrometry provide a means to quickly quantify hundreds of these proteins simultaneously from a small volume of fluids. The identification of proteins that change their levels upon cellular transformation provides biomarkers for pancreas malignancy.

We have determined that the risk of pancreatic cancer increases when the proteomic analysis shows an increase in several combinations of biomarkers: (1) When

several isoforms of mucins 1, 2, 5AC, 5B, and 6 are present; (2) when mucins are present, the risk further increases when isoforms of CEA, including CEACAM 1, 6, 7, and 8 are present; (3) when mucins and CEA are either present or absent, the risk increases when CEACAM8 is present and when certain of the biomarkers S100-A6, A8, A9, or A11 are present.

The invention also provides a biomarker, Mucin-5B and a method of use thereof for differentiating mucinous cysts from non-mucinous cysts. The difficulty of diagnosing a cyst as mucinous arises because of frequent mucin contamination which occurs when the needle used in the Fine Needle Aspiration of a EUS procedure passes through the wall of the stomach to the pancreas. We discovered, using as small a sample as one microliter fluid in mass spectrometry proteomics (500 microliters samples are conventionally used in the clinic, requiring large cysts), that mucin 5B is a significantly better pancreas-specific marker of mucin than the abundant mucin 5AC that arises from both pancreas and stomach. Previous data suggests that mucin 5B is not expressed in stomach. Fluid samples will be tested for mucin 5B and the findings correlated with surgical pathology diagnosis. Mucins 1, 5AC, 6, and 13 will also be correlated. Accurate identification of mucins of pancreas origin will help eliminate the frequent false positives which arise due to sample contamination.

The biomarkers of the invention include genes and proteins, and variants and fragments thereof. Such biomarkers include DNA comprising the entire or partial sequence of the nucleic acid sequence encoding the biomarker, or the complement of such a sequence. The biomarker nucleic acids also include RNA comprising the entire or partial sequence of any of the nucleic acid sequences of interest. A biomarker protein is a protein encoded by or corresponding to a DNA biomarker of the invention. A biomarker protein comprises the entire or partial amino acid sequence of any of the biomarker proteins or polypeptides.

A "biomarker" is any gene or protein whose level of expression in a tissue or cell is altered compared to that of a normal or healthy cell or tissue. Biomarkers of the invention are selective for underlying risk of progression to pancreatic cancer. By "selectively overexpressed in pancreatic cyst fluid" is intended that the biomarker of interest is overexpressed in neoplastic cysts relative to benign or non-malignant cysts. Thus, detection of the biomarkers of the invention permits the differentiation of samples indicative of increased risk of developing neoplasms of the pancreas from samples that

are indicative of benign proliferation. Representative biomarkers for pancreatic cell transformation include one or more or a plurality of the following proteins:

- Mucin-5B precursor - Homo sapiens (Human)
- Mucin-5AC - Homo sapiens (Human)
- 5 Mucin-1 precursor - Homo sapiens (Human)
- Gelsolin precursor - Homo sapiens (Human)
- Carcinoembryonic antigen-related cell adhesion molecule 5 precursor - Homo sapiens (Human)
- Ezrin - Homo sapiens (Human)
- 10 Galectin-3-binding protein precursor - Homo sapiens (Human)
- Mucin-13 precursor - Homo sapiens (Human)
- Leukocyte elastase inhibitor - Homo sapiens (Human)
- Annexin A1 - Homo sapiens (Human)
- Annexin A2 - Homo sapiens (Human)
- 15 Carcinoembryonic antigen-related cell adhesion molecule 6 precursor - Homo sapiens (Human)
- Annexin A3 - Homo sapiens (Human)
- Annexin A4 - Homo sapiens (Human)
- Galectin-4 - Homo sapiens (Human)
- 20 Annexin A5 - Homo sapiens (Human)
- Phosducin - Homo sapiens (Human)
- Tetraspanin-8 - Homo sapiens (Human)
- Galectin-3 - Homo sapiens (Human)
- Neutrophil gelatinase-associated lipocalin precursor - Homo sapiens (Human)
- 25 Anterior gradient protein 2 homolog precursor - Homo sapiens (Human)
- Protein S100-A11 - Homo sapiens (Human)
- Protein S100-A6 - Homo sapiens (Human)
- Protein S100-A8 - Homo sapiens (Human)
- Protein S100-A9 - Homo sapiens (Human)

- 30 Expression of the biomarkers described herein is indicative of cyst fluid protein profiles that are associated with benign pancreatic disease, pre-malignancy, and neoplastic lesions of the pancreas.

The phrase “genetic signature” refers to a plurality of nucleic acid molecules whose expression levels are indicative of a given metabolic or pathological state. The

genetic signatures described herein can be employed to characterize at the molecular level the condition of the pancreatic cyst that is associated with an increased risk of pancreatic cancer, thus providing a useful molecular tool for predicting outcomes, for identifying patients at risk, and for use in biomarker in assays for evaluating cancer preventive agents.

For purposes of the present invention, “a” or “an” entity refers to one or more of that entity; for example, “a cDNA” refers to one or more cDNA or at least one cDNA. The terms “a” or “an,” “one or more” and “at least one” can be used interchangeably herein. It is also noted that the terms “comprising,” “including,” and “having” can be used interchangeably. Furthermore, a compound “selected from the group consisting of” refers to one or more of the compounds in the list that follows, including mixtures (i.e. combinations) of two or more of the compounds. According to the present invention, an isolated, or biologically pure molecule is a compound that has been removed from its natural milieu. As such, “isolated” and “biologically pure” do not necessarily reflect the extent to which the compound has been purified. An isolated compound of the present invention can be obtained from its natural source, can be produced using laboratory synthetic techniques or can be produced by any such chemical synthetic route.

The term “genetic alteration” as used herein refers to a change from the wild-type or reference sequence of one or more nucleic acid molecules. Genetic alterations include without limitation, base pair substitutions, additions and deletions of at least one nucleotide from a nucleic acid molecule of known sequence.

The term “solid matrix” as used herein refers to any format, such as beads, microparticles, a microarray, the surface of a microtitration well or a test tube, a dipstick or a filter. The material of the matrix may be polystyrene, cellulose, latex, nitrocellulose, nylon, polyacrylamide, dextran or agarose.

“Sample” or “patient sample” or “biological sample” generally refers to a sample which may be tested for a particular molecule, preferably a genetic signature specific marker molecule, such as a marker shown in the tables provided below. Samples may include but are not limited to cells, cyst fluids, body fluids, including blood, serum, plasma, urine, saliva, tears, pleural fluid and the like.

The phrase “consisting essentially of” when referring to a particular nucleotide or amino acid means a sequence having the properties of a given SEQ ID NO. For example, when used in reference to an amino acid sequence, the phrase includes the

sequence per se and molecular modifications that would not affect the functional and novel characteristics of the sequence.

With regard to nucleic acids used in the invention, the term "isolated nucleic acid" is sometimes employed. This term, when applied to DNA, refers to a DNA molecule that is separated from sequences with which it is immediately contiguous (in the 5' and 3' directions) in the naturally occurring genome of the organism from which it was derived. For example, the "isolated nucleic acid" may comprise a DNA molecule inserted into a vector, such as a plasmid or virus vector, or integrated into the genomic DNA of a prokaryote or eukaryote. An "isolated nucleic acid molecule" may also comprise a cDNA molecule. An isolated nucleic acid molecule inserted into a vector is also sometimes referred to herein as a recombinant nucleic acid molecule.

With respect to RNA molecules, the term "isolated nucleic acid" primarily refers to an RNA molecule encoded by an isolated DNA molecule as defined above. Alternatively, the term may refer to an RNA molecule that has been sufficiently separated from RNA molecules with which it would be associated in its natural state (i.e., in cells or tissues), such that it exists in a "substantially pure" form. By the use of the term "enriched" in reference to nucleic acid it is meant that the specific DNA or RNA sequence constitutes a significantly higher fraction (2-5 fold) of the total DNA or RNA present in the cells or solution of interest than in normal cells or in the cells from which the sequence was taken. This could be caused by a person by preferential reduction in the amount of other DNA or RNA present, or by a preferential increase in the amount of the specific DNA or RNA sequence, or by a combination of the two. However, it should be noted that "enriched" does not imply that there are no other DNA or RNA sequences present, just that the relative amount of the sequence of interest has been significantly increased.

It is also advantageous for some purposes that a nucleotide sequence be in purified form. The term "purified" in reference to nucleic acid does not require absolute purity (such as a homogeneous preparation); instead, it represents an indication that the sequence is relatively purer than in the natural environment (compared to the natural level, this level should be at least 2-5 fold greater, e.g., in terms of mg/ml). Individual clones isolated from a cDNA library may be purified to electrophoretic homogeneity. The claimed DNA molecules obtained from these clones can be obtained directly from total DNA or from total RNA. The cDNA clones are not naturally occurring, but rather are preferably obtained via manipulation of a partially purified naturally occurring

substance (messenger RNA). The construction of a cDNA library from mRNA involves the creation of a synthetic substance (cDNA) and pure individual cDNA clones can be isolated from the synthetic library by clonal selection of the cells carrying the cDNA library. Thus, the process which includes the construction of a cDNA library from
5 mRNA and isolation of distinct cDNA clones yields an approximately 10^{-6} -fold purification of the native message. Thus, purification of at least one order of magnitude, preferably two or three orders, and more preferably four or five orders of magnitude is expressly contemplated. Thus, the term "substantially pure" refers to a preparation comprising at least 50-60% by weight the compound of interest (e.g., nucleic acid,
10 oligonucleotide, etc.). More preferably, the preparation comprises at least 75% by weight, and most preferably 90-99% by weight, the compound of interest. Purity is measured by methods appropriate for the compound of interest.

The term "complementary" describes two nucleotides that can form multiple favorable interactions with one another. For example, adenine is complementary to
15 thymine as they can form two hydrogen bonds. Similarly, guanine and cytosine are complementary since they can form three hydrogen bonds. Thus if a nucleic acid sequence contains the following sequence of bases, thymine, adenine, guanine and cytosine, a "complement" of this nucleic acid molecule would be a molecule containing adenine in the place of thymine, thymine in the place of adenine, cytosine in the place of
20 guanine, and guanine in the place of cytosine. Because the complement can contain a nucleic acid sequence that forms optimal interactions with the parent nucleic acid molecule, such a complement can bind with high affinity to its parent molecule.

With respect to single stranded nucleic acids, particularly oligonucleotides, the term "specifically hybridizing" refers to the association between two single-stranded
25 nucleotide molecules of sufficiently complementary sequence to permit such hybridization under pre-determined conditions generally used in the art (sometimes termed "substantially complementary"). In particular, the term refers to hybridization of an oligonucleotide with a substantially complementary sequence contained within a single-stranded DNA or RNA molecule of the invention, to the substantial exclusion of
30 hybridization of the oligonucleotide with single-stranded nucleic acids of non-complementary sequence. For example, specific hybridization can refer to a sequence which hybridizes to any specific marker gene or nucleic acid, but does not hybridize to other human nucleotides. Also polynucleotide which "specifically hybridizes" may hybridize only to a specific marker, such a genetic signature-specific marker shown in

the Tables below. Appropriate conditions enabling specific hybridization of single stranded nucleic acid molecules of varying complementarity are well known in the art.

For instance, one common formula for calculating the stringency conditions required to achieve hybridization between nucleic acid molecules of a specified sequence
5 homology is set forth below (Sambrook et al., Molecular Cloning, Cold Spring Harbor Laboratory (1989):

$$T_m = 81.5^{\circ}\text{C} + 16.6\text{Log} [\text{Na}^+] + 0.41(\% \text{ G+C}) - 0.63 (\% \text{ formamide}) - 600/\text{\#bp in duplex}$$

10 As an illustration of the above formula, using $[\text{Na}^+] = [0.368]$ and 50% formamide, with GC content of 42% and an average probe size of 200 bases, the T_m is 57°C . The T_m of a DNA duplex decreases by 1 - 1.5°C with every 1% decrease in homology. Thus, targets with greater than about 75% sequence identity would be observed using a hybridization temperature of 42°C .

15 The stringency of the hybridization and wash depend primarily on the salt concentration and temperature of the solutions. In general, to maximize the rate of annealing of the probe with its target, the hybridization is usually carried out at salt and temperature conditions that are $20\text{-}25^{\circ}\text{C}$ below the calculated T_m of the hybrid. Wash conditions should be as stringent as possible for the degree of identity of the probe for
20 the target. In general, wash conditions are selected to be approximately $12\text{-}20^{\circ}\text{C}$ below the T_m of the hybrid. In regards to the nucleic acids of the current invention, a moderate stringency hybridization is defined as hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 $\mu\text{g/ml}$ denatured salmon sperm DNA at 42°C , and washed in 2X SSC and 0.5% SDS at 55°C for 15 minutes. A high stringency hybridization is defined as
25 hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 $\mu\text{g/ml}$ denatured salmon sperm DNA at 42°C , and washed in 1X SSC and 0.5% SDS at 65°C for 15 minutes. A very high stringency hybridization is defined as hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 $\mu\text{g/ml}$ denatured salmon sperm DNA at 42°C , and washed in 0.1X SSC and 0.5% SDS at 65°C for 15 minutes.

30 The term "oligonucleotide" or "oligo" as used herein means a short sequence of DNA or DNA derivatives typically 8 to 35 nucleotides in length, primers, or probes. An oligonucleotide can be derived synthetically, by cloning or by amplification. An oligo is defined as a nucleic acid molecule comprised of two or more ribo- or

deoxyribonucleotides, preferably more than three. The exact size of the oligonucleotide will depend on various factors and on the particular application and use of the oligonucleotide. The term "derivative" is intended to include any of the above described variants when comprising an additional chemical moiety not normally a part of these
5 molecules. These chemical moieties can have varying purposes including, improving solubility, absorption, biological half life, decreasing toxicity and eliminating or decreasing undesirable side effects.

The term "probe" as used herein refers to an oligonucleotide, polynucleotide or nucleic acid, either RNA or DNA, whether occurring naturally as in a purified restriction
10 enzyme digest or produced synthetically, which is capable of annealing with or specifically hybridizing to a nucleic acid with sequences complementary to the probe. A probe may be either single-stranded or double-stranded. The exact length of the probe will depend upon many factors, including temperature, source of probe and use of the method. For example, for diagnostic applications, depending on the complexity of the
15 target sequence, the oligonucleotide probe typically contains 15-25 or more nucleotides, although it may contain fewer nucleotides. The probes herein are selected to be complementary to different strands of a particular target nucleic acid sequence. This means that the probes must be sufficiently complementary so as to be able to "specifically hybridize" or anneal with their respective target strands under a set of pre-
20 determined conditions. Therefore, the probe sequence need not reflect the exact complementary sequence of the target. For example, a non-complementary nucleotide fragment may be attached to the 5' or 3' end of the probe, with the remainder of the probe sequence being complementary to the target strand. Alternatively, non-complementary bases or longer sequences can be interspersed into the probe, provided that the probe
25 sequence has sufficient complementarity with the sequence of the target nucleic acid to anneal therewith specifically.

The term "primer" as used herein refers to an oligonucleotide, either RNA or DNA, either single-stranded or double-stranded, either derived from a biological system, generated by restriction enzyme digestion, or produced synthetically which, when placed
30 in the proper environment, is able to functionally act as an initiator of template-dependent nucleic acid synthesis. When presented with an appropriate nucleic acid template, suitable nucleoside triphosphate precursors of nucleic acids, a polymerase enzyme, suitable cofactors and conditions such as a suitable temperature and pH, the

primer may be extended at its 3' terminus by the addition of nucleotides by the action of a polymerase or similar activity to yield a primer extension product. The primer may vary in length depending on the particular conditions and requirement of the application. For example, in diagnostic applications, the oligonucleotide primer is typically 15-25 or
5 more nucleotides in length. The primer must be of sufficient complementarity to the desired template to prime the synthesis of the desired extension product, that is, to be able anneal with the desired template strand in a manner sufficient to provide the 3' hydroxyl moiety of the primer in appropriate juxtaposition for use in the initiation of synthesis by a polymerase or similar enzyme. It is not required that the primer sequence
10 represent an exact complement of the desired template. For example, a non-complementary nucleotide sequence may be attached to the 5' end of an otherwise complementary primer. Alternatively, non-complementary bases may be interspersed within the oligonucleotide primer sequence, provided that the primer sequence has sufficient complementarity with the sequence of the desired template strand to
15 functionally provide a template-primer complex for the synthesis of the extension product.

Polymerase chain reaction (PCR) has been described in US Patents 4,683,195, 4,800,195, and 4,965,188, the entire disclosures of which are incorporated by reference herein.

20 An "siRNA" refers to a molecule involved in the RNA interference process for a sequence-specific post-transcriptional gene silencing or gene knockdown by providing small interfering RNAs (siRNAs) that has homology with the sequence of the targeted gene. Small interfering RNAs (siRNAs) can be synthesized in vitro or generated by ribonuclease III cleavage from longer dsRNA and are the mediators of sequence-specific
25 mRNA degradation. Preferably, the siRNA of the invention are chemically synthesized using appropriately protected ribonucleoside phosphoramidites and a conventional DNA/RNA synthesizer. The siRNA can be synthesized as two separate, complementary RNA molecules, or as a single RNA molecule with two complementary regions. Commercial suppliers of synthetic RNA molecules or synthesis reagents include Applied
30 Biosystems (Foster City, CA, USA), Proligo (Hamburg, Germany), Dharmacon Research (Lafayette, Colo., USA), Pierce Chemical (part of Perbio Science, Rockford, Ill., USA), Glen Research (Sterling, Va., USA), ChemGenes (Ashland, Mass., USA) and Cruachem (Glasgow, UK). Specific siRNA constructs for inhibiting elevated mRNA

levels associated with pancreatic cancer may be between 15-35 nucleotides in length, and more typically about 21 nucleotides in length.

The term "vector" relates to a single or double stranded circular nucleic acid molecule that can be infected, transfected or transformed into cells and replicate
5 independently or within the host cell genome. A circular double stranded nucleic acid molecule can be cut and thereby linearized upon treatment with restriction enzymes. An assortment of vectors, restriction enzymes, and the knowledge of the nucleotide sequences that are targeted by restriction enzymes are readily available to those skilled in the art, and include any replicon, such as a plasmid, cosmid, bacmid, phage or virus, to
10 which another genetic sequence or element (either DNA or RNA) may be attached so as to bring about the replication of the attached sequence or element. A nucleic acid molecule of the invention can be inserted into a vector by cutting the vector with restriction enzymes and ligating the two pieces together.

Many techniques are available to those skilled in the art to facilitate
15 transformation, transfection, or transduction of the expression construct into a prokaryotic or eukaryotic organism. The terms "transformation", "transfection", and "transduction" refer to methods of inserting a nucleic acid and/or expression construct into a cell or host organism. These methods involve a variety of techniques, such as treating the cells with high concentrations of salt, an electric field, or detergent, to render
20 the host cell outer membrane or wall permeable to nucleic acid molecules of interest, microinjection, peptide-tethering, PEG-fusion, and the like.

The term "promoter element" describes a nucleotide sequence that is incorporated into a vector that, once inside an appropriate cell, can facilitate transcription factor and/or polymerase binding and subsequent transcription of portions of the vector DNA into
25 mRNA. In one embodiment, the promoter element of the present invention precedes the 5' end of the pancreatic cancer specific marker nucleic acid molecule(s) such that the latter is transcribed into mRNA. Host cell machinery then translates mRNA into a polypeptide.

Those skilled in the art will recognize that a nucleic acid vector can contain
30 nucleic acid elements other than the promoter element and the pancreatic cancer specific marker gene nucleic acid molecule(s). These other nucleic acid elements include, but are not limited to, origins of replication, ribosomal binding sites, nucleic acid sequences encoding drug resistance enzymes or amino acid metabolic enzymes, and nucleic acid

sequences encoding secretion signals, localization signals, or signals useful for polypeptide purification.

A "replicon" is any genetic element, for example, a plasmid, cosmid, bacmid, plastid, phage or virus that is capable of replication largely under its own control. A
5 replicon may be either RNA or DNA and may be single or double stranded.

An "expression operon" refers to a nucleic acid segment that may possess transcriptional and translational control sequences, such as promoters, enhancers, translational start signals (e.g., ATG or AUG codons), polyadenylation signals, terminators, and the like, and which facilitate the expression of a polypeptide coding
10 sequence in a host cell or organism.

As used herein, the terms "reporter," "reporter system", "reporter gene," or "reporter gene product" shall mean an operative genetic system in which a nucleic acid comprises a gene that encodes a product that when expressed produces a reporter signal that is a readily measurable, e.g., by biological assay, immunoassay, radio
15 immunoassay, or by colorimetric, fluorogenic, chemiluminescent or other methods. The nucleic acid may be either RNA or DNA, linear or circular, single or double stranded, antisense or sense polarity, and is operatively linked to the necessary control elements for the expression of the reporter gene product. The required control elements will vary according to the nature of the reporter system and whether the reporter gene is in the
20 form of DNA or RNA, but may include, but not be limited to, such elements as promoters, enhancers, translational control sequences, poly A addition signals, transcriptional termination signals and the like.

The introduced nucleic acid may or may not be integrated (covalently linked) into nucleic acid of the recipient cell or organism. In bacterial, yeast, plant and mammalian
25 cells, for example, the introduced nucleic acid may be maintained as an episomal element or independent replicon such as a plasmid. Alternatively, the introduced nucleic acid may become integrated into the nucleic acid of the recipient cell or organism and be stably maintained in that cell or organism and further passed on or inherited to progeny cells or organisms of the recipient cell or organism. Finally, the introduced nucleic acid
30 may exist in the recipient cell or host organism only transiently.

The term "selectable marker gene" refers to a gene that when expressed confers a selectable phenotype, such as antibiotic resistance, on a transformed cell.

The term "operably linked" means that the regulatory sequences necessary for expression of the coding sequence are placed in the DNA molecule in the appropriate

positions relative to the coding sequence so as to effect expression of the coding sequence. This same definition is sometimes applied to the arrangement of transcription units and other transcription control elements (e.g. enhancers) in an expression vector.

The terms "recombinant organism," or "transgenic organism" refer to organisms
5 which have a new combination of genes or nucleic acid molecules. A new combination of genes or nucleic acid molecules can be introduced into an organism using a wide array of nucleic acid manipulation techniques available to those skilled in the art. The term "organism" relates to any living being comprised of a least one cell. An organism can be as simple as one eukaryotic cell or as complex as a mammal. Therefore, the phrase "a
10 recombinant organism" encompasses a recombinant cell, as well as eukaryotic and prokaryotic organism.

The term "isolated protein" or "isolated and purified protein" is sometimes used herein. This term refers primarily to a protein produced by expression of an isolated genetic signature nucleic acid or biomarker molecule of the invention. Alternatively, this
15 term may refer to a protein that has been sufficiently separated from other proteins with which it would naturally be associated, so as to exist in "substantially pure" form. "Isolated" is not meant to exclude artificial or synthetic mixtures with other compounds or materials, or the presence of impurities that do not interfere with the fundamental activity and that may be present, for example, due to incomplete purification, addition of
20 stabilizers, or compounding into, for example, immunogenic preparations or pharmaceutically acceptable preparations.

A "specific binding pair" comprises a specific binding member (sbm) and a binding partner (bp) which have a particular specificity for each other and which in normal conditions bind to each other in preference to other molecules. Examples of
25 specific binding pairs are antigens and antibodies, ligands and receptors and complementary nucleotide sequences. The skilled person is aware of many other examples. Further, the term "specific binding pair" is also applicable where either or both of the specific binding member and the binding partner comprise a part of a large molecule. In embodiments in which the specific binding pair comprises nucleic acid
30 sequences, they will be of a length to hybridize to each other under conditions of the assay, preferably greater than 10 nucleotides long, more preferably greater than 15 or 20 nucleotides long.

"Sample" or "patient sample" or "biological sample" generally refers to a sample which may be tested for a particular molecule or combination of molecules, preferably a

5 The terms "agent" and "test compound" are used interchangeably herein and denote a chemical compound, a mixture of chemical compounds, a biological macromolecule, or an extract made from biological materials such as bacteria, plants, fungi, or animal (particularly mammalian) cells or tissues. Biological macromolecules include siRNA, shRNA, antisense oligonucleotides, small molecules, antibodies, 10 peptides, peptide/DNA complexes, and any nucleic acid based molecule, for example an oligo, which exhibits the capacity to modulate the activity of the genetic signature nucleic acids described herein or their encoded proteins. Agents are evaluated for potential biological activity by inclusion in screening assays described herein below.

20 **METHODS OF USING THE BIOMARKERS AND GENETIC SIGNATURES OF**
THE INVENTION

Further, assays for detecting the genetic signature may be conducted on any type of biological sample, but is most preferably performed on cyst fluid. From the foregoing

discussion, it can be seen that genetic signature containing nucleic acids, vectors expressing the same, genetic signature encoded proteins and anti-genetic signature encoded protein specific antibodies of the invention can be used to detect the signature in body tissue, cells, or fluid, and alter genetic signature containing marker protein expression for purposes of assessing the genetic and protein interactions involved in pancreatic cancer.

In certain embodiments for screening for genetic signature containing nucleic acid(s), the sample will initially be amplified, e.g. using PCR, to increase the amount of the template as compared to other sequences present in the sample. This allows the target sequences to be detected with a high degree of sensitivity if they are present in the sample. This initial step may be avoided by using highly sensitive array techniques that are becoming increasingly important in the art.

Alternatively, alternative detection technologies will be employed which detect the pancreatic cancer biomarker proteins directly. Such methods include geLC/MS/MS proteomics analysis. This approach provides a full panel of the protein biomarkers present in cyst fluid and allows the clinician to predict outcomes based on the panel of biomarkers present in a sample.

Thus, any of the aforementioned techniques may be used to detect or quantify genetic signature expression and or protein expression levels and accordingly, diagnose patient susceptibility for developing pancreatic cancer.

KITS AND ARTICLES OF MANUFACTURE

Any of the aforementioned products can be incorporated into a kit which may contain genetic signature polynucleotides or one or more such markers immobilized on a Gene Chip, an oligonucleotide, a polypeptide, a peptide, an antibody, a label, marker, or reporter, a pharmaceutically acceptable carrier, a physiologically acceptable carrier, instructions for use, a container, a vessel for administration, an assay substrate, or any combination thereof.

METHODS OF USING THE GENETIC SIGNATURE OR BIOMARKER PROTEINS FOR DEVELOPMENT OF THERAPEUTIC AGENTS

Since the genetic signature identified herein and the proteins encoded thereby has been associated with the etiology of pancreatic cancer, methods for identifying agents that modulate the activity of the genes and their encoded products should result in the

generation of efficacious therapeutic agents for the treatment of a cancer, particularly pancreatic cancer.

The nucleic acids comprising the signature contain regions which provide suitable targets for the rational design of therapeutic agents which modulate their activity. Small peptide molecules corresponding to these regions may be used to advantage in the design of therapeutic agents which effectively modulate the activity of the encoded proteins. Molecular modeling should facilitate the identification of specific organic molecules with capacity to bind to the active site of the proteins encoded by the genetic signature nucleic acids based on conformation or key amino acid residues required for function. A combinatorial chemistry approach will be used to identify molecules with greatest activity and then iterations of these molecules will be developed for further cycles of screening. In certain embodiments, candidate agents can be screening from large libraries of synthetic or natural compounds. Such compound libraries are commercially available from a number of companies including but not limited to Maybridge Chemical Co., (Trevillet, Cornwall, UK), Comgenex (Princeton, NJ), Microsour (New Milford, CT) Aldrich (Milwaukee, WI) Akos Consulting and Solutions GmbH (Basel, Switzerland), Ambinter (Paris, France), Asinex (Moscow, Russia) Aurora (Graz, Austria), BioFocus DPI (Switzerland), Bionet (Camelford, UK), Chembridge (San Diego, CA), Chem Div (San Diego, CA). The skilled person is aware of other sources and can readily purchase the same. Once therapeutically efficacious compounds are identified in the screening assays described herein, they can be formulated in to pharmaceutical compositions and utilized for the treatment of pancreatic cancer.

The polypeptides or fragments employed in drug screening assays may either be free in solution, affixed to a solid support or within a cell. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant polynucleotides expressing the biomarker polypeptide or fragment, preferably in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may determine, for example, formation of complexes between the polypeptide or fragment and the agent being tested, or examine the degree to which the formation of a complex between the polypeptide or fragment and a known substrate is interfered with by the agent being tested.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity for the encoded polypeptides and is

described in detail in Geysen, PCT published application WO 84/03564, published on September 13, 1984. Briefly stated, large numbers of different, small peptide test compounds, such as those described above, are synthesized on a solid substrate, such as plastic pins or some other surface. The peptide test compounds are reacted with the
5 target polypeptide and washed. Bound polypeptide is then detected by methods well known in the art.

A further technique for drug screening involves the use of host eukaryotic cell lines or cells (such as described above) which have a nonfunctional or altered pancreatic cancer associated gene. These host cell lines or cells are defective at the polypeptide
10 level. The host cell lines or cells are grown in the presence of drug compound. The effect on cellular morphology and/or proliferation of the host cells is measured to determine if the compound is capable of regulating the same in the defective cells. Host cells contemplated for use in the present invention include but are not limited to bacterial cells, fungal cells, insect cells, mammalian cells, particularly pancreatic cells. The
15 genetic signature encoding DNA molecules may be introduced singly into such host cells or in combination to assess the phenotype of cells conferred by such expression. Methods for introducing DNA molecules are also well known to those of ordinary skill in the art. Such methods are set forth in Ausubel et al. eds., Current Protocols in Molecular Biology, John Wiley & Sons, NY, N.Y. 1995, the disclosure of which is
20 incorporated by reference herein.

Pancreatic cells and pancreatic cell lines suitable for studying the effects of genetic signature expression on cellular morphology and signaling methods of use thereof for drug discovery are provided. Such cells and cell lines will be transfected with genetic signature encoding nucleic acids described herein and the effects on pancreatic
25 cell functions and/or cyst formation can be determined. Such cells and cell lines can also be contacted with the siRNA molecules provided herein to assess the effects thereof on malignant transformation. The siRNA molecules will be tested alone and in combination of 2, 3, 4, and 5 siRNAs to identify the most efficacious combination for down regulating target nucleic acids.

30 A wide variety of expression vectors are available that can be modified to express the novel DNA or RNA sequences of this invention. The specific vectors exemplified herein are merely illustrative, and are not intended to limit the scope of the invention. Expression methods are described by Sambrook et al. Molecular Cloning: A Laboratory Manual or Current Protocols in Molecular Biology 16.3-17.44 (1989). Expression

methods in *Saccharomyces* are also described in *Current Protocols in Molecular Biology* (1989).

Suitable vectors for use in practicing the invention include prokaryotic vectors such as the pNH vectors (Stratagene Inc., 11099 N. Torrey Pines Rd., La Jolla, Calif. 92037), pET vectors (Novogen Inc., 565 Science Dr., Madison, Wis. 53711) and the pGEX vectors (Pharmacia LKB Biotechnology Inc., Piscataway, N.J. 08854). Examples of eukaryotic vectors useful in practicing the present invention include the vectors pRc/CMV, pRc/RSV, and pREP (Invitrogen, 11588 Sorrento Valley Rd., San Diego, Calif. 92121); pcDNA3.1/V5&His (Invitrogen); baculovirus vectors such as pVL1392, pVL1393, or pAC360 (Invitrogen); and yeast vectors such as YRP17, YIP5, and YEP24 (New England Biolabs, Beverly, Mass.), as well as pRS403 and pRS413 Stratagene Inc.); *Pichia* vectors such as pHIL-D1 (Phillips Petroleum Co., Bartlesville, Okla. 74004); retroviral vectors such as PLNCX and pLPCX (Clontech); and adenoviral and adeno-associated viral vectors.

Promoters for use in expression vectors of this invention include promoters that are operable in prokaryotic or eukaryotic cells. Promoters that are operable in prokaryotic cells include lactose (lac) control elements, bacteriophage lambda (pL) control elements, arabinose control elements, tryptophan (trp) control elements, bacteriophage T7 control elements, and hybrids thereof. Promoters that are operable in eukaryotic cells include Epstein Barr virus promoters, adenovirus promoters, SV40 promoters, Rous Sarcoma Virus promoters, cytomegalovirus (CMV) promoters, baculovirus promoters such as AcMNPV polyhedrin promoter, *Pichia* promoters such as the alcohol oxidase promoter, and *Saccharomyces* promoters such as the gal4 inducible promoter and the PGK constitutive promoter, as well as neuronal-specific platelet-derived growth factor promoter (PDGF).

In addition, a vector of this invention may contain any one of a number of various markers facilitating the selection of a transformed host cell. Such markers include genes associated with temperature sensitivity, drug resistance, or enzymes associated with phenotypic characteristics of the host organisms.

Host cells expressing the genetic signature of the present invention or functional fragments thereof provide a system in which to screen potential compounds or agents for the ability to modulate the development of pancreatic cancer

Another approach entails the use of phage display libraries engineered to express fragment of the polypeptides encoded by the genetic signature containing nucleic acids

on the phage surface. Such libraries are then contacted with a combinatorial chemical library under conditions wherein binding affinity between the expressed peptide and the components of the chemical library may be detected. US Patents 6,057,098 and 5,965,456 provide methods and apparatus for performing such assays.

5 The goal of rational drug design is to produce structural analogs of biologically active polypeptides of interest or of small molecules with which they interact (e.g., agonists, antagonists, inhibitors) in order to fashion drugs which are, for example, more active or stable forms of the polypeptide, or which, e.g., enhance or interfere with the function of a polypeptide in vivo. See, e.g., Hodgson, (1991) *Bio/Technology* 9:19-21.

10 In one approach, discussed above, the three-dimensional structure of a protein of interest or, for example, of the protein-substrate complex, is solved by x-ray crystallography, by nuclear magnetic resonance, by computer modeling or most typically, by a combination of approaches. Less often, useful information regarding the structure of a polypeptide may be gained by modeling based on the structure of homologous proteins. An example

15 of rational drug design is the development of HIV protease inhibitors (Erickson et al., (1990) *Science* 249:527-533). In addition, peptides may be analyzed by an alanine scan (Wells, (1991) *Meth. Enzym.* 202:390-411). In this technique, an amino acid residue is replaced by Ala, and its effect on the peptide's activity is determined. Each of the amino acid residues of the peptide is analyzed in this manner to determine the important regions

20 of the peptide.

 It is also possible to isolate a target-specific antibody, selected by a functional assay, and then to solve its crystal structure. In principle, this approach yields a pharmacophore upon which subsequent drug design can be based.

 One can bypass protein crystallography altogether by generating anti-idiotypic

25 antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original molecule. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced banks of peptides. Selected peptides would then act as the pharmacophore.

30 Thus, one may design drugs which have, e.g., improved polypeptide activity or stability or which act as inhibitors, agonists, antagonists, etc. of polypeptide activity. By virtue of the availability of the genetic signature containing nucleic acid sequences described herein, sufficient amounts of the encoded polypeptide may be made available to perform such analytical studies as x-ray crystallography. In addition, the knowledge

of the protein sequence provided herein will guide those employing computer modeling techniques in place of, or in addition to x-ray crystallography.

In another embodiment, the availability of genetic signature containing nucleic acids enables the production of strains of laboratory mice carrying the signature(s) of the invention. Transgenic mice expressing the genetic signature of the invention provide a model system in which to examine the role of the protein(s) encoded by the signature containing nucleic acid in the development and progression towards pancreatic cancer. Methods of introducing transgenes in laboratory mice are known to those of skill in the art. Three common methods include: (1) integration of retroviral vectors encoding the foreign gene of interest into an early embryo; (2) injection of DNA into the pronucleus of a newly fertilized egg; and (3) the incorporation of genetically manipulated embryonic stem cells into an early embryo. Production of the transgenic mice described above will facilitate the molecular elucidation of the role that a target protein plays in various cellular metabolic processes. Such mice provide an in vivo screening tool to study putative therapeutic drugs in a whole animal model and are encompassed by the present invention.

The term "animal" is used herein to include all vertebrate animals, except humans. It also includes an individual animal in all stages of development, including embryonic and fetal stages. A "transgenic animal" is any animal containing one or more cells bearing genetic information altered or received, directly or indirectly, by deliberate genetic manipulation at the subcellular level, such as by targeted recombination or microinjection or infection with recombinant virus. The term "transgenic animal" is not meant to encompass classical cross-breeding or in vitro fertilization, but rather is meant to encompass animals in which one or more cells are altered by or receive a recombinant DNA molecule. This molecule may be specifically targeted to a defined genetic locus, be randomly integrated within a chromosome, or it may be extra-chromosomally replicating DNA. The term "germ cell line transgenic animal" refers to a transgenic animal in which the genetic alteration or genetic information was introduced into a germ line cell, thereby conferring the ability to transfer the genetic information to offspring. If such offspring, in fact, possess some or all of that alteration or genetic information, then they, too, are transgenic animals.

The alteration of genetic information may be foreign to the species of animal to which the recipient belongs, or foreign only to the particular individual recipient, or may be genetic information already possessed by the recipient. In the last case, the altered or

introduced gene may be expressed differently than the native gene. Such altered or foreign genetic information would encompass the introduction of genetic signature containing nucleotide sequences.

5 The DNA used for altering a target gene may be obtained by a wide variety of techniques that include, but are not limited to, isolation from genomic sources, preparation of cDNAs from isolated mRNA templates, direct synthesis, or a combination thereof.

10 A preferred type of target cell for transgene introduction is the embryonal stem cell (ES). ES cells may be obtained from pre-implantation embryos cultured in vitro (Evans et al., (1981) Nature 292:154-156; Bradley et al., (1984) Nature 309:255-258; Gossler et al., (1986) Proc. Natl. Acad. Sci. 83:9065-9069). Transgenes can be efficiently introduced into the ES cells by standard techniques such as DNA transfection or by retrovirus-mediated transduction. The resultant transformed ES cells can thereafter be combined with blastocysts from a non-human animal. The introduced ES cells
15 thereafter colonize the embryo and contribute to the germ line of the resulting chimeric animal.

One approach to the problem of determining the contributions of individual genes and their expression products is to use genetic signature associated genes as insertional cassettes to selectively inactivate a wild-type gene in totipotent ES cells (such as those
20 described above) and then generate transgenic mice. The use of gene-targeted ES cells in the generation of gene-targeted transgenic mice was described, and is reviewed elsewhere (Frohman et al., (1989) Cell 56:145-147; Bradley et al., (1992) Bio/Technology 10:534-539).

25 Techniques are available to inactivate or alter any genetic region to a mutation desired by using targeted homologous recombination to insert specific changes into chromosomal alleles. However, in comparison with homologous extra-chromosomal recombination, which occurs at a frequency approaching 100%, homologous plasmid-chromosome recombination was originally reported to only be detected at frequencies between 10^{-6} and 10^{-3} . Non-homologous plasmid-chromosome interactions are more
30 frequent occurring at levels 10^5 -fold to 10^2 fold greater than comparable homologous insertion.

To overcome this low proportion of targeted recombination in murine ES cells, various strategies have been developed to detect or select rare homologous recombinants. One approach for detecting homologous alteration events uses the polymerase chain

reaction (PCR) to screen pools of transformant cells for homologous insertion, followed by screening of individual clones. Alternatively, a positive genetic selection approach has been developed in which a marker gene is constructed which will only be active if homologous insertion occurs, allowing these recombinants to be selected directly. One of the most powerful approaches developed for selecting homologous recombinants is the positive-negative selection (PNS) method developed for genes for which no direct selection of the alteration exists. The PNS method is more efficient for targeting genes which are not expressed at high levels because the marker gene has its own promoter. Non-homologous recombinants are selected against by using the Herpes Simplex virus thymidine kinase (HSV-TK) gene and selecting against its nonhomologous insertion with effective herpes drugs such as gancyclovir (GANC) or (1-(2-deoxy-2-fluoro-B-D arabinofluranosyl)-5-iodou- racil, (FIAU). By this counter selection, the number of homologous recombinants in the surviving transformants can be increased. Utilizing genetic signature containing nucleic acid as a targeted insertional cassette provides means to detect a successful insertion as visualized, for example, by acquisition of immunoreactivity to an antibody immunologically specific for the polypeptide encoded genetic signature nucleic acid(s) and, therefore, facilitates screening/selection of ES cells with the desired genotype.

As used herein, a knock-in animal is one in which the endogenous murine gene, for example, has been replaced with human genetic signature -associated gene(s) of the invention. Such knock-in animals provide an ideal model system for studying the development of pancreatic cancer.

As used herein, the expression of a genetic signature containing nucleic acid, fragment thereof, or genetic signature fusion protein can be targeted in a "tissue specific manner" or "cell type specific manner" using a vector in which nucleic acid sequences encoding all or a portion of genetic signature-associated protein are operably linked to regulatory sequences (e.g., promoters and/or enhancers) that direct expression of the encoded protein in a particular tissue or cell type. Such regulatory elements may be used to advantage for both in vitro and in vivo applications. Promoters for directing tissue specific expression of proteins are well known in the art and described herein.

Methods of use for the transgenic mice of the invention are also provided herein. Transgenic mice into which a nucleic acid containing the genetic signature or its encoded protein(s) have been introduced are useful, for example, to develop screening methods to

screen therapeutic agents to identify those capable of modulating the development of pancreatic cancer.

PHARMACEUTICALS AND PEPTIDE THERAPIES

5 The elucidation of the role played by the gene products described herein in pancreatic cancer progression facilitates the development of pharmaceutical compositions useful for treatment and diagnosis of pancreatic cancer. These compositions may comprise, in addition to one of the above substances, a pharmaceutically acceptable excipient, carrier, buffer, stabilizer or other materials well
10 known to those skilled in the art. Such materials should be non-toxic and should not interfere with the efficacy of the active ingredient.

 Whether it is a polypeptide, antibody, peptide, nucleic acid molecule, small molecule or other pharmaceutically useful compound according to the present invention that is to be given to an individual, administration is preferably in a “prophylactically
15 effective amount” or a “therapeutically effective amount” (as the case may be, although prophylaxis may be considered therapy), this being sufficient to show benefit to the individual.

 As it is presently understood, RNA interference involves a multi-step process. Double stranded RNAs are cleaved by the endonuclease Dicer to generate nucleotide
20 fragments (siRNA). The siRNA duplex is resolved into 2 single stranded RNAs, one strand being incorporated into a protein-containing complex where it functions as guide RNA to direct cleavage of the target RNA (Schwarz et al, Mol. Cell. 10:537 548 (2002), Zamore et al, Cell 101:25 33 (2000)), thus silencing a specific genetic message (see also Zeng et al, Proc. Natl. Acad. Sci. 100:9779 (2003)).

25 Pharmaceutical compositions that are useful in the methods of the invention may be administered systemically in parenteral, oral solid and liquid formulations, ophthalmic, suppository, aerosol, topical or other similar formulations. These pharmaceutical compositions may contain pharmaceutically-acceptable carriers and other ingredients known to enhance and facilitate drug administration. Thus such
30 compositions may optionally contain other components, such as adjuvants, e.g., aqueous suspensions of aluminum and magnesium hydroxides, and/or other pharmaceutically acceptable carriers, such as saline. Other possible formulations, such as nanoparticles, liposomes, resealed erythrocytes, and immunologically based systems may also be used to administer the appropriate agent to a patient according to the methods of the invention.

The use of nanoparticles to deliver agents, as well as cell membrane permeable peptide carriers that can be used are described in Crombez et al., Biochemical Society Transactions v35:p44 (2007).

5 In order to treat an individual having pancreatic cancer, to alleviate a sign or symptom of the disease, the pharmaceutical agents of the invention should be administered in an effective dose. The total treatment dose can be administered to a subject as a single dose or can be administered using a fractionated treatment protocol, in which multiple doses are administered over a more prolonged period of time, for example, over the period of a day to allow administration of a daily dosage or over a longer period of time to administer a dose over a desired period of time. One skilled in the art would know that the amount of agent required to obtain an effective dose in a subject depends on many factors, including the age, weight and general health of the subject, as well as the route of administration and the number of treatments to be administered. In view of these factors, the skilled artisan would adjust the particular dose so as to obtain an effective dose for treating an individual having pancreatic cancer.

15 In an individual suffering from pancreatic cancer, in particular a more severe form of the disease, administration of agent can be particularly useful when administered in combination, for example, with a conventional agent for treating such a disease. The skilled artisan would administer the agent alone or in combination and would monitor the effectiveness of such treatment using routine methods such as sonogram, radiologic, immunologic or, where indicated, histopathologic methods. Other conventional agents for the treatment of pancreatic cancer include anti cancer agents, such as gemcitabine and erlotinib. Administration of the pharmaceutical preparation is preferably in an "effective amount" this being sufficient to show benefit to the individual. This amount prevents, alleviates, abates, or otherwise reduces the severity of pancreatic cancer symptoms in a patient.

20 The pharmaceutical preparation is formulated in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form, as used herein, refers to a physically discrete unit of the pharmaceutical preparation appropriate for the patient undergoing treatment. Each dosage should contain a quantity of active ingredient calculated to produce the desired effect in association with the selected pharmaceutical carrier. Procedures for determining the appropriate dosage unit are well known to those skilled in the art.

Dosage units may be proportionately increased or decreased based on the weight of the patient. Appropriate concentrations for alleviation of a particular pathological condition may be determined by dosage concentration curve calculations, as known in the art.

5

The Examples below are provided to illustrate certain embodiments of the invention. They are not intended to limit the invention in any way.

EXAMPLE I

10

Proteomic Analysis of Pancreatic Cancer Fluids

The following materials and methods are provided to facilitate the practice of the present invention.

15

Sample Acquisition. Aliquots of cyst fluid that were used for this project were obtained from materials that were aspirated for clinical purposes. The study was approved by the Institutional Review Board of the Fox Chase Cancer Center. EUS-FNA (14) was performed under conscious sedation using a linear echoendoscope. When a lesion was identified, EUS-FNA was performed with a 22 or 19-gauge needle through either a transduodenal or transgastric approach, depending on the location of the lesion within the pancreas. The highest priority was given to procuring a volume of fluid that was adequate to perform the necessary clinically indicated diagnostic assays (e.g., cytology, CEA in ng/mL, Mayo Medical Laboratories, code # 84074), amylase (in units/L, Mayo Medical Laboratories, code #5079). As little as 40 μ L of cyst fluids per patient were allocated for the proteomic study. For the purpose of this study, cyst fluid cytology findings were grouped into the following categories: A- Benign: No evidence of benign mucinous epithelium, atypical cells or carcinoma; B- Benign mucinous epithelium; C- Atypical/suspicious cytology; D- Malignant.

20

25

30

Proteomics Analysis. Standard Operating Procedures were established and followed for all steps of cyst fluid collection and analysis. The cyst fluid was diluted with three volumes of PBS, mixed, and centrifuged for 10 minutes at 13,000 x g at 4 °C to remove cells and any insoluble materials, snap frozen in liquid nitrogen in aliquots and

banked at -80 °C. To remove small peptides bound to larger proteins, the cyst fluid was treated with three volumes of 0.1M glycine pH 2.3 and acetonitrile was added to 25% v/v final concentration. The solution was filtered by ultrafiltration (pre-washed Amicon YM-30 Centricon #4208) at 4000 x g at 4 °C for about one hour to reach minimum retention volume designed for the unit. The retained proteins above the filter were solubilized with 200 µL of 0.2% SDS solution and transferred to a 1.5 mL microcentrifuge tube. Three volumes of cold acetone were added to precipitate the proteins overnight at -20 °C and the suspension was then centrifuged at 21,000 x g for 40 min. The pellet was washed once with 80 % cold acetone, centrifuged, and air dried, then resolubilized in 2D PAGE sample buffer (7 M urea, 2 M thiourea, 4% (w/v) CHAPS). Protein concentration was determined as previously described (15).

Protein (15 µg) was reduced with dithiothreitol and alkylated by iodoacetamide at 25 °C for 1 hr (15) and then resolved in a pre-cast Novex 4-12% gradient PAGE with 3 mm wide wells (Invitrogen™, CA, USA). Electrophoresis was performed in MOPS buffer at 150V at room temperature for about 20 min until the tracking dye was 1.5 cm from the top of the gel. The gel cassette was opened in a laminar flow hood. Each sample lane, two per gel, was cut into 11 slices from the well to about 2 mm beyond the dye front. Each gel slice was again subjected to reduction and alkylation. Porcine trypsin (Sigma proteomic grade #T6567) was added as 63 ng in 7 µL 25 mM ammonium bicarbonate and incubated for 30 min. Unabsorbed trypsin of about 2 µL was removed and 20 µL of 25 mM ammonium bicarbonate was added and incubated at 37° C for about 16 hours. 10 µL of the peptide solution was mixed with 2.5 µL of 25% acetonitrile 1% formic acid, and 2 µL was injected into the LC/MS/MS system for protein identification. A LC/MS/MS system consisted of an Applied Biosystems QSTAR XL hybrid quadrupole TOF mass spectrometer supported by an Agilent nanoLC system. For 15 µg gel loading, 10% of the digest of each gel slice was auto-injected onto a trap column (Agilent Zorbax 300SB-C18, 5 µm, 5x0.3 mm), washed, and eluted at 0.3 µL/min through an analytical column (Agilent Zorbax 300SB-C18, 3.5 µm, 150x0.1 mm) at room temperature. The elution gradient was in 0.2% formic acid with linear segments of 4.5%, 4.5%, 28%, 54%, 90% acetonitrile at 0, 4, 8, 80, 85 min, respectively. An IDA protocol using MS periods of 2 s of TOF-MS and three cycles of 4 s of MS/MS each was used to obtain highly accurate spectra for protein identification for the three most intense peptide ions in each cycle. For discovery of more proteins and peptides in cyst fluids and to overcome the

possibility of false-negatives due to under-sampling of co-eluting peptides, an exclusion list of the peptides in the first LC/MS/MS run of a cyst fluid was used to direct the second LC/MS/MS run to sequence new peptides. The two peak lists were combined for database searching for protein identification and for relative quantitation of the proteins by emPAI score (exponentially modified protein abundance index) without isotope labeling (16). The emPAI score, $[10^{(\# \text{ observed peptides} / \# \text{ theoretical peptides})} - 1]$, is roughly proportional to the abundance of a protein in a complex mixture. Almost every protein identified in Tables 1-3 was abundant enough to be identified, and its relative abundance quantified for comparison, in the first LC/MS/MS run.

10

RESULTS

Pancreatic cyst fluids were obtained by EUS from 20 patients for geLC/MS/MS proteomics analysis. The proteins of pancreatic cyst fluids can be subjected to proteolysis in some situations if the pancreatic proteases are inadvertently activated, and if inhibition by serum protease inhibitors is ineffective. To avoid this problem, proteins that were larger than 10,000 in molecular weight were analyzed and quantification performed at the level of tryptic peptides. For the analysis described herein, pancreatic cyst fluids appeared robust and stable, providing the same mass spectrometry information after multiple freeze-thaw cycles. The samples were unaffected by room temperature incubation (data not shown). However, cyst fluids are rich in small peptides bound to other carrier proteins in the sample. Our fractionation procedure removed these small peptides (data not shown), simplified the mass spectra obtained by geLC/MS/MS, and significantly increased the sensitivity of biomarker detection.

Clinical Information on the Cyst Fluids. Demography of the patients, dimensions of the cysts, and the results of traditional clinical tests performed on the cyst fluids are shown in Table 1. In Tables 1 to 4, because CEA measurements by clinical immunoassays are believed to be the strongest indicators of mucinous versus non-mucinous cysts in the absence of direct measurement of the mucins, the cysts are presented in the order of increasing CEA. Two samples without CEA values were located on the right side because high CEA values would be anticipated based on the histopathology findings.

30

All the patients were Caucasians. The various diagnostic assays, commercial amylase and CEA levels, were not obtained for all study patients. These absent values

are represented by empty boxes in Table 1. For cysts 17, 14, 20, 5, and 21, subsequent surgical resection led to definitive histopathologic diagnosis as shown. Cyst 19B, diagnosed by histology after surgical resection as an IPMN adenoma, was the same patient as cyst 19A except the latter occurred five months before the surgical resection, providing a view of the biomarker transition.

The pancreatic cyst fluid proteome. Samples were purified and analyzed by geLC/MS/MS as described in Materials and Methods. The cyst fluids in this study vary in the amounts of plasma proteins versus pancreatic enzymes. About 137 proteins normally found in plasma were observed among 13 of the pancreatic cyst fluids. A partial list of these proteins is shown in Table 2. Hemoglobin, IgG, serum albumin, apolipoprotein A1 and AII, and serotransferrin were among the most abundant serum proteins when present. Hemoglobin was found in significant quantities only in five of the 20 cysts, suggesting that there was minimal contamination of blood from needle puncture during EUS-FNA collection of the cyst fluids. If red blood cells were present, they were successfully removed by centrifugation. Eight of the cyst fluids that contained the most plasma infiltration did not contain pancreatic enzymes. For example cysts 15 and 1 contained only plasma proteins (Table 2), no detectable pancreatic enzymes (Table 3), no mucins, no CEACAM, and no S100 homologs (Table 4). Seven of the cyst fluid samples were essentially free of proteins from blood. Most of these contained abundant pancreatic enzymes. The distribution of some of the 29 pancreatic enzymes among the cysts in this study is shown in Table 3. These enzymes included digestive enzymes and proteins important to pancreatic function. The latter included the pancreatic stone protein Lithostathine 1, the Regenerating islet-derived protein 3 alpha that has multiple functions, and Pancreatic secretory granule membrane major glycoprotein GP2. Amylase is not always observed in cysts that contained abundant levels of other pancreatic enzymes.

Data Analysis. Samples were purified and analyzed by geLC/MS/MS as described above. The mass spectrometry “wiff” data files were used to search the SwissProt protein database release 54.1 using MASCOT 2.2 (Matrix Sciences, London, U.K.), analyzing the MS/MS sequencing spectra of the +2 and +3 ions. Fixed modification of carbamidomethylcysteine, variable oxidation of methionine, and one trypsin miss were allowed for protein identification, but the latter two were disallowed

for calculating the emPAI scores. Peptide mass tolerance was +/-150 ppm and fragment mass tolerance was 0.5 Da. False discovery rate was less than 3.5% for individual peptides as judged by hits at a decoy database with randomized sequences in each entry. Thus the confidence of correct protein identification is very high when three or more
5 unique peptides with high quality sequencing spectra, and from the same position in the gel, are congruent in their identification of a protein in this project.

The presence of major protein classes of blood proteins, pancreatic enzymes, and keratins, in each sample, and the limited number of definitively histopathologically identified samples in this study, confound effective classification of the potential
10 biomarkers by typical statistical approaches that include unsupervised hierarchical clustering (17) and principal component analysis (18). Low abundance proteins with an emPAI score average for the expressing samples of less than 0.01 were first removed, leaving 466 proteins identified with confidence. This emPAI score represents about one peptide sequence identified in a protein of about 250,000 molecular weight thus some of
15 the lower score protein identifications were within the approximately 3% false positive identification rate for this data. Next, 34 keratins, 137 blood proteins, and 29 pancreatic enzymes, were filtered from the proteome of each cyst fluid sample. The remaining 295 proteins were sorted by the average emPAI score calculated from the samples expressing each protein. Among the most abundant ones in this list of pancreatic cyst fluid proteins
20 were the homologs of three families of proteins previously proposed to be biomarkers of pancreatic cancer, namely mucins, CEACAM's (19), and S100's (20-22).

Proteomics Biomarkers. Several biomarkers, some of whose homologs are known to be elevated in pancreatic cancer, were identified in the cyst fluids (Table 4).
25 Ten of the cyst fluids contained one or more mucin homologs, some of which have low amino acid sequence homology to each other. Cyst fluid from seven of the patients revealed the presence of CEACAM homologs by mass spectrometry detection. Five of the patients showed expression of S100 protein homologs in their cyst fluid. The relative abundances of CEACAM5 (CEA) determined by emPAI score were in rough agreement
30 with the clinical assays performed on the samples shown in Table 1, bearing in mind the differences in CEA measurement procedures. More specifically, the emPAI score for CEACAM5 was determined as score per unit protein used in mass spectrometry while the clinical immunoassay CEA unit was concentration in ng per mL cyst fluid. In each case where the identification of a proteomic biomarker was at low abundance in a given

cyst fluid, we ruled out the possibility of sample carry over by verifying that the same biomarker was not detected in the cyst fluid loaded onto the HPLC column in the preceding sample.

5

DISCUSSION

Pancreatic cyst fluid aspirated via EUS-FNA are used clinically to provide biomarkers that facilitate the diagnosis of the potential of pancreatic cancer in patients. The number of assays feasible for each patient is often limited by the quantity of cyst fluids available which is partly a function of the cyst size. For example, the volume of cyst fluids required to submit to either the clinical amylase assay or the clinical CEA assay used in this study is 0.5 mL. Moreover, cytologic diagnosis is facilitated using as large a cyst fluid sample as possible. The scarcity of cyst fluids in cysts smaller than one centimeter in diameter is one of the reasons why such small cysts are often not referred to EUS for evaluation. Thus, a new assay that provides for the measurement of mucins, amylase, and CEA in a minute volume of fluid, provides clinically relevant information in situations where the cyst fluid volumes are small. The proteome of pancreatic cyst fluids as elucidated by LC/MS/MS mass spectrometry proteomics provides comprehensive information on cancer biomarkers in pancreatic cyst fluids using a minimal volume of fluids. Interesting observations made on four classes of biomarkers are described below.

Amylase biomarker. Although the measurement of amylase in blood has been a traditional biomarker of pancreatitis, the basis for using amylase measurements in pancreatic cyst fluid as an indicator of pancreatitis, non-mucinous cyst, or the absence of cancer has not been well studied. We show here that the pancreatic amylase activity in the cyst fluids is divided into two isozymes, alpha amylase 2B and pancreatic alpha amylase, encoded by two separate genes, AMY2B and AMY2A, respectively. The two isozymes have 98% sequence identity, but may differ in their regulation as shown in cysts 8, 10, 19B, and 14. Thus it may be pertinent to consider the levels of the two amylase isozymes individually. Amylase by itself is not always a good indicator of the presence of pancreatic enzymes as in the cases of 19B, cyst 3 and cyst 14 (Table 3). Although pancreatic lipases have been suggested as a substitute for amylase in the analysis of pancreatic cyst fluids, carboxypeptidases A1 and B may be equally effective as indicators of the presence of pancreatic enzymes in this set of samples. The simultaneous measurement of many pancreatic enzymes, made feasible by the use of

mass spectrometry proteomics, may provide more complete information without the limitation of choosing one pancreatic enzyme as biomarker.

Abnormal expression of mucins and changes in their post-translational modification patterns have long been recognized as potential biomarkers of malignancy (3, 23). Table 4 shows that five soluble mucin homologs in cysts 2, 11, 9, and 19 can be distinguished and conveniently measured via LC/MS/MS proteomics and may assist in future classification of cysts. Soluble mucins were detected in these cases where the cytologists were unable to detect mucinous epithelial cells.

CEACAM biomarkers. There are at least seven carcinoembryonic antigen homologs in humans (24, 25). The widely used CEA in clinical tests for various cancers (26) and in pancreatic cyst fluids is CEACAM5 (Table 1). For example, CEA levels of >400 ng/mL appear to be specific for mucin-producing cystic neoplasms (27). However, the CEA levels in these tumors is frequently lower, thus using a cutoff of 400 ng/mL may result in an unacceptably high “miss rate” for diagnosing these potentially malignant tumors (27). Alternatively, a level of 192 ng/mL has been cited as the “optimal” cutoff value (i.e., provides the best combined sensitivity and specificity for distinguishing mucinous from non-mucinous pancreatic cysts); however, this value results in an accuracy rate of only 79% (3). For example, cyst 14 had a CEA level of 582 ng/mL and amylase of 2853 U/L via clinical assays (Table 1) was found to be a MCA upon surgical histopathology. Accordingly, proteomics showed that there was down-regulation of amylase in this enzyme cyst (Table 2) plus higher levels of the proteins of CEA homolog CEACAM6 and CEACAM7 than CEA (Table 4). Thus it appears that a combined high level of the CEACAM homologs are as important an indication as a high level of CEA by itself. As discussed below, the expression of S100A6 and S100A9 in this cyst are indicative of further progression for this neoplasm (Table 4). Cyst 3 with a clinical CEA assay of 63,830 had declined a Whipple procedure and thus had no pathology information. This high CEA value is consistent with its high mucin content and high CEACAM 5 and CEACAM 6 seen in proteomics (Table 4). However, the absence of S100 expression distinguish it from cysts 5, 14, and 21. Thus CEACAM homologs are markers that can assist in risk-stratifying pancreatic cysts.

S100 biomarkers. The S100 protein family includes small Ca^{++} binding proteins that are soluble in 100% saturated ammonium sulfate solution and have long been

recognized as biomarkers of brain cancer. A recent review provides references to the many cellular functions in which S100 homologs appear to participate (20). Although S100S8, S9 and S12 have been implicated as biomarkers of inflammation (28-30), no clinical pancreatitis was observed among the samples used in this study. Lu et al. observed that S100A9 was elevated in pancreatic carcinoma tissue compared with adjacent control tissue, and proposed that other S100 proteins may also serve as markers of pancreatic cancer (31). Recently, Ohuchida et al. extended this finding and showed that S100A6 and S100A11 are also elevated in pancreatic cancer tissue compared with controls and also in the ductal juices (21, 22). The confidence of identification of these S100 homologs A6, A8, A9, and A11 in the cyst fluids in our current study is very high. No peptides overlapped among the 15 peptides sequenced for the four S100 homologs. S100A8 was detected with 5 peptides sequenced and 54% amino acid sequence coverage while homolog S100A9 was distinguished using 6 peptides sequenced and 64% coverage. 43% sequence coverage was obtained for 3 peptides sequenced for S100A11. Only one peptide of excellent sequence quality and reproducibility was detected for S100A6.

Detectable significant levels of CEACAM homologs and S100 homologs in pancreatic cyst fluids, in addition to the presence of mucins and the loss of amylase, are useful biomarkers for the presence or potential of adenoma and carcinoma. This conclusion is supported by their presence in the five cysts, #17, 20, 5, 14, 21 that were confirmed by histopathology and cytopathology to be adenomas or carcinomas, but not at significant levels in cysts #13, 16, 18, 7, 19, 15, and 1. Cyst 5, an adenocarcinoma, is similar in cyst fluid proteome to non-mucinous cysts except for the presence of significant levels of CEACAM1, CEACAM5, and four S100 homologs. Cyst 21, a MCA, is similar in cyst fluid proteome to cyst 5 but with less of these proteins. For cyst 19B, a MCA, multiple potential biomarkers of mucins, CEACAM, and S100 are apparent. Five months earlier, when cyst 19B was aspirated as cyst 19A, mass spectrometry had detected two CEACAMs and multiple mucins, consistent with the suggestion that CEA of greater than 192 ng/mL may indicate the presence of mucinous neoplasm (3). Thus S100 homologs are also useful markers of cyst progression.

Although the high mass accuracy and platform stability of the mass spectrometer used in this study facilitated biomarker discovery, once the protein names are known, another mass spectrometry method called Multiple Reaction Monitoring (MRM) can accurately quantify multiple biomarkers against internal standards at the same time with

much higher sensitivity (32-35). Importantly, the quantitation of mucin homologs, CEACAM homologs, S100 homologs, amylase, and other marker proteins, can be performed at the same time using the same method, so their combination will provide valuable diagnostic and prognostic information to the clinician. The biomarkers described herein are obtainable from less than 40 μ L of cyst fluids and detection of their presence facilitates earlier pancreatic cancer detection in cysts than heretofore previously possible.

EXAMPLE 2

10 **Xenograft model of pancreatic cancer and pancreatic cyst fluid secretion**

The use of clinical samples for studying the biology of pancreatic cyst to cancer is difficult because of inability to obtain time course material in most instances, and because most invasive techniques of laboratory investigation cannot be used on patients. A mouse model, if valid and available, can accelerate pancreatic cancer research. For example, mouse stroma cells infiltrating the tumor and supports the tumor growth can be marked with Green Fluorescence Protein by using a GFP transgenic mouse as the host of the xenograft.

In the laboratory of Dr. Repasky, about 33% of pancreatic tumors engraftment resulted in successful tumor propagation for three passages for both adenocarcinomas and neuroendocrine tumors (38, 39). These xenograft tissues contain complex cell types from both cancer and stroma. The cancer cells form glands that hold secretions similar to what is seen in the parent tumors. We obtained three mice for use that harbored tumors soft to the touch. These mice are normally not used for preclinical drug treatment experiments. About 500 μ L of fluids were collected from each of these xenograft and we analyzed them using the same protocol of GelC/MS/MS described in Example 1.

The high mass accuracy and sensitivity of our QSTAR XL mass spectrometer using LC/MS/MS produced numerous peptide sequences, many of which easily distinguished human homologs of proteins from mouse homologs. This ability allowed us to interpret whether and how much mouse plasma and stroma are infiltrating the human tumors, and the origin of biomarker proteins detected in the xenograft fluids. Although these liquid cysts are often believed to result from necrosis, a process called cystic degeneration of tumors, in our mass spectrometry analysis of three independent xenograft fluids, we saw no evidence of general release of high abundance cellular

proteins normally happening in necrosis. Similar to the pancreatic cyst fluids that did not contain pancreatic enzyme secretion, the xenograft fluids contained hundreds of plasma infiltrate proteins of mouse origin. However, human proteins were present, including abundant levels of pancreatic cyst fluid neoplasm biomarker proteins described in

- 5 Example 1: mucin 1, mucins 5AC, mucin 5B, CEA, CEACAM 6, and S100A6 and S100A11. (Tables 3 and 4). Thus these "cancer cells" were secreting the same biomarkers as for the pancreatic cyst fluids from patients harboring cystadenoma, IPMN, and adenocarcinoma. Several other proteins in Table 5 below, not assigned as biomarkers in Example 1 are also seen in both cyst fluid and xenograft fluids, indicating
- 10 that they can be functional biomarkers as well. Validating that our cyst fluid biomarkers can be secreted by pancreatic cancer xenograft was exciting, but cyst proteins conspicuously absent were CEACAM7, S100A8 and A9.

Table 5

Swiss-Prot Entry name	Title (Bold protein names are also found in pancreatic cyst fluids harboring cystadenoma, IPMN, or adenocarcinoma but not in benign cysts)	Mass (Da)	emPAI score		
			Xeno-graft Sample 1	Xeno-graft Sample 2	Xeno-graft Sample 3
ANXA2_HUMAN	Annexin A2	38808	6.61	9.04	3.38
S10A6_HUMAN	Protein S100-A6	10230	1.71	0.94	0.94
EZRI_HUMAN	Ezrin	69484	1.42	1.83	0.52
LG3BP_HUMAN	Galectin-3-binding protein	66202	1.03	1.15	0.64
S10AB_HUMAN	Protein S100-A11	11847	0.78	3.23	1.37
GELS_HUMAN	Gelsolin	86043	0.52	0.52	0.13
MOES_HUMAN	Moesin	67892	0.45	0.53	0.17
ANXA1_HUMAN	Annexin A1	38918	0.44	0.58	0.74
PIGR_HUMAN	Polymeric-immunoglobulin receptor	84429	0.41	0.9	0.24
NGAL_HUMAN	Neutrophil gelatinase-associated lipocalin	22745	0.36	3.04	0.86
ANXA3_HUMAN	Annexin A3	36524	0.34	0.34	0.48
CEAM6_HUMAN	Carcinoembryonic antigen-related cell adhesion molecule 6	37499	0.33	0.46	0.21
1433S_HUMAN	14-3-3 protein sigma	27871	0.29	0.89	0.29
MUC5A_HUMAN	Mucin-5AC	135404	0.27	1.01	0.21
CEAM5_HUMAN	Carcinoembryonic antigen-related cell adhesion molecule 5	77489	0.21	0.26	0.15

MUC1_HUMAN	Mucin-1	122170	0.2	0.27	0.16
MUC13_HUMAN	Mucin-13	55710	0.14	0.3	0.21
MUC5B_HUMAN	Mucin-5B	605803	0.08	0.16	0.21
AGR2_HUMAN	Anterior gradient protein 2	20024	1.41	3.09	0.42
ANXA5_HUMAN	Annexin A5	35971	1.22	1.7	0.82

Table 5. A partial list of proteins found in xenograft fluids from all three separate xenograft experiments. The numbers under each sample are EMPAI scores roughly proportional to the protein abundance in the sample. For comparison, serum albumin from mouse plasma infiltration or blood contamination has an average value of 12. The first 18 proteins, in bold, are also found among patient-derived pancreatic cyst fluids when cystadenoma, IPMN, or adenocarcinoma were indicated.

Other biomarkers of interest include anterior gradient protein 2 which has been proposed as a marker of pancreatic cancer tissue because of its over-expression in most pancreatic cancers (40). Another interesting biomarkers is NGAL (Neutrophil gelatinase-associated lipocalin), a new early biomarker of acute kidney injury in rats. Its level in blood rises within two hours of renal injury. The protein is a member of the large lipocalin family of extracellular proteins which transports or binds small hydrophilic molecules, but when located inside a cell may become protease inhibitors. Its role in pancreatic cyst fluids is may be partly associated with inflammation.

Proteomics has resulted in the identification of biomarkers present in cysts, a better understanding of the basic biological features of cysts and their natural history, thereby providing a better understanding of the molecular profile within these cysts. Such information can be used to advantage to identify clinically relevant targets for early diagnosis and treatment of pancreatic cancer.

EXAMPLE 3

As mentioned above, pancreatic cancer kills about 40,000 patients each year. Absent the present discovery, there is no early detection. Current diagnosis is neither completely accurate nor confident, with both unavoidable false positives and false negatives. There are several types of pancreatic cancer with different biology and outcome. Some rarer varieties are less aggressive than adenocarcinoma which comprises 85% of pancreatic cancer arising in the pancreatic duct. Not all adenocarcinoma are observed to originate from a cyst. While most pancreatic cysts are benign in the short term, there is no certainty that the cyst

which appears benign today will still be benign a few years later. Thus all pancreatic cyst patients are followed periodically with more scanning.

A patient suspected of having a mucinous pancreatic cyst is often referred for surgical resection because of significant risk that cancer is present. Liquid from the cyst is collected by a needle that goes through the stomach wall in “endoscopic ultrasound – fine needle aspiration (EUS-FNA) (Figure 5). A misdiagnosis of mucin in the cyst can dramatically change the fate of an otherwise healthy individual. It is an object of the present method to remove this uncertainty as mucin 5B is a pancreas-specific mucin marker that cannot come from stomach contamination.

A second concern is the staging of pancreatic cancer development. The results presented in Example I correlated about 30 biomarker proteins whose mechanisms suggest that they may be important in different stages of progression to cancer.

The third concern is the detection of adenocarcinoma while it is still small. Cancer can give rise to expression of both cancer-causing biomarkers and/or reporter proteins induced by the presence of cancer.

Cystic Neoplasms of the Pancreas.

Cystic neoplasms of the pancreas arise from the pancreatic ductal epithelium and produce fluids (e.g., mucinous or serous fluids) that lead to the formation of cystic cavities within the tumor (Figure 5). Cystic neoplasms of the pancreas are initially benign lesions. However, those which produce mucin, broadly referred to as mucinous cystic neoplasms, can progress through a series of histological stages with the eventual development of adenocarcinoma. These stages have been defined by the World Health Organization (7, 8). Some cysts may drain into the pancreatic duct.

The biologic nature and histopathologic features of pancreatic cysts are varied (3, 6) (Figure 6). Ten to twenty percent of pancreatic cysts are neoplastic, including neoplasms which grow as cystic structures (i.e., primary cystic neoplasms of the pancreas), and solid neoplasms that have undergone cystic degeneration. Serous cystadenomas (microcystic adenomas) account for approximately 32-39% of the primary cystic neoplasms and have very low malignant potential. Mucinous cystic neoplasms which includes mucinous cystadenomas (MCAs) and intraductal papillary mucinous neoplasms (IPMNs) are a subgroup of primary cystic neoplasms that have malignant potential (7, 8), accounting for approximately 10-45% and 21-33% of primary cystic neoplasms, respectively (6, 9-11). Two subtypes of IPMN have been described (1, 12), a main duct variant and a branch duct variant; the latter may have a more indolent course.

In the absence of reliable methods of quantifying the malignant potential of suspected pre-malignant cystic neoplasms of the pancreas, existing clinical approaches include recommendation of partial or total pancreatectomy even if only a single cyst is present. This approach is associated with significant morbidity and about 1% mortality (13) in the best of hands.

Diagnosis of Cystic Lesions of the Pancreas.

The sensitivity of pancreatic cyst fluid cytology has been reported as only 27-64%. The endosonographic morphologic findings by themselves are inadequate for distinguishing amongst the various cyst types or determining whether the process of malignant transformation has begun (36, 37). For example, elevated fluid amylase levels may be only 55-60% specific for differentiating pseudocysts from certain cystic neoplasms, and the sensitivity of fluid cytology has been reported to be only 27-64% in most series. Immunoassays of a variety of tumor markers, e.g., CEA (carcinoembryonic antigen CEACAM5, a carcinoembryonic antigen-related cell adhesion molecule) (38), CA 19-9 (carbohydrate antigen 19-9), CA 15-3 (cancer antigen MUC1, mucin 1) may distinguish mucinous from non-mucinous cystic lesions, and also may have predicted whether a cyst harbors areas of malignant transformation (3, 4, 5); however, optimal threshold values for each tumor marker have not been well-established, limiting their utility in evaluating these lesions. For example, pancreatic cyst fluid CEA levels over 192 ng/mL to 400 ng/mL appear to be specific for mucin-producing cystic neoplasms (27) most of the time (3), but not always. Moreover, the CEA levels in these tumors is frequently lower, thus using a cutoff of 400 ng/mL may result in an unacceptably high “miss rate” for diagnosing these potentially malignant tumors (27). Thus a method to directly quantify the mucins present would be more ideal if stomach contamination were not a problem. Other investigators have begun to evaluate the role of other molecular assays, including K-ras point mutations, telomerase activity, mucin isoform analysis, Claudin 4, CXCR4, S100A4, mesothelin, tumor suppressor gene methylation and alterations (e.g., p16, p53, DPC4) and mutational allelotyping of PCR-amplified DNA from cyst fluid. None of above assays is able to determine biomarker levels in the small quantities of sample used for mass spectrometry.

Mucin 5B.

Most studies of mucin 5 in cancer are focused on mucin 5AC which is often more abundant than mucin 5B. Mucin 5B is expressed in the normal salivary gland (39) and

airway (40) but is emerging as a potential marker of some tumors, in particular, gastric cancer (41, 42), lung cancer (43), breast cancer (44), nasopharyngeal cancer (45), bladder (46), but apparently not ovarian cancer (47) or colon cancer (48). Altered mucin 5AC expression was not predictive of an increased risk for pancreatic cancer while mucin 2 is
5 (49). Mucin 5B was not investigated in this study. A monoclonal antibody LUM5B-2 to the non-glycosylated domains of mucin 5B has been made in Sweden (50). However, mucin5B from pancreas has not yet been characterized with this antibody. The present approach entails the use of mass spectrometry to avoid any possible differences between the glycosylation of mucin 5B in saliva versus in pancreas that may lead to different
10 antibody reactivity.

Mucin 5B is not found in the mucosa of normal stomach. Other studies reached the same conclusion in the normal gastric mucosa are reported in the tissue gene expression database of SAGE and the in situ hybridization study of Perrais et al. (42) and Buisine et al. (51). On the other hand, mucin 5AC and mucin 6 are the major gastric
15 mucins (40).

Proteomics of Pancreatic Cyst Fluids

Example I describes the proteome of pancreatic cyst fluid. A total of about 30 biomarker proteins, including homologs of mucins, CEACAMs, S100s, can be employed
20 to determine whether a cyst is mucinous or cancerous. Although our panel of samples was small, the conclusions correlated with CEA levels from immunoassays and pathology diagnosis where available. An important aspect of this study was the demonstration that numerous cyst fluid biomarker proteins can be quantified at the same time from a tiny amount of cyst fluid, thereby opening the door for future quantitative
25 mass spectrometry to provide early detection. Pertinent to this proposal was our discovery of ample amounts of mucin 5B in several cysts that were known to be mucinous neoplasms, indicating the mucin 5B may be a good biomarker for excluding the possibility of stomach mucin contamination.

GeLC/MS/MS proteomics was used in Example I to show that some cyst fluids
30 are abundant with about 250 serum infiltrate proteins while others are abundant with about 70 pancreatic enzymes and some have both. By subtracting these protein names, we arrived at about 250 proteins that could be correlated with the clinical information on the cysts and gave about 30 potential biomarkers. See Figures 2, 3, and 4.

Ten of the cyst fluids contained one or more soluble mucin homologs. Four soluble mucin homologs can be distinguished. Mucin 1 is MUC1, CA15-3, a known pancreatic cancer marker. Mucin 1 and mucin 5AC are found in the stomach as well as in the pancreas and thus can potentially appear as contaminant at low levels in pancreatic cyst fluids due to the Fine Needle Aspiration puncturing through the stomach wall. However, mucin 5B is found in the pancreas but not in the stomach. We have observed cyst fluids in which mucin 5B was the second most abundant protein. Importantly, this specificity of identification of soluble mucin 5B by mass spectrometry overcomes the concern of potential gastric contamination. The absolute specificity of mass spectrometry is demonstrated by the fact that none of the 11 unique non-glycosylated peptides which identified mucin 5AC with confidence were found among the 33 unique non-glycosylated peptides that identified mucin 5B.

Interestingly, CEACAM 5, 6, 7, 8 are only found in primates and not in rodents, and literature implicated the levels of CEACAM6 and CEACAM 7 as opposite to each other.

The S100 protein family includes small Ca^{++} binding proteins that are soluble in 100% saturated ammonium sulfate solution and participate in many cellular functions (28, 29, 30) including tumor promotion. S100A9, A6 and A11 were elevated in pancreatic carcinoma tissue. S100A6 was elevated in the ductal epithelium of pancreatic cancer obtained by laser-capture microdissection. Our study extended S100 detection to pancreatic cyst fluids. No peptides were identical among the 17 peptides sequenced for the four S100 homologs. Cyst 21 may be untypical, a false diagnosis, or an experimental error.

To further confirm our biomarkers as originating from pancreatic cancer tissue, in unpublished studies, we examined the proteome of three samples of fluids produced by the xenograph implantation of pancreatic cancer in mice. These fluids contained the same biomarkers as observed in the pancreatic cyst fluids of human mucinous cysts and adenocarcinoma.

Quantitation of Biomarker Proteins by Multiple Reaction Monitoring.

We have prepared the synthetic protein that contained three peptides for each biomarker protein, and can be isotope-labeled as internal standards for quantitative mass spectrometry. A Thermo Quantum Access TSQ triple-quadrupole mass spectrometer has been

purchased in Fox Chase Cancer Center and will be used for further validation and quantitation of the biomarkers in our samples.

Methods

5 *Sample Acquisition*

Dr. John Hoffman and partners perform the majority of pancreaticobiliary and gastric surgeries at Fox Chase Cancer Center. He and the interventional endoscopists Drs. Jeffrey Tokar, Oleh Haluszka, and others can collect about 50 pancreatic cyst fluid samples during EUS and post-surgery (free of stomach contamination) and about 50
10 gastric samples a year for proteomic analyses. Human subjects are not involved in this proposed research because we are using coded private information/data and/or coded human biological specimens. IRB approval for these specific experiments will be obtained.

GeLC/MS/MS Comparison Of Protein Expression.

15 While mRNA gene expression literature asserts that mucin 5B is not made in the stomach mucosa, we intend to confirm this finding at the protein level. Here we propose to use samples of stomach mucosa from 30 patients to confirm at the protein level that mucin 5B is absent from the normal stomach mucosa, thus ensuring that mucin 5B seen in pancreatic cyst fluid does not from stomach contamination. We will collect multiple
20 specimens per patient because the location of any given lesion is what determines what part of the stomach the EUS needle must traverse – can be body, antrum, rarely fundus. The detection of mucin 5B will entail use of a tandem mass spectrometry in a protocol called GeLC/MS/MS as described below. This procedure is routine in our laboratory.

25 *Method*

15 µg of protein obtained from normal stomach mucosa solubilized by SDS and dithiothreitol will again be reduced with dithiothreitol and alkylated by iodoacetamide at 25 °C for 1 hr and then resolved in a pre-cast Novex 4-12% gradient PAGE with 3 mm wide wells (Invitrogen™, CA, USA). Electrophoresis was performed in MOPS buffer
30 at 150V at room temperature until the tracking dye was 1.5 cm from the top of the gel. The gel cassette was opened in a laminar flow hood. Each sample lane, two per gel, was cut into about 11 slices from the well to about 2 mm beyond the dye front. Each gel slice was again subjected to reduction and alkylation. Porcine trypsin (Sigma proteomic grade #T6567) was added as 63 ng in 7 µL 25 mM ammonium bicarbonate and incubated for

30 min. Unabsorbed trypsin of about 2 μ L was removed and 20 μ L of 25 mM ammonium bicarbonate was added and incubated at 37°C for about 16 hours. 10 μ L of the peptide solution was mixed with 2.5 μ L of 25% acetonitrile 1% formic acid, and 2 μ L was injected into the LC/MS/MS system for protein identification as described in

5 Example I. For discovery of more proteins and peptides in a sample and to overcome the possibility of false-negatives due to under-sampling of co-eluting peptides, after the first LC/MS/MS run of a sample, an exclusion list composed of the peptides sequenced in the first LC/MS/MS run was assembled and used to direct another LC/MS/MS run to

10 sequence only new peptides in the same sample. The two peak lists were combined for database searching for protein identification and for relative quantitation of the proteins by emPAI score (exponentially modified protein abundance index) without isotope labeling (47). The emPAI score, $[10^{(\text{number of observed peptides per protein/number of theoretical peptides per protein}) - 1}]$, is roughly proportional to the abundance of a protein in a complex mixture and is believed to be more accurate than a routine dye-

15 binding protein assay. Details of the HPLC and mass spectrometry operation were previously described in Example I.

Data Analysis.

The mass spectrometry “wiff” data files will be used to search the SwissProt protein database release 54.1 using MASCOT 2.2 (Matrix Sciences, London, U.K.),

20 analyzing the MS/MS sequencing spectra of the +2 and +3 ions. Fixed modification of carbamidomethylcysteine, variable oxidation of methionine, and one trypsin miss were allowed for protein identification in GeLC/MS/MS, but the latter two will be disallowed for calculating the emPAI scores. Peptide mass tolerance will be +/-150 ppm and fragment mass tolerance was 0.5 Da. The protein identifications require more than one

25 “bold red” peptide for each protein. False discovery rate due to coincidence in database will be less than 3.5% for individual peptides as judged by hits at a decoy database containing randomized sequences in each entry. We anticipate identification of 500 to 1000 proteins in these samples.

For better assurance of detection dynamic range, we will also be developing a

30 MRM mass spectrometry procedure.

We can also determine the particular stage of pancreatic lesion marked by mucin 5B expression. For example, we will determine those mucin 5B levels associated with cancer, or indicative of MCN and IPMN lesions that are still benign at the time of sample collection. To start to do this, we will analyze 21 samples that we already have that each

has diagnosis of the lesion made by surgical pathology. Each of these samples is from a dilated pancreatic duct downstream from a pancreatic lesion including MCN, IPMN, or adenocarcinoma. Some of the samples are control ductal fluids from pancreatitis patients.

5 *Data analysis:*

The dominance of two major protein classes in pancreatic ductal fluids, blood proteins versus pancreatic enzymes, does not allow effective classification of the potential biomarkers by typical statistical approaches that include unsupervised hierarchical clustering and principal component analysis. Therefore, low abundance proteins with an emPAI score average for the expressing samples of less than 0.01 will first be removed, leaving several hundred proteins with confident identification. Next, 34 keratins, 137 blood proteins, and 29 pancreatic enzymes, will be filtered from the proteome of each ductal fluid sample. The remaining list of about 300 proteins can be analyzed by cluster analysis and principal component analysis. Among the most abundant ones in this list of pancreatic ductal fluid proteins is anticipated to be three families of proteins some of whose homologs were previously proposed to be biomarkers of pancreatic cancer, namely mucins, CEACAM's (50), and S100's and which we found to be potential biomarkers in pancreatic cyst fluids.

We anticipate that biomarker quantitation can be improved over the EMPAI score (spectral counting) approach via use of a qTOF instrument which will increase the dynamic range of quantitation. Thus in future, we will establish the MRM (multiple

MRM (Multiple Reaction Monitoring, also called SRM for Selected Reaction Monitoring is a mass spectrometry biomarker assay which simultaneously monitors multiple biomarkers and is widely believed to be the way of the future for clinical diagnostics. The MRM mass spectrometry method proposed to measure the biomarker peptides has been demonstrated to be reproducible in multiple laboratories in parallel. Based on data obtained to date, we have designed groups of three peptides that represent each of our biomarkers to be used as internal standards in quantitation assays. We have made these peptides stoichiometrically by joining them into a recombinant protein coded for by a synthetic gene which can be used to make isotope-labeled peptides to be used as internal standards. This recombinant protein has been cloned and expressed in *E. coli*. We have also installed a Thermo Quantum Access TSQ triple-quad mass spectrometer. We will

also employ qTOF mass spectrometry. The two datasets will be complementary. Figure 7 shows the construction of the synthetic gene of the internal standard peptides.

In conclusion, we have provided the means for improving the current diagnosis for mucinous cysts, without false positives and false negatives and for cysts that yield little fluid. The compositions and methods disclosed herein may rescue a subset of the pancreatic patients for whom early detection at the small cystic stages can lead to a cure by surgery. It may also substantially decrease the cost of health care for patients with pancreatic cysts for whom current diagnosis is not successful, are perhaps false positives, and prompt future intervention methods that may become feasible at the earlier stages of the disease.

References

1. Yeo, C. J., and Sarr, M. G. Cystic and pseudocystic diseases of the pancreas. *Curr Probl Surg*, 31: 165-243, 1994.
2. Shami, V. M., Sundaram, V., Stelow, E. B., Conaway, M., Moskaluk, C. A., White, G. E., Adams, R. B., Yeaton, P., and Kahaleh, M. The level of carcinoembryonic antigen and the presence of mucin as predictors of cystic pancreatic mucinous neoplasia. *Pancreas*, 34: 466-9, 2007.
3. Brugge, W. R., Lewandrowski, K., Lee-Lewandrowski, E., Centeno, B. A., Szydlo, T., Regan, S., del Castillo, C. F., and Warshaw, A. L. Diagnosis of pancreatic cystic neoplasms: a report of the cooperative pancreatic cyst study. *Gastroenterology*, 126: 1330-6, 2004.
4. Lewandrowski, K. B., Southern, J. F., Pins, M. R., Compton, C. C., and Warshaw, A. L. Cyst fluid analysis in the differential diagnosis of pancreatic cysts. A comparison of pseudocysts, serous cystadenomas, mucinous cystic neoplasms, and mucinous cystadenocarcinoma. *Ann Surg*, 217: 41-7, 1993.
5. Levy, M., Levy, P., Hammel, P., Zins, M., Vilgrain, V., Amouyal, G., Amouyal, P., Molas, G., Flejou, J. F., Voitot, H., and et al. [Diagnosis of cystadenomas and cystadenocarcinomas of the pancreas. Study of 35 cases]. *Gastroenterol Clin Biol*, 19: 189-96, 1995.
6. Brugge, W. R., Lauwers, G. Y., Sahani, D., Fernandez-del Castillo, C., and Warshaw, A. L. Cystic neoplasms of the pancreas. *N Engl J Med*, 351: 1218-26, 2004.
7. Hamilton, S. R., and Aaltonen, L. A. WHO Classification of Tumours, Pathology and Genetics of Tumours of Digestive System: . Lyon, France: IARC Press, 2000.
8. Kloppel, G., Solcia, E., Longnecker, D., Capella, C., and Sobin, L. Histological typing of tumours of the exocrine pancreas: World Health Organization international histological classification of tumours: Springer-Verlag, 1998.
9. Wilentz, R. E., Albores-Saavedra, J., Zahurak, M., Talamini, M. A., Yeo, C. J., Cameron, J. L., and Hruban, R. H. Pathologic examination accurately predicts prognosis in mucinous cystic neoplasms of the pancreas. *Am J Surg Pathol*, 23: 1320-7, 1999.
10. Kloppel, G. Clinicopathologic view of intraductal papillary-mucinous tumor of the pancreas. *Hepatogastroenterology*, 45: 1981-5, 1998.
11. Sarr, M. G., Carpenter, H. A., Prabhakar, L. P., Orchard, T. F., Hughes, S., van Heerden, J. A., and DiMagno, E. P. Clinical and pathologic correlation of 84 mucinous cystic neoplasms of the pancreas: can one reliably differentiate benign from malignant (or premalignant) neoplasms? *Ann Surg*, 231: 205-12, 2000.
12. Yeo, T. P., Hruban, R. H., Leach, S. D., Wilentz, R. E., Sohn, T. A., Kern, S. E., Iacobuzio-Donahue, C. A., Maitra, A., Goggins, M., Canto, M. I., Abrams, R. A., Laheru, D., Jaffee, E. M., Hidalgo, M., and Yeo, C. J. Pancreatic cancer. *Curr Probl Cancer*, 26: 176-275, 2002.
13. Cullen, J. J., Sarr, M. G., and Ilstrup, D. M. Pancreatic anastomotic leak after pancreaticoduodenectomy: incidence, significance, and management. *Am J Surg*, 168: 295-8, 1994.

14. Jacobson, B. C., Adler, D. G., Davila, R. E., Hirota, W. K., Leighton, J. A., Qureshi, W. A., Rajan, E., Zuckerman, M. J., Fanelli, R. D., Baron, T. H., and Faigel, D. O. ASGE guideline: complications of EUS. *Gastrointest Endosc*, *61*: 8-12, 2005.
15. Li, X. M., Patel, B. B., Blagoi, E. L., Patterson, M. D., Seechozer, S. H., Zhang, T., Damle, S.,
5 Gao, Z., Boman, B., and Yeung, A. T. Analyzing alkaline proteins in human colon crypt proteome. *J Proteome Res*, *3*: 821-33, 2004.
16. Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., and Mann, M. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics*, *4*: 1265-72, 2005.
- 10 17. Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A*, *95*: 14863-8, 1998.
18. Peterson, L. E. Partitioning large-sample microarray-based gene expression profiles using principal components analysis. *Comput Methods Programs Biomed*, *70*: 107-119, 2003.
19. Schmid, R. M. [Pancreatic cancer]. *Schweiz Rundsch Med Prax*, *95*: 1709-12, 2006.
- 15 20. Salama, I., Malone, P. S., Mihaimed, F., and Jones, J. L. A review of the S100 proteins in cancer. *Eur J Surg Oncol*, 2007.
21. Ohuchida, K., Mizumoto, K., Ishikawa, N., Fujii, K., Konomi, H., Nagai, E., Yamaguchi, K., Tsuneyoshi, M., and Tanaka, M. The role of S100A6 in pancreatic cancer development and its clinical implication as a diagnostic marker and therapeutic target. *Clin Cancer Res*, *11*: 7785-93, 2005.
- 20 22. Ohuchida, K., Mizumoto, K., Ohhashi, S., Yamaguchi, H., Konomi, H., Nagai, E., Yamaguchi, K., Tsuneyoshi, M., and Tanaka, M. S100A11, a putative tumor suppressor gene, is overexpressed in pancreatic carcinogenesis. *Clin Cancer Res*, *12*: 5417-22, 2006.
23. Hammel, P. R., Forgue-Lafitte, M. E., Levy, P., Voitot, H., Vilgrain, V., Flejou, J. F., Molas, G., Gespach, C., Ruzsniwski, P., Bernades, P., and Bara, J. Detection of gastric mucins (M1 antigens) in cyst
25 fluid for the diagnosis of cystic lesions of the pancreas. *Int J Cancer*, *74*: 286-90, 1997.
24. Kuespert, K., Pils, S., and Hauck, C. R. CEACAMs: their role in physiology and pathophysiology. *Curr Opin Cell Biol*, *18*: 565-71, 2006.
25. Redefined Nomenclature for Members of the Carcinoembryonic Antigen Family.pdf.
26. Gold, P., and Freedman, S. O. Specific carcinoembryonic antigens of the human digestive system.
30 *J Exp Med*, *122*: 467-81, 1965.
27. Hammel, P., Voitot, H., Vilgrain, V., Levy, P., Ruzsniwski, P., and Bernades, P. Diagnostic value of CA 72-4 and carcinoembryonic antigen determination in the fluid of pancreatic cystic lesions. *Eur J Gastroenterol Hepatol*, *10*: 345-8, 1998.
28. Roth, J., Goebeler, M., and Sorg, C. S100A8 and S100A9 in inflammatory diseases. *Lancet*, *357*:
35 1041, 2001.
29. Ryckman, C., Vandal, K., Rouleau, P., Talbot, M., and Tessier, P. A. Proinflammatory activities of S100: proteins S100A8, S100A9, and S100A8/A9 induce neutrophil chemotaxis and adhesion. *J Immunol*, *170*: 3233-42, 2003.
30. Leach, S. T., Yang, Z., Messina, I., Song, C., Geczy, C. L., Cunningham, A. M., and Day, A. S. Serum and mucosal S100 proteins, calprotectin (S100A8/S100A9) and S100A12, are elevated at diagnosis
40 in children with inflammatory bowel disease. *Scand J Gastroenterol*: 1-11, 2007.
31. Lu, Z., Hu, L., Evers, S., Chen, J., and Shen, Y. Differential expression profiling of human pancreatic adenocarcinoma and healthy pancreatic tissue. *Proteomics*, *4*: 3975-88, 2004.
32. Zappacosta, F., Collingwood, T. S., Huddleston, M. J., and Annan, R. S. A quantitative results-driven approach to analyzing multisite protein phosphorylation: the phosphate-dependent phosphorylation
45 profile of the transcription factor Pho4. *Mol Cell Proteomics*, *5*: 2019-30, 2006.
33. Keshishian, H., Addona, T., Burgess, M., Kuhn, E., and Carr, S. A. Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. *Mol Cell Proteomics*, 2007.
- 50 34. Wolf-Yadlin, A., Hautaniemi, S., Lauffenburger, D. A., and White, F. M. Multiple reaction monitoring for robust quantitative proteomic analysis of cellular signaling networks. *Proc Natl Acad Sci U S A*, *104*: 5860-5, 2007.
35. Anderson, L., and Hunter, C. L. Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Mol Cell Proteomics*, *5*: 573-88, 2006.
- 55 36. Ahmad NA, Kochman ML, Brensinger C, et al. Interobserver agreement among endosonographers for the diagnosis of neoplastic versus non-neoplastic pancreatic cystic lesions. *Gastrointest Endosc* 2003; *58*: 59-64.
37. Ahmad NA, Kochman ML, Lewis JD, Ginsberg GG. Can EUS alone differentiate between malignant and benign cystic lesions of the pancreas? *Am J Gastroenterol* 2001; *96*: 3295-300.

38. Shami VM, Sundaram V, Stelow EB, et al. The level of carcinoembryonic antigen and the presence of mucin as predictors of cystic pancreatic mucinous neoplasia. *Pancreas* 2007; 34: 466-9.
39. Wickstrom C, Davies JR, Eriksen GV, Veerman EC, Carlstedt I. MUC5B is a major gel-forming, oligomeric mucin from human salivary gland, respiratory tract and endocervix: identification of glycoforms and C-terminal cleavage. *Biochem J* 1998; 334 (Pt 3): 685-93.
- 5 40. Andrianifahanana M, Moniaux N, Batra SK. Regulation of mucin expression: mechanistic aspects and implications for cancer and inflammatory diseases. *Biochim Biophys Acta* 2006; 1765: 189-222.
41. Pinto-de-Sousa J, Reis CA, David L, Pimenta A, Cardoso-de-Oliveira M. MUC5B expression in gastric carcinoma; relationship with clinico-pathological parameters and with expression of mucins MUC1, MUC2, MUC5AC and MUC6. *Virchows Arch* 2004; 444: 224-30.
- 10 42. Perrais M, Pigny P, Buisine MP, Porchet N, Aubert JP, Van Seuningem-Lempire I. Aberrant expression of human mucin gene MUC5B in gastric carcinoma and cancer cells. Identification and regulation of a distal promoter. *J Biol Chem* 2001; 276: 15386-96.
43. Yu CJ, Yang PC, Shun CT, Lee YC, Kuo SH, Luh KT. Overexpression of MUC5 genes is associated with early post-operative metastasis in non-small-cell lung cancer. *Int J Cancer* 1996; 69: 457-65.
- 15 44. Sonora C, Mazal D, Berois N, et al. Immunohistochemical analysis of MUC5B apomucin expression in breast cancer and non-malignant breast tissues. *J Histochem Cytochem* 2006; 54: 289-99.
45. Tang FQ, Duan CJ, Huang DM, et al. HSP70 and mucin 5B: novel protein targets of N,N'-dinitrosopiperazine-induced nasopharyngeal tumorigenesis. *Cancer Sci* 2008.
- 20 46. Ahn EK, Kim WJ, Kwon JA, et al. Variants of MUC5B minisatellites and the susceptibility of bladder cancer. *DNA Cell Biol* 2009; 28: 169-76.
47. Giuntoli RL, 2nd, Rodriguez GC, Whitaker RS, Dodge R, Voynow JA. Mucin gene expression in ovarian cancers. *Cancer Res* 1998; 58: 5546-50.
- 25 48. Sylvester PA, Myerscough N, Warren BF, et al. Differential expression of the chromosome 11 mucin genes in colorectal cancer. *J Pathol* 2001; 195: 327-35.
49. Pantano F, Baldi A, Santini D, et al. MUC2 but not MUC5 expression correlates with prognosis in radically resected pancreatic cancer patients. *Cancer Biol Ther* 2009; 8: 996-9.
50. Rousseau K, Wickstrom C, Whitehouse DB, Carlstedt I, Swallow DM. New monoclonal antibodies to non-glycosylated domains of the secreted mucins MUC5B and MUC7. *Hybrid Hybridomics* 2003; 22: 293-9.
- 30 51. Buisine MP, Devisme L, Maunoury V, et al. Developmental mucin gene expression in the gastroduodenal tract and accessory digestive glands. I. Stomach. A relationship to gastric carcinoma. *J Histochem Cytochem* 2000; 48: 1657-66.
- 35

EXAMPLE 4

Proteomic Biomarkers for the Early Detection of Pancreatic Cancer –

An Analysis of Pancreatic Duct Fluids

As mentioned in the previous examples, pancreatic cancer has been a challenging oncology problem because of late detection and a lack of understanding of the biology of different pancreatic lesions. Increasing use of high resolution computerized tomography and magnetic resonance imaging in clinical practice has discovered the presence of pancreatic cysts in about 2% of adults in the U.S.². Not all pancreatic cancers arise from cysts, some may only come from a dilated duct or a small mass. However, there are currently no diagnostic indicators that are consistently reliable, obtainable, and conclusive for diagnosing and risk-stratifying of these precursor lesions³. Thus if biomarker diagnosis can be successful either from small amount of specimen or in the blood, there is a potential to improve the early management of pancreatic lesions. We previously performed proteomic analysis of a collection of pancreatic cyst fluids to shed

light on the protein compositions and on their biology. See Example I. In the cyst study, 30 biomarkers were discovered but only 16 well understood ones were discussed. That biomarker set is expanded to 34 in the current study because of increased confidence in some new candidates. Using a set of pancreatic duct fluids obtained at the time of surgery near pancreatic lesions, the biomarkers discovered can be correlated with pathology findings. By studying duct fluid near the lesion, we also addressed pancreatic cancer situations where a cyst is not observed but where a minute specimen can still be obtained from the pancreatic duct. We report herein the congruent biomarkers in duct fluids, in cyst fluids, and in fluids of pancreatic cancer when grafted in mice, indicating that the set of about 34 biomarker proteins we describe are useful for the understanding and staging of early pancreatic lesions in the future.

The following materials and methods are provided to facilitate the practice of Example 4.

15 **Sample Acquisition**

This study was approved by the Institutional Review Board of the Fox Chase Cancer Center. We prospectively aspirated pancreatic duct fluid at the time of surgical resection performed for various pancreatic pathologies for 28 patients. Briefly, the pancreas head was separated from the rest of the pancreas, exposing the main duct. A syringe with needle was used to aspire the fluid from the main duct and put on ice. Patient information was blinded and deidentified prior to performing proteomics on the duct fluids. Each duct fluid was photographed, diluted with three volumes of ice cold PBS, mixed, and centrifuged for 10 min at 13,000 x g at 4 °C to remove cells and any insoluble materials, photographed, and snap frozen in liquid nitrogen in aliquots and banked at -80 °C. Fluids from pancreatic cancer tissue grafted in scid mice⁵ was a generous gift of Dr. Repasky of the Roswell Park Cancer Center. These samples were analyzed identical to the ductal specimens.

GeLC-MS/MS comparison of ductal protein expression

15 µg of duct fluid proteins was analyzed by GeLC-MS/MS exactly as previously described in detail for cyst fluids in Example 1. Protein determination was performed using the BioRad protein assay⁶. About 15 µg of protein was reduced with Dithiothreitol (DTT) and alkylated with Iodoacetamide (IAA) as previously described⁶ and solubilized

with lithium dodecylsulfate (LDS solution, Novex, Invitrogen) at 70 °C for resolution by PAGE⁴.

The proteome for each sample was deduced exactly as described in Example 1. Protein relative abundance is expressed as the emPAI score for each protein⁷. This number, shown in all the tables in this report is $\{10^{(\text{number of observed peptides per protein} / \text{number of theoretical peptides per protein}) - 1}\}$, is roughly proportional to the abundance of a protein in a complex mixture.

Data Analysis

The proteome data was processed exactly as previously described in Example 1. The mass spectrometry “wiff” data files were used to search the SwissProt protein database release 54.1 using MASCOT 2.2 (Matrix Sciences, London, U.K.), analyzing the MS/MS sequencing spectra of the +2 and +3 ions. Fixed modification of carbamidomethylcysteine, variable oxidation of methionine, and one trypsin miss were allowed for protein identification in GeLC-MS/MS, but the latter two were disallowed for calculating the emPAI scores. Peptide mass tolerance was +/-150 ppm and fragment mass tolerance was 0.5 Da. False discovery rate due to coincidence in database was less than 3.5% for individual peptides as judged by hits at a decoy database containing randomized sequences in each entry. Data visualization was assisted by Scaffold 3.0 with Q+ software⁸. An automated spreadsheet handled data compilation and alignment to the cyst biomarkers to confirm the validity of the 34 biomarker proteins congruent between the two systems.

Clinical Information on the Duct Fluids

The patients were divided into three groups as shown in Figure 8 and presented in the order of increasing tumor stage. All the patients were Caucasian. Of the 28 patients, 3 patients had benign disease (pancreatitis or serous cystadenoma); 4 had ampullary adenocarcinoma; 2 had intraductal mucinous papillary neoplasm (IPMN); 17 patients had pancreatic ductal adenocarcinoma (PDAC); 2 had non-PDAC pancreatic carcinomas. Of the 17 patients with PDAC, 6 had associated IPMN. Four patients had radiation treatment (denoted by y in staging) prior to surgery. We noticed no obvious effect by the radiation treatment on the biomarkers observed in the ductal fluids.

Distribution of the 34 Biomarker Proteins in Duct Fluids

The biological nature and histopathologic features of pancreatic cysts are varied⁹,
¹⁰. Mucinous cystadenomas (MCA/MCN - mucinous cystic neoplasm) and intraductal
papillary mucinous neoplasms (IPMN) are two forms of mucinous cystic neoplasms with
5 malignant potential that may or may not be related to each other¹⁰⁻¹⁵. IPMN may be in
the main duct or branch duct where it may have a more indolent course^{2, 16}.
About 503 proteins were identified with confidence among the duct fluids in this study.
Approximately 137 proteins normally found in plasma were observed in different
quantities in all of the duct fluids. See Figure 9. About 30 pancreatic enzymes were
10 identified within the duct fluids. The rest of the proteins varied in abundance from being
comparable to albumin to being barely detectable by our mass spectrometer. The 34
ductal fluid proteins that appeared most relevant as biomarkers are shown in Figures 10A
and 10B.

Samples were grouped into 5 categories for presentation in Figures 9, and 10
15 based on pathology diagnosis and tumor staging. Group 1, duct fluid #25, 45, 47, 48, 23
were specimens thought to be ampullary cancer. The origin of these cancers are clinically
difficult to determine^{17, 18}. They in general have little or no biomarkers that we described
in Table 8, shown in Figures 10A and 10B. Group 2 include duct fluids #29, 46, 49, 28,
42, 27, and 30. Samples #29, 46, and 49 are benign lesions expressing no biomarkers
20 (Figure 10A). Sample #28 was thought to be an ampullary cancer but pathology
determined it to be a ductal adenocarcinoma. The location of this sample may have made
it difficult to contain biomarker proteins. Sample #42 is a mucinous cystic adenoma with
an abundance of biomarkers. The patient chose palliative care and died within about one
year. Sample #27 and 30 are IPMN with no evidence of cancer. Sample 30 had CEA of
25 about 10,000 units/mL.

Group 3, #31, 41, 32, 36, 43, and 39 are six ductal fluid samples that were IPMN
plus some form of ductal adenocarcinoma. Thus one would expect to see biomarkers
from both types of lesions (Figures 10A and 10B). However, the adenocarcinoma stages
were early and their signal may be weaker than the IPMN component.

30 Group 4, #33, 26, 24, 35, 22, 34, 40, 38, 44 are 9 samples of ductal
adenocarcinoma. These exhibited an abundance of the biomarker proteins described in
Figure 10B.

Group 5, #50 is an adenosquamous carcinoma with abundant biomarkers.

Relative abundance of pancreatic enzymes in the Duct Fluids

Although all samples in this study were collected at the pancreatic main duct and in general the levels of most of the pancreatic digestive enzymes were high, some samples had moderate levels and some had no enzyme detected by our mass

5 spectrometry. The reason for enzyme level drop off may be the limits of sensitivity of the mass, but a decrease or absence of enzymes in a duct fluid sample may also have biological and clinical explanations to be uncovered. For example, in our previous study of cyst fluids, the enzyme levels in mucinous cystic adenoma fluids were relatively low as well.

10

Proteome of Fluids from Pancreatic Cancer Tissue Grafted in Mice

About 300 unique human proteins and 208 unique mouse proteins were discovered in three samples of fluids from pancreatic cancer tissue grafted in mice. Most of these proteins were congruent with the biomarker proteins of the cyst fluid study and

15 with the current duct fluid study. See Table 10 below.

Table 10: Proteomic Biomarkers Produced by Mice Xenograft Fluid from Human Pancreatic Cancer

Swiss-Prot ID	Biomarkers in pancreatic duct fluids & cyst fluids of IPMN or adenocarcinoma	Mass (Da)	Xenograft Sample 1	Xenograft Sample 2	Xenograft Sample 3
MUC1_HUMAN	Mucin-1	122170	0.2	0.27	0.16
MUC5A_HUMAN	Mucin-5AC	135404	0.27	1.01	0.21
MUC5B_HUMAN	Mucin-5B	605803	0.08	0.16	0.21
MUC13_HUMAN	Mucin-13	55710	0.14	0.3	0.21
CEAM1_HUMAN	CEACAM1	57981	0.13	-	-
CEAM5_HUMAN	CEACAM 5	37499	0.33	0.46	0.21
CEAM6_HUMAN	CEACAM 6	77489	0.21	0.26	0.15
CEAM8_HUMAN	CEACAM 8	38415	0.1	-	-
S10A6_HUMAN	Protein S100-A6	10230	1.71	0.94	0.94
S10AB_HUMAN	Protein S100-A11	11847	0.78	3.23	1.37
LG3BP_HUMAN	Galectin-3-binding protein	66202	1.03	1.15	0.64
GELS_HUMAN	Gelsolin	86043	0.52	0.52	0.13
DMBT1_HUMAN	Deleted in malignant brain tumors 1 protein	268039	-	0.01	0.06
PIGR_HUMAN	Polymeric-immunoglobulin receptor	84429	0.41	0.9	0.24
EZRI_HUMAN	Ezrin	69484	1.42	1.83	0.52

ANXA2_HUMAN	Annexin A2	38808	6.61	9.04	3.38
ANXA4_HUMAN	Annexin A4	36088	1.98	1.98	0.64
ANXA1_HUMAN	Annexin A1	38918	0.44	0.58	0.74
ANXA3_HUMAN	Annexin A3	36524	0.34	0.34	0.48
ANXA5_HUMAN	Annexin A5	35971	1.22	1.7	0.82
ANXA6_HUMAN	Annexin A6	76168	0.1	0.1	-
LEG3_HUMAN	Galectin-3	26229	0.5	0.97	0.5
LEG4_HUMAN	Galectin-4	36032	0.64	1.44	0.1
ENOA_HUMAN	Alpha-enolase	47481	0.46	1.48	1.13

These proteins are representative of about 130 serum proteins observed. Serum albumin but not hemoglobin represents serum contamination. The numbers shown are emPAI roughly proportional to protein abundance. No proteins were detected for the empty boxes.

5 DISCUSSION

The sensitivity of pancreatic cyst fluid cytology has been reported as only 27-64% but it can add value to the diagnosis¹⁹. Several studies have suggested that a variety of tumor markers (e.g., CEA (carcinoembryonic antigen CEACAM5, a carcinoembryonic antigen-related cell adhesion molecule)^{20, 21}, CA 19-9 (carbohydrate antigen 19-9), CA 15-3 (cancer antigen MUC1, mucin 1)) may distinguish mucinous from non-mucinous cystic lesions, and also may predict whether a cyst harbors areas of malignant transformation^{9, 22, 23}. However, no marker by itself is sufficiently reliable. Recently there has been significant interest in pancreatic cyst fluids²⁴. For example, cyst fluid interleukin -1 beta was proposed as an indicator of the risk of carcinoma in IPMN²⁵ in addition to the traditional interest in mucin levels in these lesions²⁶. Incidentally, mucin 5B, a promising biomarker we discovered⁴, was not studied in the latter report. IPMN lesions can be benign or can progress to malignancy²⁷⁻²⁹. Distinguishing these two possibilities is not always possible. A recent study suggested that CEA levels is by itself a good correlation with the mucinous nature of the cyst, perhaps also for malignancy²¹.

20 The following observations can be made from this data:

(1) Homologs of biomarkers are important. Homologs of biomarkers in pancreatic cyst and duct fluids are present in an interchangeable manner, perhaps reflecting that they may be interchangeable in certain ways *in vivo*. The current duct fluid study has confirmed many of the biomarkers discovered in pancreatic cysts. Tables 8, 9A and 9B show that the four classes of biomarkers we previously reported⁴, namely, ductal enzymes including amylase, mucins, CEACAMs, and S100s, are also observed in the duct fluids and correlated with the lesions being mucinous or ductal adenocarcinoma.

Ductal Biomarkers

Elevated duct fluid amylase suggest communication with the pancreatic duct, as in the case of IPMNs or cancer, but cannot distinguish between these entities. Amylase by itself is not always a good indicator of ductal involvement, lipases and
5 carboxypeptidases are among the 30 enzymes that can each suggest ductal involvement. In this study, we observed that at least in the duct, which may not be as restricted in fluid flow as in the cyst, the level of pancreatic enzymes can vary from very high to very low to undetectable. Cyst fluids amylase levels can vary from non-detectable to 50,000 units/mL. Hence it is believable that the same may be true for other pancreatic enzymes.
10 The qTOF mass spectrometry used in this study is unable to deliver four logs of sensitivity of peptide detection in a complex mixture. Hence where enzymes are not observed, a more sensitive approach including multiple-reaction monitoring mass spectrometry³⁰⁻³³ or ELISA assay may be able to provide a more complete picture of the presence and distribution of these proteins.

15

Mucin Biomarkers

Abnormal expression of mucins and changes in the post-translational modification patterns of the amino acids or carbohydrates were recognized as potential biomarkers of malignancy^{9, 34-36}. Mucin 1 is MUC1, CA15-3, a known pancreatic cancer
20 marker³⁷. Mucin 1 and mucin5AC are found in the stomach as well as in the pancreas³⁸ and thus can potentially appear as contaminant at low levels in pancreatic duct fluids due to the FNA needle puncturing through the stomach wall. However, mucin 5B is found in the pancreas but not in the stomach³⁸. These two mucins are encoded by two different genes, MUC5AC (coding for both MUC5A and MUC5C) and MUC5B. Mucins 1, 5 and
25 AC, but not mucin 5B have been demonstrated in pancreatic lesions by immunohistochemistry³⁹. None of the 11 unique peptides identifying mucin 5AC with confidence in this report are found among the 33 unique peptides that identified mucin 5B. Importantly, this specificity of identification of soluble mucin 5B by mass
30 spectrometry overcomes the concerns of potential gastric contamination. One of the important purposes of the current duct fluid study is elimination of gastric contamination in our samples because no needle puncture through the stomach wall was involved in the collection of duct fluids during surgical resection.

Mucin 6 (MUC 6) (Table 8) is of great interest in pancreatic disease stratification. For example, it was shown that MUC6 expression was limited to the very early areas of

PanIN-1A and expression was lost in later stages⁴⁰. It was reported to be seen in advanced lesions more in the “cuboidal-cell” but not in the “columnar-cell” phenotype⁴⁰.

5 CEACAM Biomarkers

There are at least seven carcinoembryonic antigen homologs in humans^{41, 42}. CEACAM5 (CEA) is often elevated in several cancers⁴³. Pancreatic cyst fluid CEA levels of 192 ng/mL to 400 ng/mL appear to be specific for mucin-producing ductic neoplasms⁴⁴ most of the time⁹, but not always. Our data in duct fluids, as in cyst fluids⁴,
 10 demonstrated that CEACAM 5 is not the only homolog that can be elevated in pancreatic cancer. In fact, CEACAM 6 and sometimes CEACAM 8 may replace CEACAM 5 in its mechanism if expression levels were an indicator of their *in vivo* functions. There is little known about the functions of CEACAM homologs in pancreatic duct fluids and how they become solubilized. Importantly, CEACAM 5, 6, 7, 8 are only found in primates
 15 and not in rodants⁴¹, supporting our finding that these four CEACAMs may all be biomarkers of cancer, in human pancreatic duct fluids as well as in cyst fluids.

S100 Biomarkers

The S100 protein family are small Ca⁺⁺ binding proteins with many cellular
 20 functions⁴⁵⁻⁴⁸ including tumor promotion⁴⁹. S100A9⁵⁰, A6 and A11^{51, 52} were elevated in pancreatic carcinoma tissue. S100A6 was elevated in the ductal epithelium of pancreatic cancer obtained by laser capture microdissection⁵³. S100 homologs with roles in inflammation may influence the etiology of pancreatic ducts⁴⁹. This study extended S100 detection to pancreatic duct fluids. As previously discussed⁴, mass spectrometry
 25 has very high confidence of distinguishing these four S100s.

(2) Many biomarkers can substitute for each other for diagnosis and prognosis.

Interchangeable members of our first set of biomarkers can differentiate benign pancreatic disease and ampullary neoplasms from IPMN and PDAC in this study. This
 30 first set includes: Anterior gradient protein 2 homolog, Galectin-3-binding protein, Gelsolin, Neutrophil gelatinase-associated lipocalin, Leukocyte elastase inhibitor, Deleted in malignant brain tumors 1 protein, Polymeric-immunoglobulin receptor¹, Histone H4, Tetraspanin-1, Tetraspanin-8, Ezrin, Annexin A2, Annexin A4, Annexin A1, Annexin A3, Annexin A5, Galectin-3, Galectin-4, and Alpha-enolase. Some of these

biomarkers of more advanced pancreatic lesions were identified in our previous study⁴, but their significance is greatly assured by the current study in duct fluids because each of the samples in this larger study has pathology conclusion. Moreover, our new biomarkers: alpha enolase, polymeric immunoglobulin receptor, histone H4, and deleted in malignant brain tumor 1 protein, are rediscovered in the pancreatic cyst study data as a result of the duct biomarker study. That these set of biomarkers are congruent between duct and cyst fluids is shown in Tables 9A and 9B; (Figure 11), and many are further supported by their presence being demonstrated in the fluid of pancreatic cancer grafted into scid mice (Table 10, shown above).

Many of these biomarkers have interesting biology and relevance in disease. For example, alpha enolase has been used as a cancer indicator⁵⁴⁻⁵⁶ possibly related to inflammation. Anterior gradient protein 2 has been observed as elevated in pancreatic adenocarcinoma and proposed as a biomarker⁵⁷. Other biomarkers proposed for various cancers include annexins 3 and 4^{58,59}, tetraspanin 8⁶⁰, gelsolin⁶¹ and galactins⁶²⁻⁶⁴.

The second set of biomarkers has prognostic significance in PDAC and is variably expressed in PDAC. This set includes: Mucin-1, Mucin-2, Mucin-5AC, Mucin-5B, Mucin-6, Mucin-13, CEACAM 1, CEACAM 5, CEACAM 6, CEACAM 7, CEACAM 8, S100-A6, S100-A8, S100-A9, and S100-A11. This set of biomarkers was expressed in 65-100% of patients with IPMN, PDAC and other pancreatic carcinomas, compared with 0-25% of patients with benign pancreatic disease or ampullary adenocarcinoma. An example of the prognosis indication of S100s is discussed below.

Mucins distribution in ductal lesions

The current study makes use of the availability of detailed patient information and pathology results since each sample was obtained at the time of surgery. The distribution of the patients and disease classification with respect to mucin protein being detected is shown in Table 11. We conclude that in this data, non-mucinous and non-adenocarcinomatous tumors do not express Mucins compared to IPMN, PDAC or mucinous cystadenomas. We noted that one case of ampullary adenocarcinoma, duct 47, expressed MUC5AC; and on pathology was found to have signet cells with mucinous features.

TABLE 11. Summary of Mucins Expression in the Ductal Lesions

Pathology	n	MUC1	MUC5AC	MUC5B	Any MUC
Pancreatitis	1	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Ampullary Adenocarcinoma	4	0 (0%)	1 (25%)	0 (0%)	1 (25%)
Serous Cystadenoma	2	0 (0%)	0 (0%)	0 (0%)	0 (0%)
IPMN	2	0 (0%)	2 (100%)	1 (50%)	2 (100%)
IPMN+PDAC	6	2 (33%)	5 (83%)	3 (50%)	6 (100%)
PDAC	11	4 (36%)	6 (54%)	7 (65%)	7 (65%)
Mucinous cystadenoma	1	1 (100%)	1 (100%)	1 (100%)	1 (100%)
Adenosquamous	1	0 (0%)	1 (100%)	1 (100%)	1 (100%)

% in brackets denote the frequency of observations in the relevant cases.

Mucin 1 assists in the diagnosis of ductal lesions

We investigated the characteristics of IPMN and PDAC Patients (Table 12).

- 5 MUC1 expression appeared an early occurrence in the development of PDAC. No IPMN expressed MUC1 (Table 13). 4 patients, #32, 43, 41, and 39, preoperatively thought to have IPMN were found to have PDAC on final pathology (Table 8). Importantly, in this group of patients, 2/4 patients (50%) expressed MUC1, an information that if known in light of the observations in this report would have altered diagnosis.

10

TABLE 12. Characteristics of IPMN and PDAC Patients

	N	Neoadjuvant Chemorads	Reached 3 year survival?	Median CA19-9 at Dx	Median age
IPMN	2	0 (0%)	2 (100%)	16	66
PDAC					
T0N0	4	2 (50%)	4 (100%)	57	59
T1N0	2	1 (50%)	0 (0%)*	5	61
T3N0	2	0 (0%)	0 (0%)	446	65
TxN1	9	1 (11%)	0 (0%)+	347	75

*1 still alive

+1 still alive; 1 died of post-op complications

15

TABLE 13. Expression of Mucin 1 (MUC1) and S100A8/A9 in IPMN and PDAC cases.

Duct	Pre-op Dx	Path Dx	Stage	MUC1	S100A8/S100A9	DFS	OS
30	IPMN	IPMN	NA	-	+	NED	Alive
27	IPMN	IPMN	NA	-	-	NED	Alive
32	IPMN	IPMN+PDAC-	T0N0	-	-	NED	Alive

		is					
33	PDAC	PDAC	yT0N0	-	+	NED	Alive
36	Mucinous adenoCA	IPMN+PDAC- is	yT0N0	-	-	NED	Alive
28	PDAC	PDAC	T3N0	-	-	22.2	28.3
35	PDAC	PDAC	T3N0	-	-	17	21.7
43	IPMN	IPMN+PDAC	T2N1	-	-	NED	Alive
24	PDAC	PDAC	T2N1	-	-	17.3	21
34	PDAC	PDAC	yT3N1	-	+	2.3	9
44	PDAC	PDAC	T3N1	-	-	NED	Alive
23	Cholangio	PDAC	T3N1	-	-	NED	27
41	IPMN	IPMN+PDAC	T1N0	+	+	NED	Alive
39	IPMN	IPMN+PDAC	T3N1	+	+	1.7	16.7
26	PDAC	PDAC	yT1N0	+	+	5.8	22.3
22	PDAC	PDAC	T3N1	+	+	6.3	7.7
40	PDAC	PDAC	T3N1	+	+	11.1	12.6
38	PDAC	PDAC	T3N1	+	-	NED	2.7*

*Died of post op complications

5 **S100 A8 and S100 A8 are implicated in the prognosis of ductal lesions for advanced pancreatic cancer**

Protein S100A8, S100A9, S100A6 and S100A11 expression have all been associated with pancreatic ductal adenocarcinoma. Table 14 shows that 47-100% of patients with IPMN/PDAC had expression of both S100A8 and S100A9. Non-

10 IPMN/PDAC pathologies were less likely to express either S100A8 or S100A9, and in no instance was there expression of both S100A8 and S100A9.

TABLE 14. Summary of S100s Expression in the Ductal Lesions

Pathology	n	Protein S100A8	Protein S100A9	Protein S100A8 + S100A9	Protein S100A6	Protein S100A11
Pancreatitis	1	0 (0%)	1 (100%)	0 (0%)	0 (0%)	0 (0%)
Ampullary Adenocarcinoma	4	0 (0%)	1 (25%)	0 (0%)	0 (0%)	0 (0%)
Serous Cystadenoma	2	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
IPMN	2	2 (100%)	2 (100%)	2 (100%)	0 (0%)	0 (0%)
IPMN+PDAC	6	3 (50%)	3 (50%)	2 (33%)	1 (17%)	0 (0%)
PDAC	11	6 (55%)	7 (64%)	6 (55%)	3 (27%)	2 (18%)
Mucinous cystadenoCA	1	0 (0%)	0 (0%)	0 (0%)	1 (100%)	0 (0%)
Adenosquamous	1	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)

Legends are the same as in Table 11.

We investigated whether the expression of S100 A8 and A9 may have prognostic value in the management of PDAC patients with stage II and above given that S100 A8 and A9 over-expression is associated with poor pathological parameters in many invasive cancers. Table 10 shows that 7 out of 16 patients with PDAC (44%) had at least moderate or strong expression of either A8 or A9 or both. Patients with A8 or A9 expression had significantly worse median DFS and OS by Kaplan-Meier estimates (Table 15). This observation illustrates that the use of pancreatic lesion biomarkers may be more useful than whether cancer is present or the prognosis of an IPMN lesion.

TABLE 15. S100 A8 and S100 A9 are implicated in the prognosis of ductal lesions for stage II and above PDAC

Duct	Path Dx	Stage	MUC1	S100A8/ S100A9	DFS	OS	Median DFS	Median OS
28	PDAC	T3N0	-	-	22.2	28.3	28	28
35	PDAC	T3N0	-	-	17	21.7		
43	IPMN+PDAC	T2N1	-	-	NED	Alive		
24	PDAC	T2N1	-	-	17.3	21		
44	PDAC	T3N1	-	-	NED	Alive		
23	PDAC	T3N1	-	-	NED	27		
38	PDAC	T3N1	+	-	NED	2.7*		
34	PDAC	yT3N1	-	+	2.3	9	8	16
41	IPMN+PDAC	T1N0	+	+	NED	Alive		
39	IPMN+PDAC	T3N1	+	+	1.7	16.7		
26	PDAC	yT1N0	+	+	5.8	22.3		
22	PDAC	T3N1	+	+	6.3	7.7		
40	PDAC	T3N1	+	+	11.1	12.6		

*Patient died 2nd/post op complications – excluded from OS calculations

(3) Biomarkers of different stages of the disease may be present together when sampled at late stages of disease. The lesions studied in this report are relatively late lesions; meaning more than one feature may be present at variable magnitudes, constituting the set of 34 candidate proteins. The prognostic value of the S100s illustrated above supports the suggestion that different diagnostic and prognostic values of these biomarkers can be

elucidated with a larger study involving samples collected at different disease stages to better define stage specificity. In that light, cysts may be better earlier lesions where these biomarkers appear to be related with earlier lesion forms. However, we are mindful that many pancreatic cancers did not start from cysts. In those cases, sampling of a minute sample at the dilated duct by EUS-FNA should allow proteomic characterization. Interestingly, in many such specimens analyzed in this report, the same cyst-specific sets of biomarkers were discovered, illustrating common mechanisms of biomarker release in two apparently different lesion anatomies. Early lesions are more difficult to obtain but not impossible through collaborations. Moreover, even with late stage specimens, with the panel of biomarker proteins reduced to a manageable number, it is possible to use quantitative immunohistochemistry to investigate the relationship of each biomarker protein in specific tissue features and cells⁶⁵.

This study has illustrated that mass spectrometry tryptic peptide sequencing can provide comprehensive information on cancer biomarkers within pancreatic duct fluids and cyst fluids using minimal volumes of fluid, including small cyst size and increased viscosity, where current clinical assays cannot be easily performed.

In summary, proteomics of duct fluids revealed about 34 biomarker proteins whose presence or absence in a sample could be correlated with mucinous cyst adenoma and ductal adenocarcinoma. They are congruent with those discovered in pancreatic cyst fluids containing cystadenoma or adenocarcinoma, and fluids produced by pancreatic cancer engrafted in mice

We identified at least two 2 biomarker sets. One set included: Anterior gradient protein 2 homolog, Galectin-3-binding protein, Gelsolin, Neutrophil gelatinase-associated lipocalin, Leukocyte elastase inhibitor, Deleted in malignant brain tumors 1 protein, Polymeric-immunoglobulin receptor¹, Histone H4, Tetraspanin-1, Tetraspanin-8, Ezrin, Annexin A2, Annexin A4, Annexin A1, Annexin A3, Annexin A5, Galectin-3, Galectin-4, and Alpha-enolase. The second set included: Mucin-1, Mucin-2, Mucin-5AC, Mucin-5B, Mucin-6, Mucin-13, CEACAM 1, CEACAM 5, CEACAM 6, CEACAM 7, CEACAM 8, S100-A6, S100-A8, S100-A9, and S100-A11. They are expressed in 65-100% of patients with IPMN, PDAC and other pancreatic carcinomas vs. 0-25% of patients with benign disease or ampullary adenocarcinoma. Members of this set are variably expressed in PDAC; however, in stage II or higher PDAC, positive expression, for example, S100A8 and A9, was associated with a median disease-free survival of 8 months, vs. 28 months in patients without expression ($p=0.045$). Median

overall survival was 16 vs. 28 months with positive vs. negative expression ($p=0.033$).

Members of 34 pancreatic biomarker genetic signature provide suitable candidates to facilitate future pancreatic cancer diagnosis and cancer risk-stratification, whether cysts are involved or not. The small sample requirement can facilitate early
5 detection when the lesions are small and presumably less transformed.

REFERENCES FOR EXAMPLE 4

1. Asano M, Komiyama K. Polymeric immunoglobulin receptor. *Journal of oral science*
10 2011;53:147-56.
2. Yeo CJ, Sarr MG. Cystic and pseudocystic diseases of the pancreas. *Curr Probl Surg*
1994;31:165-243.
3. de Jong K, Bruno MJ, Fockens P. Epidemiology, diagnosis, and management of cystic lesions of
the pancreas. *Gastroenterology research and practice* 2012;2012:147465.
- 15 4. Ke E, Patel BB, Liu T, et al. Proteomic analyses of pancreatic cyst fluids. *Pancreas* 2009;38:e33-
42.
5. Hylander BL, Pitoniak R, Penetrante RB, et al. The anti-tumor effect of Apo2L/TRAIL on patient
pancreatic adenocarcinomas grown as xenografts in SCID mice. *J Transl Med* 2005;3:22.
6. Li XM, Patel BB, Blagoi EL, et al. Analyzing alkaline proteins in human colon crypt proteome. *J*
20 *Proteome Res* 2004;3:821-33.
7. Ishihama Y, Oda Y, Tabata T, et al. Exponentially modified protein abundance index (emPAI) for
estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol*
Cell Proteomics 2005;4:1265-72.
8. Searle BC. Scaffold: a bioinformatic tool for validating MS/MS-based proteomic studies.
25 *Proteomics* 2010;10:1265-9.
9. Brugge WR, Lewandrowski K, Lee-Lewandrowski E, et al. Diagnosis of pancreatic cystic
neoplasms: a report of the cooperative pancreatic cyst study. *Gastroenterology* 2004;126:1330-6.
10. Brugge WR, Lauwers GY, Sahani D, Fernandez-del Castillo C, Warshaw AL. Cystic neoplasms
of the pancreas. *N Engl J Med* 2004;351:1218-26.
- 30 11. Hamilton SR, Aaltonen LA. WHO Classification of Tumours, Pathology and Genetics of
Tumours of Digestive System. . Lyon, France: IARC Press, 2000.
12. Kloppel G, Solcia E, Longnecker D, Capella C, Sobin L. Histological typing of tumours of the
exocrine pancreas: World Health Organization international histological classification of tumours.
Springer-Verlag, 1998.
- 35 13. Wilentz RE, Albores-Saavedra J, Zahurak M, et al. Pathologic examination accurately predicts
prognosis in mucinous cystic neoplasms of the pancreas. *Am J Surg Pathol* 1999;23:1320-7.
14. Kloppel G. Clinicopathologic view of intraductal papillary-mucinous tumor of the pancreas.
Hepatogastroenterology 1998;45:1981-5.

15. Sarr MG, Carpenter HA, Prabhakar LP, et al. Clinical and pathologic correlation of 84 mucinous cystic neoplasms of the pancreas: can one reliably differentiate benign from malignant (or premalignant) neoplasms? *Ann Surg* 2000;231:205-12.
16. Yeo TP, Hruban RH, Leach SD, et al. Pancreatic cancer. *Curr Probl Cancer* 2002;26:176-275.
- 5 17. Ohike N, Kim GE, Tajiri T, et al. Intra-ampullary papillary-tubular neoplasm (IAPN): characterization of tumoral intraepithelial neoplasia occurring within the ampulla: a clinicopathologic analysis of 82 cases. *The American journal of surgical pathology* 2010;34:1731-48.
18. Lowe MC, Coban I, Adsay NV, et al. Important prognostic factors in adenocarcinoma of the ampulla of Vater. *The American surgeon* 2009;75:754-60; discussion 761.
- 10 19. Genevay M, Mino-Kenudson M, Yaeger K, et al. Cytology adds value to imaging studies for risk assessment of malignancy in pancreatic mucinous cysts. *Annals of surgery* 2011;254:977-83.
20. Shami VM, Sundaram V, Stelow EB, et al. The level of carcinoembryonic antigen and the presence of mucin as predictors of cystic pancreatic mucinous neoplasia. *Pancreas* 2007;34:466-9.
21. Cizginer S, Turner B, Bilge AR, et al. Cyst fluid carcinoembryonic antigen is an accurate
- 15 22. Lewandrowski KB, Southern JF, Pins MR, Compton CC, Warshaw AL. Cyst fluid analysis in the differential diagnosis of pancreatic cysts. A comparison of pseudocysts, serous cystadenomas, mucinous cystic neoplasms, and mucinous cystadenocarcinoma. *Ann Surg* 1993;217:41-7.
23. Levy M, Levy P, Hammel P, et al. [Diagnosis of cystadenomas and cystadenocarcinomas of the
- 20 24. Kwon RS, Simeone DM. The use of protein-based biomarkers for the diagnosis of cystic tumors of the pancreas. *International journal of proteomics* 2011;2011:413646.
25. Maker AV, Katabi N, Qin LX, et al. Cyst fluid interleukin-1beta (IL1beta) levels predict the risk of carcinoma in intraductal papillary mucinous neoplasms of the pancreas. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2011;17:1502-8.
- 25 26. Maker AV, Katabi N, Gonen M, et al. Pancreatic cyst fluid and serum mucin levels predict dysplasia in intraductal papillary mucinous neoplasms of the pancreas. *Annals of surgical oncology* 2011;18:199-206.
27. Yopp AC, Katabi N, Janakos M, et al. Invasive carcinoma arising in intraductal papillary
- 30 28. Mino-Kenudson M, Fernandez-del Castillo C, Baba Y, et al. Prognosis of invasive intraductal papillary mucinous neoplasm depends on histological and precursor epithelial subtypes. *Gut* 2011;60:1712-20.
- 35 29. Inui K, Yoshino J, Miyoshi H, Kobayashi T, Yamamoto S. Development of pancreatic ductal adenocarcinoma associated with intraductal papillary mucinous neoplasia. *ISRN gastroenterology* 2011;2011:940378.
30. DeSouza LV, Romaschin AD, Colgan TJ, Siu KW. Absolute quantification of potential cancer markers in clinical tissue homogenates using multiple reaction monitoring on a hybrid triple
- 40 31. quadrupole/linear ion trap tandem mass spectrometer. *Anal Chem* 2009;81:3462-70.

31. Kitteringham NR, Jenkins RE, Lane CS, Elliott VL, Park BK. Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics. *J Chromatogr B Analyt Technol Biomed Life Sci* 2009;877:1229-39.
32. Unwin RD, Griffiths JR, Whetton AD. A sensitive mass spectrometric method for hypothesis-driven detection of peptide post-translational modifications: multiple reaction monitoring-initiated detection and sequencing (MIDAS). *Nat Protoc* 2009;4:870-7.
33. James A, Jorgensen C. Basic design of MRM assays for peptide quantification. *Methods Mol Biol* 2010;658:167-85.
34. Hammel PR, Forgue-Lafitte ME, Levy P, et al. Detection of gastric mucins (M1 antigens) in cyst fluid for the diagnosis of cystic lesions of the pancreas. *Int J Cancer* 1997;74:286-90.
35. Adsay NV. Role of MUC genes and mucins in pancreatic neoplasia. *The American journal of gastroenterology* 2006;101:2330-2.
36. Haab BB, Porter A, Yue T, et al. Glycosylation variants of mucins and CEACAMs as candidate biomarkers for the diagnosis of pancreatic cystic neoplasms. *Annals of surgery* 2010;251:937-45.
37. Grote T, Logsdon CD. Progress on molecular markers of pancreatic cancer. *Curr Opin Gastroenterol* 2007;23:508-14.
38. Andrianifahanana M, Moniaux N, Batra SK. Regulation of mucin expression: mechanistic aspects and implications for cancer and inflammatory diseases. *Biochim Biophys Acta* 2006;1765:189-222.
39. Wang Y, Gao J, Li Z, et al. Diagnostic value of mucins (MUC1, MUC2 and MUC5AC) expression profile in endoscopic ultrasound-guided fine-needle aspiration specimens of the pancreas. *Int J Cancer* 2007;121:2716-22.
40. Basturk O, Khayyata S, Klimstra DS, et al. Preferential expression of MUC6 in oncocytic and pancreatobiliary types of intraductal papillary neoplasms highlights a pyloropancreatic pathway, distinct from the intestinal pathway, in pancreatic carcinogenesis. *The American journal of surgical pathology* 2010;34:364-70.
41. Kuespert K, Pils S, Hauck CR. CEACAMs: their role in physiology and pathophysiology. *Curr Opin Cell Biol* 2006;18:565-71.
42. Varnum SM, Covington CC, Woodbury RL, et al. Proteomic characterization of nipple aspirate fluid: identification of potential biomarkers of breast cancer. *Breast Cancer Res Treat* 2003;80:87-97.
43. Gold P, Freedman SO. Specific carcinoembryonic antigens of the human digestive system. *J Exp Med* 1965;122:467-81.
44. Hammel P, Voitot H, Vilgrain V, et al. Diagnostic value of CA 72-4 and carcinoembryonic antigen determination in the fluid of pancreatic cystic lesions. *Eur J Gastroenterol Hepatol* 1998;10:345-8.
45. Salama I, Malone PS, Mihaimeed F, Jones JL. A review of the S100 proteins in cancer. *Eur J Surg Oncol* 2007.
46. Roth J, Goebeler M, Sorg C. S100A8 and S100A9 in inflammatory diseases. *Lancet* 2001;357:1041.
47. Ryckman C, Vandal K, Rouleau P, Talbot M, Tessier PA. Proinflammatory activities of S100: proteins S100A8, S100A9, and S100A8/A9 induce neutrophil chemotaxis and adhesion. *J Immunol* 2003;170:3233-42.

48. Leach ST, Yang Z, Messina I, et al. Serum and mucosal S100 proteins, calprotectin (S100A8/S100A9) and S100A12, are elevated at diagnosis in children with inflammatory bowel disease. *Scand J Gastroenterol* 2007;1-11.
49. Dougan M, Dranoff G. Inciting inflammation: the RAGE about tumor promotion. *J Exp Med* 2008;205:267-70.
50. Lu Z, Hu L, Evers S, Chen J, Shen Y. Differential expression profiling of human pancreatic adenocarcinoma and healthy pancreatic tissue. *Proteomics* 2004;4:3975-88.
51. Ohuchida K, Mizumoto K, Ishikawa N, et al. The role of S100A6 in pancreatic cancer development and its clinical implication as a diagnostic marker and therapeutic target. *Clin Cancer Res* 2005;11:7785-93.
52. Ohuchida K, Mizumoto K, Ohhashi S, et al. S100A11, a putative tumor suppressor gene, is overexpressed in pancreatic carcinogenesis. *Clin Cancer Res* 2006;12:5417-22.
53. Shekouh AR, Thompson CC, Prime W, et al. Application of laser capture microdissection combined with two-dimensional electrophoresis for the discovery of differentially regulated proteins in pancreatic ductal adenocarcinoma. *Proteomics* 2003;3:1988-2001.
54. Ferrigno D, Buccheri G, Giordano C. Neuron-specific enolase is an effective tumour marker in non-small cell lung cancer (NSCLC). *Lung Cancer* 2003;41:311-20.
55. Burghuber OC, Worofka B, Schernthaner G, et al. Serum neuron-specific enolase is a useful tumor marker for small cell lung cancer. *Cancer* 1990;65:1386-90.
56. Adewole IF, Newlands ES. Neuron-specific enolase (NSE) as a tumour marker and comparative evaluation with carcinoembryonic antigen (CEA) in small-cell lung cancer. *Med Oncol Tumor Pharmacother* 1987;4:11-5.
57. Riener MO, Pilarsky C, Gerhardt J, et al. Prognostic significance of AGR2 in pancreatic ductal adenocarcinoma. *Histol Histopathol* 2009;24:1121-8.
58. Schostak M, Schwall GP, Poznanovic S, et al. Annexin A3 in urine: a highly specific noninvasive marker for prostate cancer early detection. *J Urol* 2009;181:343-53.
59. Lin LL, Chen CN, Lin WC, et al. Annexin A4: A novel molecular marker for gastric cancer with *Helicobacter pylori* infection using proteomics approach. *Proteomics Clin Appl* 2008;2:619-34.
60. Voss MA, Gordon N, Maloney S, et al. Tetraspanin CD151 is a novel prognostic marker in poor outcome endometrial cancer. *Br J Cancer*;104:1611-8.
61. Habeck M. Gelsolin: a new marker for breast cancer? *Mol Med Today* 1999;5:503.
62. Demers M, Rose AA, Grosset AA, et al. Overexpression of galectin-7, a myoepithelial cell marker, enhances spontaneous metastasis of breast cancer cells. *Am J Pathol*;176:3023-31.
63. Ahmed H, Cappello F, Rodolico V, Vasta GR. Evidence of heavy methylation in the galectin 3 promoter in early stages of prostate adenocarcinoma: development and validation of a methylated marker for early diagnosis of prostate cancer. *Transl Oncol* 2009;2:146-56.
64. Endo K, Kohnoe S, Tsujita E, et al. Galectin-3 expression is a potent prognostic marker in colorectal cancer. *Anticancer Res* 2005;25:3117-21.

65. Poultides GA, Reddy S, Cameron JL, et al. Histopathologic basis for the favorable survival after resection of intraductal papillary mucinous neoplasm-associated invasive adenocarcinoma of the pancreas. *Annals of surgery* 2010;251:470-6.

5

While certain of the preferred embodiments of the present invention have been described and specifically exemplified above, it is not intended that the invention be limited to such embodiments. Various modifications may be made thereto without departing from the scope and spirit of the present invention, as set forth in the following

10 claims.

What is claimed is:

1. A method of differentiating a pancreatic mucinous cyst from a non-mucinous cyst in a human test subject, said method comprising:

a) providing a pancreatic cyst fluid specimen from a human test subject;

5 b) analyzing said fluid specimen for the presence or absence of mucin 5B in said cyst fluid sample, said presence being indicative of a mucinous cyst.

2. The method of claim 1 wherein the presence of a mucinous cyst indicates the patient should be assessed over time for development of pancreatic cancer.

10

3. The method of claim 1, wherein said cyst fluid is analyzed using a Thermo Quantum Access TSQ triple-quad mass spectrometer.

4. The method of claim 1, wherein said cyst fluid is analyzed using a qTOF mass
15 spectrometer.

5. A method for predicting risk or classifying pancreatic cancer based on the presence or absence of pancreatic biomarkers in a fluid sample isolated from the pancreas of a subject or a tissue isolated therefrom in a patient in need thereof, said method
20 comprising:

a) providing a protein profile of the proteins present in said fluid sample;

b) determining whether one, two, three or four or more of at least two proteins selected from the group consisting of Mucin-5B, Mucin-5AC, Mucin-1, Gelsolin, Carcinoembryonic antigen-related cell adhesion molecule 5 precursor, Ezrin
25 Galectin-3-binding protein, Mucin-13, Leukocyte elastase inhibitor, Annexin A1, Annexin A2, Carcinoembryonic antigen-related cell adhesion molecule 6, Annexin A3, Annexin A4, Galectin-4, Annexin A5, Phosducin, Tetraspanin-8, Galectin-3, Neutrophil gelatinase-associated lipocalin, Anterior gradient protein 2 homolog, Protein S100-A11, Protein S100-A6, Protein S100-A8, and Protein S100-A9.

30

Figure 1

Cyst #	15	6	8	10	1	2	4	16	18	13	7
Histology	-	-	-	-	-	-	-	-	-	-	-
CEA	-	-	-	-	<0.1	1.8	49.4	60.5	176	182	231
Amylase	-	-	-	-	-	56	-	-	-	163400	12330
Cytology	A	A	A	C	A	A	C	A	A	A	A
Gender	F	F	F	M	F	F	F	F	F	M	M
Age	33	53	83	50	51	74	73	56	76	67	63
Cyst size	15 mm x 12 mm	14 mm x 17 mm	22 mm x 14 mm	18 mm x 28 mm	16 mm x 13 mm	20 mm x 24 mm	31 mm x 11 mm	34 mm x 32 mm	22 mm x 14 mm	24 mm x 17 mm	23 mm x 10 mm
Location	pancreatic tail	pancreatic tail	pancreatic head	pancreatic body	pancreatic body	pancreatic tail	pancreatic neck	pancreatic body	pancreatic head	pancreatic head	pancreatic body

Cyst #	11	17	14	9	19A	19B	12	5	3	21
Histology	-	IPMN adenoma	MCA	-	-	IPMN adenoma	-	Adeno-carcinoma	-	MCA
CEA	333	410	582	797	2853	-	5115	14175	63,830	-
Amylase	43152	25600	4036	75210	-	-	-	6	-	-
Cytology	A	A	A	A	A	A	A	C	A	A
Gender	F	M	F	M	M	M	F	M	F	F
Age	81	77	45	68	81	81	65	70	73	32
Cyst size	27 mm x 10 mm	23 mm x 26 mm	46 mm x 47 mm	29 mm x 21 mm	22 mm x 20 mm	-	14 mm x 9 mm	18 mm x 18 mm	17 mm x 13 mm	-
Location	pancreatic tail	uncinate portion of pancreatic head	-	pancreatic head	pancreatic tail	-	pancreatic head	pancreatic body/tail	pancreatic head	-

Table I. **Clinical information for the pancreatic cyst fluid samples.** CEA (ng/mL) and Amylase (units/mL) are results from clinical lab immuno assays. MCA=mucinous cystadenoma. Cytology categories: A- Benign: No evidence of benign mucinous epithelium, atypical cells or carcinoma. C- Atypical/suspicious cytology.

Cyst #	15	6	8	10	1	2	4	16	18	13	7	11	17	14	9	19A	19B	12	5	3	21
Histology																					
CEA	-	-	-	-	-	<0.1	2	49	61	176	182	231	333	410	582	797	2853	-	14175	63830	-
Serum albumin	10.4	9.3	7.0	5.5	5.6	5.5		1.8		2.4		0.8	1.4	7.4	0.1	1.1	7.9	0.4	7.9	7.0	9.9
Alpha-1-antitrypsin	4.9	4.0	0.6	1.7	1.5	2.7		2.2	0.5	2.2	0.2	0.5	0.4	3.6	0.3		0.2	1.3	4.4	1.3	1.7
Serotransferrin	7.2	5.8	2.4	3.1	4.9	3.7		4.2		1.0		0.7		5.8		0.1	3.3	0.1	5.8	3.7	5.5
Apolipoprotein A-I	7.1	4.7	4.7	10.4	0.8	4.7							0.3	0.6			0.3		11.8	4.7	1.3
Transferritin	1.4	1.4	0.2	0.2	0.2	1.4		1.4						0.6					1.4	0.6	0.2
Apolipoprotein A-II	0.8																		2.3		0.4
Alpha-1-acid glycoprotein 1	1.4	1.8	0.8	1.4	0.6	1.4		0.8		0.8		0.6	0.2	1.4			0.4	0.2	2.3	0.8	0.6
Alpha-1-acid glycoprotein 2	1.1	1.1	0.6	0.8		1.4		0.4		0.8				1.1			0.2		1.4		0.4
Zinc-alpha-2-glycoprotein	0.7	1.1	0.2	1.3	0.1	1.1		0.9		0.7		0.1		0.7		0.1	0.4	0.1	1.3	0.2	0.2
Alpha-2-macroglobulin	0.3	0.1	0.6	0.5		0.9		0.3						0.1			0.4		0.4	0.1	0.7
Vitamin D-binding protein	2.1	1.5	0.7	0.5	1.4	0.4								0.3			0.4		1.4	0.1	0.9
Complement C3	0.4	0.6	0.4	0.7	0.1	0.9		0.1		0.0				0.5			0.4	0.1	1.2	0.3	0.6
Hemoglobin subunit alpha		0.3	2.1	1.0		1.0		0.6											2.1		1.0
Hemoglobin subunit beta		1.4	19.7	3.6		6.0		0.2											19.7		1.4
Haptoglobin alpha beta chain	1.4		2.0	1.4		0.2				0.2							0.7		2.5	0.7	
Ig gamma-1 chain C region	2.6	2.2	1.7	1.7	1.7	1.4		1.4		1.2		1.2	0.1	1.9			1.4		2.6	1.9	2.6
Ig gamma-2 chain C region	1.9	1.9	1.2	1.7	1.2	1.2		1.7		0.2		0.5	0.1	1.4			1.2	0.1	1.7	1.7	1.9

Table 2. **Representatives of 137 plasma proteins distributed among the pancreatic cyst samples.** The numbers shown are emPAI scores which are roughly proportional to protein abundance. Dark (>1), medium (0.1 to 1), and light (<0.1) shading denotes relative protein abundance. No proteins were detected for the empty boxes. CEA = ng/mL.

Figure 2A

-	Mass	SwissProt name	Title	Cyst 18	Cyst 7	Cyst 4	Cyst 13	Cyst 16	Cyst 11	Cyst 9	Cyst 12	Cyst 17	Cyst 3	Cyst 14	Cyst 19	Cyst 20
1	71317	ALBU_HUMAN	Serum albumin	-	0.05	0.05	3.83	2.38	1.04	-	0.11	0.58	17.1	16.2	1.49	15.3
2	30759	APOA1_HUMAN	Apolipoprotein A-I	-	-	-	-	-	-	-	-	-	4.69	0.42	-	-
3	16102	HBB_HUMAN	Hemoglobin subunit beta	-	-	-	-	0.24	-	-	-	-	-	-	-	-
4	79280	TRFE_HUMAN	Serotransferrin	-	-	-	1.38	3.32	0.58	-	0.05	-	8.83	11.4	-	2.94
5	46878	A1AT_HUMAN	Alpha-1-antitrypsin	0.47	0.08	0.36	5.31	2.98	0.59	0.26	1.51	0.26	1.33	4.42	-	0.17
6	15991	TTHY_HUMAN	Transferrin	-	-	-	-	1.39	-	-	-	-	0.24	0.92	-	-
7	11773	KAC_HUMAN	Ig kappa chain C region	-	-	-	0.79	3.28	-	-	-	-	3.28	4.73	-	2.2
8	11282	APOA2_HUMAN	Apolipoprotein A-II	-	-	-	-	-	-	-	-	-	-	0.35	-	-
9	11401	LAC_HUMAN	Ig lambda chain C regions	-	-	-	2.3	3.45	1.45	-	0.35	0.35	2.3	5	-	2.3
10	23725	A1AG1_HUMAN	Alpha-1-acid glycoprotein 1	-	-	-	1.44	1.11	0.56	-	0.16	-	0.35	1.84	-	0.35
11	23873	A1AG2_HUMAN	Alpha-1-acid glycoprotein 2	-	-	-	1.1	0.56	-	-	-	-	-	1.43	-	-
12	36596	IGHG1_HUMAN	Ig gamma-1 chain C region	-	-	-	1.41	1.66	1.19	-	0.1	-	1.93	2.23	-	1.41
13	36489	IGHG2_HUMAN	Ig gamma-2 chain C region	-	-	-	0.22	1.94	0.48	-	-	-	1.66	2.24	-	2.24
14	33222	IGHG3_HUMAN	Ig gamma-3 chain C region	-	-	-	0.38	-	-	-	-	-	0.71	1.12	-	1.93
15	34079	ZA2G_HUMAN	Zinc-alpha-2-glycoprotein	-	-	-	0.69	0.37	0.11	-	0.11	-	0.23	0.37	-	0.37
16	15305	HBA_HUMAN	Hemoglobin subunit alpha	-	-	-	-	0.57	-	-	-	-	-	-	-	-
17	55347	IC1_HUMAN	Plasma protease C1 inhibitor	-	-	-	0.07	0.07	-	-	0.3	-	0.3	0.58	-	0.07
18	164600	A2MG_HUMAN	Alpha-2-macroglobulin	-	-	-	-	0.14	-	-	-	-	0.02	0.09	-	0.3
19	11882	KV302_HUMAN	Ig kappa chain V-III region SIE	-	-	-	-	-	-	-	-	-	1.37	0.78	-	0.33
20	122983	CERU_HUMAN	Ceruloplasmin	-	-	-	-	-	-	-	-	-	0.03	-	-	0.03
21	12422	KV204_HUMAN	Ig kappa chain V-II region TEW	-	-	-	-	-	-	-	-	-	0.74	-	-	-
22	188569	CO3_HUMAN	Complement C3 [	-	-	-	0.02	0.04	-	-	0.06	-	0.04	0.29	-	0.24
23	40098	FETUA_HUMAN	Alpha-2-HS-glycoprotein	-	-	-	0.09	-	-	-	-	-	-	-	-	-

Figure 2B

[illegible]

Figure 2C

[illegible]

Figure 2D

[illegible]

Figure 2E

[illegible]

Figure 2F

[illegible]

Figure 2G

125	13059	HV313_HUMAN	Ig heavy chain V-III region POM	-	-	-	-	-	-	-	-	-	-	-	-	-
126	34454	FCN2_HUMAN	Ficolin-2	-	-	-	-	-	-	-	-	-	-	-	-	-
127	40624	RAI3_HUMAN	Retinoic acid-induced protein 3	-	-	-	-	-	-	-	-	-	-	-	-	-
128	61187	CH60_HUMAN	60 kDa heat shock protein, mitochondrial	-	-	-	-	-	-	-	-	-	-	-	-	-
129	58259	MMSA_HUMAN	Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial	-	-	-	-	-	-	-	-	-	-	-	-	-
130	69484	EZRI_HUMAN	Ezrin	-	-	-	-	-	-	-	-	-	-	-	-	-
131	596494	FCGBP_HUMAN	IgG Fc-binding protein	-	-	-	-	-	-	-	-	-	-	-	-	-
132	168870	MRC1_HUMAN	Macrophage mannose receptor 1	-	-	-	-	-	-	-	-	-	-	-	-	-
133	514737	APOA_HUMAN	Apolipoprotein(a)	-	-	-	-	-	-	-	-	-	-	-	-	-
134	59703	K1C10_HUMAN	Keratin, type I cytoskeletal 10	3.82	1.79	1.48	-	5.92	1.63	0.72	1.33	3.02	0.35	0.94	0.35	0.35
135	51875	K1C14_HUMAN	Keratin, type I cytoskeletal 14	0.52	0.52	0.63	-	0.74	0.87	0.23	0.32	0.52	0.23	-	0.15	0.15
136	44065	K1C19_HUMAN	Keratin, type I cytoskeletal 19	0.18	-	0.28	0.08	0.28	-	0.28	-	-	-	-	-	0.18
137	53671	K2C8_HUMAN	Keratin, type II cytoskeletal 8	0.22	0.07	-	-	0.14	0.31	-	0.07	0.22	-	-	0	-

Figure 2H

-	Mass	SwissProt name	Title	Cyst 5	Cyst 21	Cyst 8	Cyst 15	Cyst 6	Cyst 1	Cyst 10	Cyst 2
1	71317	ALBU_HUMAN	Serum albumin	18	14.5	19	18	16.2	18	13	16.2
2	30759	APOA1_HUMAN	Apolipoprotein A-I	21.9	0.79	3.51	8.05	4.69	0.42	21.9	11.8
3	16102	HBB_HUMAN	Hemoglobin subunit beta	12.4	1.38	15.7	-	-	-	2.66	9.81
4	79280	TRFE_HUMAN	Serotransferrin	4.68	3.95	3.12	5.82	6.47	7.19	3.95	6.82
5	46878	A1AT_HUMAN	Alpha-1-antitrypsin	4.85	2.16	0.59	2.98	4.02	1.71	2.16	5.31
6	15991	TTHY_HUMAN	Transferrin	3.59	0.24	0.24	2.69	2.69	0.24	0.24	3.59
7	11773	KAC_HUMAN	Ig kappa chain C region	3.28	1.39	4.73	4.73	6.66	2.2	4.73	3.28
8	11282	APOA2_HUMAN	Apolipoprotein A-II	10.2	0.35	-	0.83	-	-	-	2.34
9	11401	LAC_HUMAN	Ig lambda chain C regions	5	2.3	3.45	3.45	2.3	3.45	2.3	2.3
10	23725	A1AG1_HUMAN	Alpha-1-acid glycoprotein 1	2.29	0.81	0.81	1.11	2.29	0.56	1.11	1.84
11	23873	A1AG2_HUMAN	Alpha-1-acid glycoprotein 2	2.8	-	0.35	0.56	1.82	-	1.82	1.82
12	36596	IGHG1_HUMAN	Ig gamma-1 chain C region	2.56	1.66	1.66	1.66	1.66	1.66	2.23	1.66
13	36489	IGHG2_HUMAN	Ig gamma-2 chain C region	1.66	1.94	1.94	2.24	2.24	1.42	2.58	1.66
14	33222	IGHG3_HUMAN	Ig gamma-3 chain C region	1.12	1.93	1.37	1.12	1.63	0.71	1.12	1.63
15	34079	ZA2G_HUMAN	Zinc-alpha-2-glycoprotein	1.57	0.11	0.11	0.37	1.08	0.11	1.08	1.57
16	15305	HBA_HUMAN	Hemoglobin subunit alpha	2.9	1.48	2.9	-	0.25	-	1.48	0.97
17	55347	IC1_HUMAN	Plasma protease C1 inhibitor	0.92	0.3	0.22	0.38	0.38	0.07	0.8	0.92
18	164600	A2MG_HUMAN	Alpha-2-macroglobulin	0.39	0.49	0.78	0.17	0.05	-	0.42	0.82
19	11882	KV302_HUMAN	Ig kappa chain V-III region SIE	1.37	0.78	-	0.78	0.78	-	0.78	0.78
20	122983	CERU_HUMAN	Ceruloplasmin	0.56	0.34	0.09	0.47	0.13	0.34	0.38	0.75
21	12422	KV204_HUMAN	Ig kappa chain V-II region TEW	0.74	1.29	0.74	-	0.74	-	0.74	0.74
22	188569	CO3_HUMAN	Complement C3 [	0.68	0.44	0.19	0.24	0.29	0.06	0.56	0.68
23	40098	FETUA_HUMAN	Alpha-2-HS-glycoprotein	0.56	0.31	0.2	0.43	1.05	0.43	0.09	0.56

Figure 21

24	54526	VTDB_HUMAN	Vitamin D-binding protein	0.94	1.21	0.22	1.21	0.7	1.52	0.39	0.49
25	54809	A1BG_HUMAN	Alpha-1B-glycoprotein	1.06	0.14	0.14	0.48	0.48	0.39	0.39	0.48
26	20064	FRIL_HUMAN	Ferritin light chain	1.01	0.19	-	-	0.19	-	-	0.42
27	266034	FINC_HUMAN	Fibronectin	0.71	0.16	-	0.55	0.18	-	0.09	0.41
28	11895	KV106_HUMAN	Ig kappa chain V-I region EU	0.78	0.78	0.33	-	-	-	0.33	0.33
29	12099	KV101_HUMAN	Ig kappa chain V-I region AG	0.76	-	0.33	-	-	-	0.33	0.33
30	12042	LV302_HUMAN	Ig lambda chain V-III region LOI	0.33	0.33	-	0.33	0.33	0.33	0.33	0.33
31	52385	HEMO_HUMAN	Hemopexin	0.86	0.32	0.41	0.62	0.51	0.73	0.62	0.32
32	39886	AMBP_HUMAN	AMBP protein [Contains: Alpha-1-microglobulin]	0.2	0.09	0.09	0.31	0.09	-	0.31	0.31
33	12681	KV309_HUMAN	Ig kappa chain V-III region VG	0.31	0.31	0.31	0.31	0.31	-	0.31	0.31
34	45861	HPT_HUMAN	Haptoglobin [Contains: Haptoglobin alpha chain; Haptoglobin beta chain]	3.1	-	3.1	1.37	-	-	2.51	0.27
35	47792	AACT_HUMAN	Alpha-1-antichymotrypsin	1.12	0.16	0.16	0.25	0.25	0.57	0.69	0.25
36	52106	FIBG_HUMAN	Fibrinogen gamma chain	1.46	1.14	-	0.07	0.07	-	-	0.23
37	95656	FIBA_HUMAN	Fibrinogen alpha chain [Contains: Fibrinopeptide A]	0.35	0.3	-	-	0.04	-	-	0.21
38	38486	IGHA1_HUMAN	Ig alpha-1 chain C region	1.54	1.31	1.31	0.92	0.1	-	0.2	0.2
39	21383	FRIH_HUMAN	Ferritin heavy chain	0.39	-	-	-	-	-	-	0.18
40	56577	FIBB_HUMAN	Fibrinogen beta chain [Contains: Fibrinopeptide B]	1.6	1.02	-	-	-	-	-	0.14
41	60510	HRG_HUMAN	Histidine-rich glycoprotein	0.2	0.06	-	0.13	0.06	-	-	0.13
42	28909	CAH1_HUMAN	Carbonic anhydrase 1	0.28	0.28	0.64	-	-	-	0.28	0.13
43	42052	ACTB_HUMAN	Actin, cytoplasmic 1	0.41	0.67	0.82	-	-	0.29	0.09	0.09
44	143654	CFAH_HUMAN	Complement factor H	0.16	-	-	0.19	-	-	-	0.08
45	103489	ITI4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4	0.23	0.19	0.04	0.11	0.07	-	0.15	0.07

Figure 2J

46	194247	CO4A_HUMAN	Complement C4-A	0.12	0.04	-	-	0.02	-	-	0.02
47	53031	CLUS_HUMAN	Clusterin	0.07	0.23	-	-	0.07	-	0.61	0.23
48	40022	IDH3A_HUMAN	Isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial	0.09	0.09	-	-	0.09	-	-	0.09
49	23337	RETBP_HUMAN	Plasma retinol-binding protein	0.83	-	-	-	0.16	0.58	0.83	0.58
50	46484	PEDF_HUMAN	Pigment epithelium-derived factor	0.08	0.08	-	-	0.26	0.86	-	0.59
51	13671	HV310_HUMAN	Ig heavy chain V-III region HIL	0.29	0.29	-	-	0.29	-	0.29	-
52	12836	HV320_HUMAN	Ig heavy chain V-III region GAL	0.71	-	-	-	0.31	0.31	-	0.31
53	38747	LUM_HUMAN	Lumican	1.52	-	-	-	0.32	0.32	-	0.2
54	12462	HV304_HUMAN	Ig heavy chain V-III region TIL	0.32	0.74	0.32	-	0.32	-	0.32	-
55	12003	LV102_HUMAN	Ig lambda chain V-I region HA	0.33	0.33	-	-	0.33	0.33	-	-
56	39584	APOH_HUMAN	Beta-2-glycoprotein 1	0.09	0.2	-	-	0.57	0.31	-	0.2
57	36431	IGHG4_HUMAN	Ig gamma-4 chain C region	1.42	0.99	0.8	0.99	-	-	0.99	-
58	53025	ANT3_HUMAN	Antithrombin-III	1.11	0.72	0.07	1.11	0.15	-	0.31	0.15
59	12746	KV402_HUMAN	Ig kappa chain V-IV region Len	0.71	0.71	-	-	1.24	1.24	-	0.71
60	66202	LG3BP_HUMAN	Galectin-3-binding protein	0.39	-	-	-	-	-	-	0.06
61	123407	SDCG1_HUMAN	Serologically defined colon cancer antigen 1	0.03	-	-	-	-	0.03	0.03	-
62	13332	HV305_HUMAN	Ig heavy chain V-III region BRO	0.3	0.68	0.3	-	-	0.3	-	-
63	11768	KV105_HUMAN	Ig kappa chain V-I region DEE	0.34	-	-	-	-	0.79	-	-
64	55069	VTNC_HUMAN	Vitronectin	0.07	-	-	-	0.22	-	-	0.07
65	50210	MUC_HUMAN	Ig mu chain C region	0.43	1.54	0.15	-	-	-	-	0.15
66	70963	AFAM_HUMAN	Afamin	0.58	0.23	-	-	0.05	-	-	0.05

Figure 2K

67	23840	TIMP1_HUMAN	Metalloproteinase inhibitor 1	0.16	0.16	-	-	-	-	-	-	-	-
68	165215	PZP_HUMAN	Pregnancy zone protein	0.07	0.05	0.12	-	-	-	-	0.09	-	-
69	72996	KNG1_HUMAN	Kininogen-1	0.28	0.05	-	0.1	0.1	0.1	0.1	0.22	-	-
70	54873	A2AP_HUMAN	Alpha-2-antiplasmin	0.07	-	-	0.07	-	-	-	-	-	-
71	101782	ITIH1_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H1	0.11	0.07	-	-	-	-	-	0.24	-	-
72	16187	HGB1_HUMAN	Hemoglobin subunit gamma-1	0.24	-	-	-	-	-	-	-	-	-
73	516666	APOB_HUMAN	Apolipoprotein B-100	0.01	-	-	-	-	-	-	-	-	-
74	189923	CO5_HUMAN	Complement C5 [Contains: Complement C5 beta chain; Complement C5 alpha chain; C5a anaphylatoxin; Complement C5 alpha~ chain]	0.02	-	-	-	-	-	-	-	-	-
75	108367	CO6_HUMAN	Complement component C6	0.03	-	-	-	-	-	-	-	-	-
76	13114	HV103_HUMAN	Ig heavy chain V-I region V35	0.3	-	-	-	-	-	-	-	-	-
77	11500	LV301_HUMAN	Ig lambda chain V-III region SH	0.81	0.81	-	-	-	-	-	0.34	-	-
78	12863	KV310_HUMAN	Ig kappa chain V-III region VH	0.31	-	-	-	-	-	-	-	-	-
79	80014	TRFL_HUMAN	Lactotransferrin	0.15	0.44	-	-	-	-	-	-	-	-
80	42829	ILEU_HUMAN	Leukocyte elastase inhibitor	0.09	-	-	-	-	-	-	-	-	-
81	8560	UBIQ_HUMAN	Ubiquitin	-	-	-	-	-	-	-	-	-	-
82	15123	LEG7_HUMAN	Galectin-7	-	-	-	-	-	-	-	-	-	-
83	45037	CATD_HUMAN	Cathepsin D	-	-	-	-	-	-	-	-	-	-
84	27947	CASPE_HUMAN	Caspase-14	-	-	-	-	-	-	-	-	-	-
85	59947	CATA_HUMAN	Catalase	-	-	-	-	-	-	-	-	-	-

Figure 2L

86	37139	APRV1_HUMAN	Retroviral-like aspartic protease 1	-	-	-	-	-	-	-
87	76926	TGM3_HUMAN	Protein-glutamine gamma-glutamyltransferase E	-	-	-	-	-	-	-
88	334021	DESP_HUMAN	Desmoplakin	-	-	-	-	-	-	-
89	105392	OCRL_HUMAN	Inositol polyphosphate 5-phosphatase OCRL-1	-	-	-	-	-	-	-
90	82376	PLAK_HUMAN	Junction plakoglobin	-	-	-	-	-	-	-
91	112640	NFH_HUMAN	Neurofilament heavy polypeptide	-	-	-	-	-	-	-
92	115417	PA24B_HUMAN	Cytosolic phospholipase A2 beta	-	-	-	-	-	-	-
93	129242	RPC2_HUMAN	DNA-directed RNA polymerase III subunit RPC2	-	-	-	-	-	-	-
94	199928	RGPD5_HUMAN	RANBP2-like and GRIP domain-containing protein 5	-	-	-	-	-	-	-
95	43543	MUCB_HUMAN	Ig mu heavy chain disease protein	-	-	0.18	-	-	-	-
96	102013	B3AT_HUMAN	Band 3 anion transport protein	-	-	-	-	-	-	-
97	16389	HV209_HUMAN	Ig heavy chain V-II region ARH-77	-	0.24	-	-	-	-	-
98	15684	HBAZ_HUMAN	Hemoglobin subunit zeta	-	0.25	-	-	-	-	-
99	13486	KV401_HUMAN	Ig kappa chain V-IV region	-	-	-	-	-	-	-
100	16159	HBD_HUMAN	Hemoglobin subunit delta	-	-	5.92	-	-	0.54	-
101	16249	HBE_HUMAN	Hemoglobin subunit epsilon	-	-	0.24	-	-	0.24	-
102	11810	KV104_HUMAN	Ig kappa chain V-I region CAR	-	-	-	-	0.34	-	-
103	12931	KV110_HUMAN	Ig kappa chain V-I region HK102	-	-	-	-	0.31	-	-

Figure 2M

104	53406	ANGT_HUMAN	Angiotensinogen	-	-	-	-	-	0.07	0.14	-	-
105	38777	FHR1_HUMAN	Complement factor H-related protein 1	-	-	-	-	-	-	-	-	0.1
106	84583	CO2_HUMAN	Complement C2	-	-	-	-	-	-	-	-	0.04
107	45371	APOA4_HUMAN	Apolipoprotein A-IV	-	0.17	-	2.84	1.59	-	-	3.16	0.61
108	11889	KV116_HUMAN	Ig kappa chain V-I region Roy	-	0.78	-	0.78	-	-	-	-	-
109	12745	HV303_HUMAN	Ig heavy chain V-III region VH26	-	-	-	0.71	-	-	-	-	-
110	11192	PLGB_HUMAN	Plasminogen-related protein B	-	-	-	0.36	-	-	-	-	-
111	11715	KV119_HUMAN	Ig kappa chain V-I region Wes	-	-	-	0.34	-	-	-	-	-
112	11941	KV113_HUMAN	Ig kappa chain V-I region Lay	-	0.33	0.33	0.33	0	0.33	0.33	0.33	0.33
113	86847	CFAB_HUMAN	Complement factor B	-	-	-	0.13	0.18	-	-	-	-
114	68072	CFAI_HUMAN	Complement factor I	-	0.05	-	0.11	-	-	-	-	0.05
115	36246	APOE_HUMAN	Apolipoprotein E	-	-	-	0.1	-	-	0.22	0.1	-
116	39895	PON1_HUMAN	Serum paraoxonase/arylesterase 1	-	-	-	-	-	-	0.2	-	-
117	11624	LV403_HUMAN	Ig lambda chain V-IV region Hil	-	-	-	-	-	-	-	0.34	-
118	21547	APOD_HUMAN	Apolipoprotein D	-	-	-	-	-	-	-	0.18	-
119	86043	GELS_HUMAN	Gelsolin	-	-	-	0.04	-	-	-	-	-
120	96650	CO7_HUMAN	Complement component C7	-	-	-	0.3	-	-	-	-	0.04
121	12782	KV201_HUMAN	Ig kappa chain V-II region Cum	-	-	-	0.31	-	-	0.31	-	-
122	93247	PLMN_HUMAN	Plasminogen	-	-	-	0.48	0.21	0.12	0.08	-	-
123	16041	IGJ_HUMAN	Immunoglobulin J chain	-	0.54	0.24	-	-	-	-	-	0.24
124	11752	CO005_HUMAN	Uncharacterized protein C-15orf5	-	-	-	-	-	-	-	-	-

Figure 2N

125	13059	HV313_HUMAN	Ig heavy chain V-III region POM	-	-	-	-	-	-	-	-	-	-	-	-	-	-
126	34454	FCN2_HUMAN	Ficolin-2	-	-	-	-	-	-	-	-	-	-	-	-	-	-
127	40624	RAI3_HUMAN	Retinoic acid-induced protein 3	-	-	-	-	-	-	-	-	-	-	-	-	-	-
128	61187	CH60_HUMAN	60 kDa heat shock protein, mitochondrial	-	-	-	-	-	-	-	-	-	-	-	-	-	-
129	58259	MMSA_HUMAN	Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial	-	-	-	-	-	-	-	-	-	-	-	-	-	-
130	69484	EZRI_HUMAN	Ezrin	-	-	-	-	-	-	-	-	-	-	-	-	-	-
131	596494	FCGBP_HUMAN	IgG Fc-binding protein	-	-	-	-	-	-	-	-	-	-	-	-	-	-
132	168870	MRC1_HUMAN	Macrophage mannose receptor 1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
133	514737	APOA_HUMAN	Apolipoprotein(a)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
134	59703	K1C10_HUMAN	Keratin, type I cytoskeletal 10	1.19	4.43	0.83	0.2	1.19	0.94	0.72	2.78						
135	51875	K1C14_HUMAN	Keratin, type I cytoskeletal 14	0.42	0.87	0.15	0.15	0.15	0.23	0.32	0.32						
136	44065	K1C19_HUMAN	Keratin, type I cytoskeletal 19	0.77	-	0.18	0.18	-	0.28	0.28	0.18						
137	53671	K2C8_HUMAN	Keratin, type II cytoskeletal 8	0.31	0.22	0.07	-	0.07	0.4	-	0.07						

Figure 20

Cyst #	15	6	8	10	1	2	4	16	18	13	7	11	17	14	9	19A	19B	12	5	3	21
Histology													IPMN adenoma	MCA			IPMN adenoma		Adeno-carcinoma		MCA
CEA	-	-	-	-	<0.1	2	49	61	176	182	231	333	410	582	797	2853	-	5115	14175	63830	-
Alpha-amylase 2B							40	3.2	2.5	2.7	2.1	1.5	2.1		2.7	1.2		1.0			
Pancreatic alpha-amylase			0.2	0.2			4.0	3.1	2.7	2.7	2.2	1.5	2.2	0.7	2.7	1.4	0.3	1.1			
Pancreatic secretory granule membrane major glycoprotein GP2																					
Pancreatic triacylglycerol lipase							0.6	0.6	0.7	0.6	0.4	0.5	0.4	0.4	0.5	0.4		0.2			
Bile salt-activated lipase							0.3	0.5		0.7	1.6	1.8	0.6	0.7	1.5	2.5	0.2	0.2			
Phospholipase A2										0.3	0.5	1.0	0.1		0.5	0.8	0.2			0.6	
Colipase							0.2	0.9	0.2	0.5	0.5	0.9	0.5		0.2	0.9		0.2			
Carboxypeptidase A1									0.3	0.7	1.3					2.0					
Carboxypeptidase A2							4.0	0.5	3.0	3.0	3.3	3.3	0.7	0.4	3.6	1.7	0.5	0.8		0.6	
Carboxypeptidase B							1.7	0.8	0.3	1.2	1.3	1.2	0.1		1.7	1.0	0.1	0.2		0.7	
Chymotrypsinogen B2							5.6	1.7	4.3	5.1	5.6	5.1	1.1	0.5	6.7	2.3	0.6	1.7		1.0	
Chymotrypsinogen B							0.5	0.9	1.1	0.5	0.3	0.7	0.3		0.5	0.9		0.7			
Chymotrypsin-C									0.9		0.3			0.3						0.7	
Elastase-2A							0.6	0.8	1.6	0.8	0.8	0.8			0.8	0.8				0.3	
Elastase-2B							2.8	1.6	2.4	1.6	1.3	1.6	0.8	0.1	2.4	0.6	0.6	0.6			
Elastase-3A			0.1				0.6	0.4	0.4	0.4	0.6	0.6	0.3		0.6		0.1	0.4			
Elastase-3B							3.7	3.1	3.1	2.7	2.7	2.7	2.7	0.1	3.7	1.6				1.3	
Lithostathine 1 alpha							1.9	1.1	1.6	1.6	0.8	1.1	1.3		1.3	1.1				0.4	
Trypsin-1					0.1		0.4		2.0	0.4	0.4	1.1	0.2		0.7	1.1		1.5		1.5	
Trypsin-2							1.2	0.5	1.9	0.7	0.9	1.2	0.9	0.5	1.5	0.7	0.7	1.2		0.9	
Trypsin-3							0.5	0.7	1.2	0.7	0.7	1.2	0.3	0.7	0.9	0.5	0.5	0.3		1.5	
Kallikrein-1							0.2	0.2	0.5	0.2	0.2	0.4	0.2		0.5	0.7	0.2	0.1		0.4	
								0.1	0.1	0.3		0.1	0.3		0.3	0.1					

Table 3. Pancreatic enzyme proteins in the pancreatic cyst samples. Legends are the same as for Table 2.

Figure 3

Cyst #	15	6	8	10	1	2	4	16	18	13	7	11	17	14	9	19A	19B	12	5	3	21
Histology													IPMN adenoma	MCA			IPMN adenoma		Adeno- carcinoma		MCA
CEA	-	-	-	-	<0.1	2	49	61	176	182	231	333	410	582	797	2853	-	5115	14175	63830	-
Mucin-1													0.01	0.06		0.01	0.06	0.03		0.16	
Mucin-2																0.01	0.01	0.01			
Mucin-5AC											0.06	0.03	0.24		0.14	0.11	0.27	0.21		0.21	
Mucin-5B													0.02	0.21		0.01	0.01	0.02		0.02	
Mucin-6						0.07							0.01	0.03		0.06	0.14	0.04		0.10	
CEACAM1									0.06							0.06			0.06		
CEACAM5 (CEA)									0.05					0.05	0.05	0.15	0.15		0.10	0.26	
CEACAM6														0.10			0.10			0.33	
CEACAM7														0.13							
CEACAM8									0.10					0.10			0.10			0.10	0.10
Protein S100-A6			0.94														0.39		0.39		0.39
Protein S100-A8								0.37						1.56					5.54		2.50
Protein S100-A9								0.68						0.68					3.73		2.65
Protein S100-A11																			1.37		0.33

Table 4. Proteomic potential biomarkers for mucins, CEACAMs, and S100s found in the pancreatic cyst samples. Legends are the same as for Table 2.

Figure 4



Figure 5

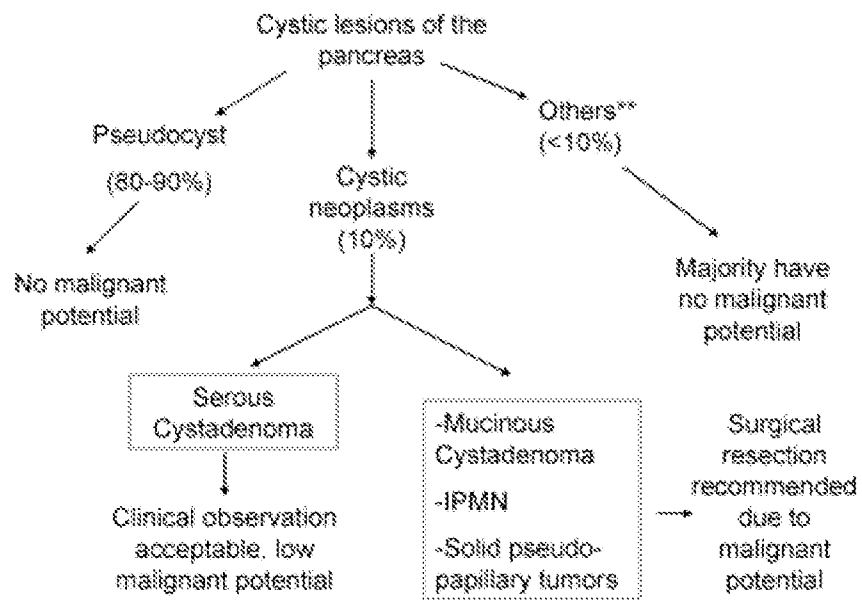


Figure 6

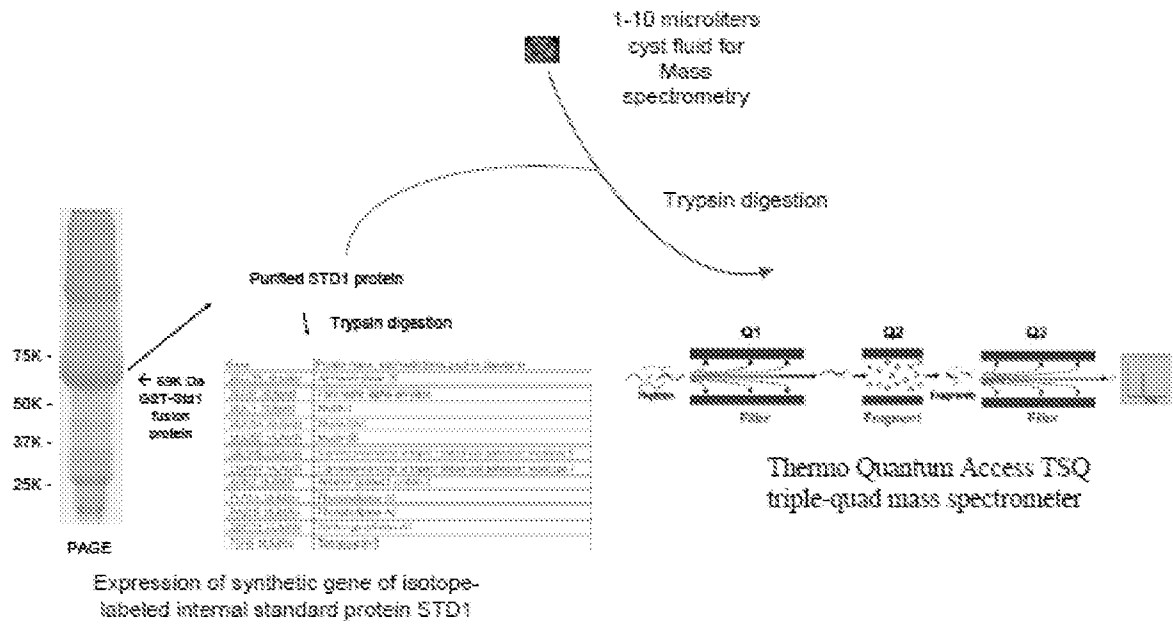


Figure 7

TABLE 6. Clinical Information for the Pancreatic Duct Fluid Samples

Duct fluid #	25	45	47	48	23	29	46	49	28	42
Pathology Report	Ampullary adeno CA	Ampullary adeno CA	Ampullary adeno CA	Ampullary adeno CA	Ductal adeno CA	Serous cystadenoma	Serous cystadenoma	Chronic pancreatitis	Ductal adeno CA	Mucinous cystic adeno
Stage of the Cancer	T2N0	T2N0	T3N0	T3N1	T3N1	NA	NA	NA	T3N0	T1N1M1
Pre-Operation Diagnosis	PDAC	Ampullary neoplasm	Ampullary neoplasm	IPMN	Cholangio-carcinoma	Serous cystadenoma	PNET	Pancreatic divisum	PDAC	Mucinous adeno CA
Surgery:	PPPD	PPPD	PPPD	PPPD	PPPD	Distal	Central	PPPD	Whipple	Palliative
Duct fluid #	27	30	31	41	32	36	43	39		
Pathology Report	IPMN	IPMN	IPMN + in situ adeno CA	IPMN + foci adeno CA	IPMN + in situ adeno CA	IPMN + in situ adeno CA	IPMN + ductal adeno CA	IPMN + ductal adeno CA		
Stage of the Cancer	NA	NA	T0N0	T1N0	TisN0	yTisN0	T2N1	T3N1		
Pre-Operation Diagnosis	IPMN	IPMN	IPMN	IPMN	IPMN	Mucinous adeno CA	IPMN	IPMN		
Surgery:	PPPD	PPPD	PPPD	Central	PPPD	Whipple	Whipple	Whipple		
Duct fluid #	33	26	24	35	22	34	40	38	44	50
Pathology Report	Ductal adeno CA	Ductal adeno CA	Ductal adeno CA	Ductal adeno CA	Ductal adeno CA	Ductal adeno CA	Ductal adeno CA	Ductal adeno CA	Ductal adeno CA	Adenosquamous CA
Stage of the Cancer	yT0N0	yT1N0	T2N1	T3N0	T3N1	yT3N1	T3N1	T3N1	T3N1	T3N0
Pre-Operation Diagnosis	PDAC	PDAC	PDAC	PDAC	PDAC	PDAC	PDAC	PDAC	PDAC	PDAC
Surgery:	PPPD	Whipple	Whipple	PPPD	Whipple	Whipple	PPPD	PPPD	PPPD	Total

MCA, mucinous cystadenoma; Sample 22 and 23 are cysts. "y" signifies pathologic staging in post-radiation patients. Patient 33 had no residual cancer at surgery.

Figure 8

TABLE 7. Blood Proteins Observed in the Pancreatic Duct Fluid Samples

Duct fluid #	25	45	47	48	23	29	46	49	28	42
Serum albumin	7.9		2.1	4.3	3.1	3.1	1.1	2.6	7.0	7.9
Alpha-1-antitrypsin	3.3	1.5	1.7	0.5	0.2	0.2	0.6	1.7	2.2	1.7
Serotransferrin	3.5		1.1	0.7	0.2	1.1	0.3	1.0	5.8	2.3
Apolipoprotein A-I	4.1			0.3		0.3			7.1	2.2
α -1-acid glycoprotein 1	1.4	1.8	1.4				0.2	1.4	0.6	0.2
α -1-acid glycoprotein 2	1.1							1.1	0.6	0.4
Hemoglobin subunit α	2.9			2.1	1.5	2.1	0.6		5.1	1.0
Hemoglobin subunit β	24.7		0.2	7.7	2.0	7.7	1.4	3.6	30.9	7.7
Hemopexin	0.73	-	-	-	-	-	-	-	0.73	0.51
Duct fluid #	27	30	31	41	32	36	43	39		
Serum albumin	5.9	7.9	8.8	7.9	7.9	5.5	4.3	7.9		
Alpha-1-antitrypsin	0.4	2.2	1.0	1.7	1.2	1.0	0.3	2.7		
Serotransferrin	1.4	5.8	3.1	5.2	4.0	2.0	0.4	5.5		
Apolipoprotein A-I	1.0	1.0	1.5	3.0	2.6	0.6		3.5		
α -1-acid glycoprotein 1		1.4	0.8	1.1	0.4	1.1		1.1		
α -1-acid glycoprotein 2		0.8						0.8		
Hemoglobin subunit α	2.9	0.6	1.0	2.1	3.9	2.9	2.1	2.9	2.9	0.6
Hemoglobin subunit β	24.7	2.0	4.7	9.8	19.7	30.9	2.7	19.7	24.7	2.0
Hemopexin	0.32	0.51	0.73	1.13	0.86	0.51	0.07	1.62	0.32	0.51
Duct fluid #	33	26	24	35	22	34	40	38	44	50
Serum albumin	7.9	8.8	5.5	7.9	6.6	5.9	9.3	6.6	6.2	7.9
Alpha-1-antitrypsin	2.7	5.8	1.0	1.9	1.7	2.7	2.4	1.3	1.2	1.9
Serotransferrin	4.0	6.8	2.4	4.0	2.4	0.8	3.5	1.9	3.5	3.1
Apolipoprotein A-I	0.6	6.2	0.1	3.5	5.4	1.5	1.0	1.0	0.6	4.1
α -1-acid glycoprotein 1		1.8	0.8	2.3	2.3	1.8	1.4	0.6	0.6	1.4
α -1-acid glycoprotein 2	0.4	0.8		0.6	0.6	0.6	0.8			
Hemoglobin subunit α	2.9	1.5	2.1	2.9	1.5	2.9	2.1	2.1	2.1	
Hemoglobin subunit β	30.9	9.8	24.7	15.7	4.7	9.8	15.7	3.6	15.7	0.5
Hemopexin	0.73	0.99	0.62	1.13	0.62	1.28	0.86	0.23	0.15	1.28

These proteins are representative of about 130 serum proteins observed. Serum albumin but not hemoglobin represents serum contamination. The numbers shown are emPAI roughly proportional to protein abundance. Noproteins were detected for the empty boxes.

Figure 9

Figure 10B: TABLE 8B. Proteomic Potential Biomarkers for Neoplasms in Pancreatic Ductal Fluids

Duct fluid #	32	36	43	39	33	26	24	35	22	34	40	38	44	50
Mucin-1				0.03		0.08	0.16		0.03		0.03	0.03		
Mucin-2														
Mucin-5AC		0.01	0.01	0.03		0.03			0.03	0.04	0.13	0.02	0.03	0.14
Mucin-5B	0.03			0.02	0.04	0.30			0.04	0.06	0.02	0.02	0.01	0.06
Mucin-6		0.12		0.03		0.22								0.10
Mucin-13														
CEACAM 1					0.06							0.06	0.06	
CEACAM 5				0.15					0.10		0.10			0.10
CEACAM 6						0.21								
CEACAM 7												0.13		
CEACAM 8														
Protein S100-A6	0.39				0.94	0.94	0.94							
Protein S100-A8	0.37			0.87	5.54	1.56			5.54	5.54	7.94			
Protein S100-A9				1.17	1.82	0.30	0.68		2.65	5.12	3.73			
Protein S100-A11						1.37	2.17							
Anterior gradient protein 2 homolog					0.42		0.42					0.19		
Galectin-3-binding protein				0.24		1.53	0.06			0.06		0.12	0.18	0.73
Gelsolin	0.09			0.04	0.04	0.46		0.29					0.04	0.13
Neutrophil gelatinase-associated lipocalin					0.17	2.46			2.46		1.54	0.59	0.17	4.52
Leukocyte elastase inhibitor					0.40		0.29			0.09	0.96			
Deleted in malignant brain tumors 1 protein	0.01			0.04	0.03	0.01	0.01		0.01	0.01	0.01	0.03	0.03	0.04
Polymeric-immunoglobulin receptor	0.35	0.35	0.09	0.35	0.47	0.75	0.04		0.82	0.29	0.47	1.07	0.75	0.99
Histone H4				0.83	1.47	2.34			1.47	1.47	1.47			0.83
Tetraspanin-1														
Tetraspanin-8					0.14	0.14						0.14		
Ezrin	0.30			0.05	0.23	0.05					0.11			
Annexin A2	0.20			0.10	0.91	3.80	0.74					0.10		
Annexin A4	1.21		1.21		3.89	3.01	1.00	0.22						
Annexin A1					0.32	0.10								
Annexin A3											0.48			
Annexin A5					0.49	1.01								
Galectin-3					0.31	0.31								
Galectin-4					0.35		0.35							
Alpha-enolase	0.70				3.22	0.16	0.35		0.16		0.16			

