

Proteomic Profiling of Pancreatic Cancer for Biomarker Discovery*

Ru Chen[‡], Sheng Pan[§], Teresa A. Brentnall[¶], and Ruedi Aebersold[§]

Pancreatic cancer is a uniformly lethal disease that is difficult to diagnose at early stage and even more difficult to cure. In recent years, there has been a substantial interest in applying proteomics technologies to identify protein biomarkers for early detection of cancer. Quantitative proteomic profiling of body fluids, tissues, or other biological samples to identify differentially expressed proteins represents a very promising approach for improving the outcome of this disease. Proteins associated with pancreatic cancer identified through proteomic profiling technologies could be useful as biomarkers for the early diagnosis, therapeutic targets, and disease response markers. In this article, we discuss recent progress and challenges for applying quantitative proteomics technologies for biomarker discovery in pancreatic cancer. *Molecular & Cellular Proteomics* 4:523–533, 2005.

Pancreatic cancer is the fourth leading cause of cancer death in the United States (1, 2). Most patients diagnosed with pancreatic cancer die within 12 months, and only 4% survive 5 years after diagnosis. Earlier diagnosis and better treatments are desperately needed to improve the survival rate of pancreatic cancer patients.

Pancreatic ductal adenocarcinomas, the majority of the exocrine pancreatic tumors, are thought to develop in a multi-step process, involving a series of specific genetic mutations in each step (3). In the past 10 years, significant progress in the identification and characterization of cancer-related gene abnormalities has been made; unfortunately, this progress has not yet effectively translated into substantial clinical improvement in the diagnosis or treatment of the disease. Current clinical pancreatic tumor markers lack the sensitivity and specificity required for screening an asymptomatic population for the purpose of early detection; an ideal biomarker would require a specificity of greater than 99% to avoid the consequences of a high rate of false-positive results. The only widely used marker for pancreatic cancer, CA19-9, is frequently elevated in pancreatic cancer but can also be ex-

pressed in other malignancies. Moreover, CA19-9 levels can be elevated in such benign conditions as acute and chronic pancreatitis, hepatitis, and biliary obstruction. The sensitivity and specificity of CA19-9 are ~80–90%, limiting its value as a screening marker for the general populace (4). Efforts to use various genetic mutations as surrogate markers of disease have also been unsuccessful due to the lack of sensitivity or specificity (5, 6). Therefore, there is an urgent need for new and better biomarkers for pancreatic cancer.

In the search for effective biomarkers, some fundamental and important aspects are worth careful consideration. First, what would be the ideal target for a biomarker? Theoretically, DNA, RNA, or protein can all be used as biomarkers. The choice of the type of target largely determines the technique applied and ultimately influences the outcome of biomarker development. Second, what biological specimens should be used in the search for the biomarkers? Several types of specimens are available for pancreatic cancer research, including pancreatic tissue, proximal body fluid such as pancreatic juice, and easily accessible body fluids such as serum or plasma. Third, should one use whole cancer tissue or isolated tumor cells for biomarker discovery? This is a concern common to all cancer researchers, but it is most manifest in pancreatic cancer where extensive fibroblastic stroma surrounds the cancer cells.

In recent years, the development of quantitative proteomics technologies has stimulated considerable interest in applying the technology for clinical applications. The capability to identify sensitive and effective biomarker proteins is critical in the battle against cancer, because the ability to treat and cure cancer, particularly pancreatic cancer, directly depends on the ability to detect it at its earliest stage (7–9). Recent proteomics studies in pancreatic cancer have identified proteins differentially regulated in cancer samples and have led to the discovery of several candidate biomarkers (12).¹ The use of proteomic profiling for pancreatic cancer biomarker discovery is still at its early stage; however, the efforts so far have been productive and the results are encouraging.

STRATEGIES FOR PROTEIN BIOMARKER DISCOVERY IN PANCREATIC CANCER

Discovery of cancer biomarkers for diagnostic purposes aims at identifying target molecules that are specific to or

From the [‡]GI Division/Department of Medicine, University of Washington, Seattle, WA 98195; [§]Institute for Systems Biology, Seattle, WA 98103; and [¶]Institute for Molecular Systems Biology, ETH Zurich and Faculty of Nature, Science Federal Institute, University of Zurich, CH-8093 Zurich, Switzerland

Received, January 14, 2005

Published, MCP Papers in Press, January 31, 2005, DOI 10.1074/mcp.R500004-MCP200

¹ R. Chen, E. C. Yi, S. Donohoe, S. Pan, D. A. Crispin, Z. Lane, D. R. Goodlett, M. P. Bronner, R. Aebersold, and T. A. Brentnall, manuscript in preparation.

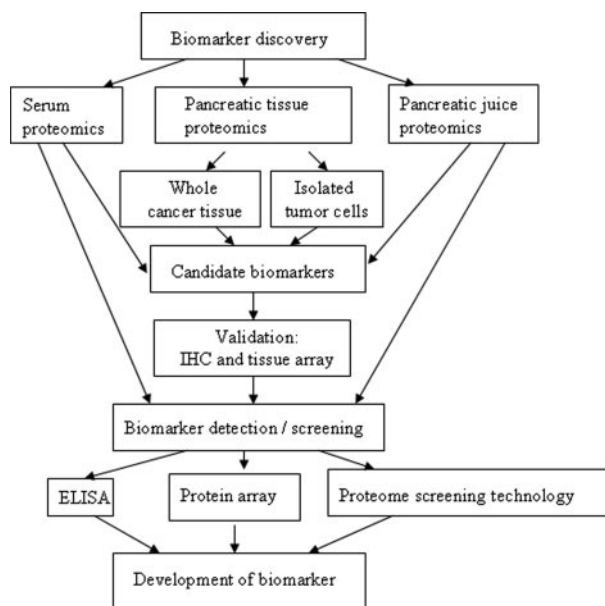


FIG. 1. **Flow chart for biomarker discovery and development using proteomic profiling.** For the discovery phase, protein profiling can be used to analyze pancreatic tissue, pancreatic juice, and serum to identify candidate biomarkers. Candidate biomarkers can be next validated by immunohistochemistry (IHC) in large cohort of patient samples, such as tissue array, and subsequently developed into diagnostic biomarkers in serum or pancreatic juice, using ELISA, protein array, and/or high-throughput proteome-screening technology.

associated with cancer, such as genetic alterations (DNA-based), differentially expressed genes (RNA-based), and differentially regulated proteins (protein-based). Previously characterized genetic alterations, including mutations and aberrant methylation, have been used to screen for DNA-based biomarkers. Currently there is not a good DNA-based biomarker used for detection of pancreatic cancer. Recent efforts using gene expression profiling have effectively identified differentially expressed genes in pancreatic cancers (14–18). RNA-based studies have reported overexpression of S100A4, prostate stem cell antigen, osteopontin, mesothelin, hTERT, and tissue inhibitor of metalloproteinase 1 (19–25). Some of these markers have diagnostic potential as serum markers or as tissue predictors of cancer. One RNA-based marker, the telomerase subunit (hTERT), is also expressed in inflammatory conditions (6, 26). The clinical application of these RNA-based markers has not yet been reported.

As it is evident that RNA levels do not necessarily correlate with protein levels (27, 28), a direct search for protein biomarkers may also be successful. Currently, protein-based biomarkers represent the majority of cancer biomarkers (6). The current approaches to search for pancreatic cancer protein biomarker using proteomic profiling are illustrated in Fig. 1. They include the investigation of differentially expressed proteins in pancreatic juice, pancreatic cancer tissue, and serum. Serum is an ideal diagnostic specimen in general, due to its

easy and inexpensive accessibility. Pathological and cancerous changes in the body can be reflected as protein changes in the serum.

Pancreatic juice is another good specimen for identifying biomarkers of pancreatic cancer. Juice is rich in proteins that are secreted directly from the pancreatic duct, where pancreatic cancers arise. Cancer cells are preferentially shed into the ductal lumen, making juice a rich source of cancer-specific proteins. While pancreatic juice may not be as easily accessible as serum, it is a more specific source for searching biomarker candidates associated with pancreatic cancer due to the proximity to the tumor. Biomarkers identified in pancreatic juice can be used to screen high-risk patients undergoing endoscopic procedures.

The most direct approach is to identify protein biomarkers in pancreatic cancer tissue. Biomarkers are presumably present in cancer tissue at higher concentration, and therefore have higher chance to be detected. However, tissue is not as clinically useful as blood or juice, because it is difficult to obtain tissue samples from the pancreas, which is in a remote location in the body. However, candidate biomarkers identified through proteomic profiling of tissue can provide a great basis for biomarker identification in serum for screening general population, for which a more sensitive and specific assay will be required. While different profiling techniques are warranted for each of these approaches to achieve an optimal outcome, a complementary strategy using different approaches and sample sources will be beneficial to facilitate the discovery of candidate biomarkers for pancreatic cancer.

CURRENT MS-BASED METHODS FOR PROTEOMIC PROFILING

One of the great challenges in analyzing the protein composition of biological samples is their enormous complexity. In biomarker discovery, much of the recent efforts in proteomics have been directed toward the development of strategies and platforms for quantitative protein profiling fitting for the needs of different types of biological samples. The traditional approach for quantitative protein profiling has been based on two-dimensional electrophoresis (2-DE)² for protein separation, followed by mass spectrometric identification of selected proteins or all proteins detected (29, 30). The two-dimensional separation method is based on isoelectric focusing and SDS-PAGE. One major advantage of the 2-DE approach is its ability to separate intact isomeric forms of protein that are useful in studies detailing posttranslational modifications (31). However, 2-DE does not work well for small proteins, very basic or acidic proteins, or for hydrophobic proteins, and is limited to a relatively low-throughput scale of research (27, 32). Since its introduction in 1975 (33, 34), 2-DE has been widely used and become one of the major techniques for

² The abbreviations used are: 2-DE, 2-dimensional electrophoresis; LCM, laser capture microdissection; HIP/PAP, hepatocarcinoma-intestine-pancreas/pancreatitis-associated protein.

protein separation and analysis, partly owing to recent technical advancements in this field, including: larger-format and higher resolving gels (35, 36), more sensitive staining methods (30, 37); and the coupling of 2-DE with MS/MS for protein sequence identification (38–40). Although the identification and accurate quantification of proteins separated by 2-DE is complex, and biased toward the abundant proteins (29, 41), the technique is robust and capable of detecting thousands of proteins in a single gel (30, 42).

The recent development of MS-based quantitative proteomics utilizes multi-dimensional chromatography for extensive separation of peptides and the incorporation of signature tags into proteins or peptides for quantitative analysis (32, 35, 43). In general, these quantitative approaches for protein profiling consist of the following major steps: 1) differentially labeling proteins/peptides from comparative samples with stable isotopes or chemical tags followed by proteolysis, 2) multi-dimensional chromatography separation, 3) ESI or MALDI MS/MS, and 4) computational analysis of the obtained spectra for the identification of peptide sequences and quantitative analysis. Relative quantification of each identified protein of the compared samples is accomplished by determining the abundance ratio from the signal intensities of the differentially labeled peptides with identical sequences. These newer approaches to quantitative proteomics significantly increase the analytical dynamic range and enhance the capability of detecting low-abundant proteins in a complex system. Quantitative proteomics has been applied to study cellular functions and pathways affected by perturbations and disease (44–49), to identify new components and changes in the composition of protein complexes and organelles (50–55), and has led to the detection of putative disease biomarkers (56).

While they differ in sample preparation and other aspects, most of the MS-based quantitative proteomic methods share the use of stable isotope labeling or mass tagging to generate the mass signatures that identify the target sample and serve as the basis for accurate quantification. These methods include the use of chemical reactions to introduce an isotopic or chemical tag at specific functional groups on polypeptides (57–60), metabolic isotope labeling using heavy amino acids in cell culture (61, 62), and methods that introduce stable isotope tags via enzymatic reactions (63, 64). Among those approaches, the most commonly used and versatile method has been the ICAT technology (59, 65), in which the proteins in two samples (*i.e.* normal and disease) representing different proteomes are labeled separately using one of the two chemically identical, but isotopically different, ICAT reagents. The labeled proteins from different proteomes are then combined, digested, purified, and separated by multi-dimensional chromatography, and analyzed by MS/MS. The incorporation of a biotin affinity tag into ICAT reagents enables the selective isolation and purification of cysteine-containing peptides, thus affording a substantial reduction in sample complexity. The presence of tagged cysteinyl residue(s) in the peptide

adds an additional powerful constraint for database searching. Because over 90% of mammalian species contain proteins with cysteinyl residue, the development of the ICAT strategy provides a widely applicable tool for comparative proteomics studies and significantly advances the technology for quantitative protein profiling—especially for human tissue samples that cannot be subjected to metabolic labeling.

Another approach for disease biomarker development is using proteomic pattern analysis. This approach primarily uses the pattern of signals observed within a mass spectrum to identify differentially abundant peaks within normal and disease samples for distinguishing the two groups (target and control). SELDI-TOF MS has been proposed in recent years for biomarker study, in combination with artificial intelligence algorithm, for the detection of discriminating signals (66). The technique uses MS to generate proteomic patterns in biological fluids, especially serum, and then applies pattern recognition algorithms to distinguish cancer patients from normal controls. In the first SELDI study, an artificial intelligence algorithm was used to identify a mass spectrometric pattern and to distinguish ovarian cancer in individuals from a second independent noncancer group with 95% specificity and 100% sensitivity (66). The study demonstrated that the technique was rapid and required minimal sample preparation and was suitable for analyzing serum and other body fluids. Because the technique did not provide quantitative information and protein identification, which can be at times critical for biomarker identification, the approach of using patterns of mass peaks to diagnose disease without knowing the identities of proteins has drawn some concern (67–69).

A comprehensive, quantitative proteomics analysis for biomarker discovery is challenging and involves considerations of different aspects. It deals with an enormously complex background that may fluctuate with the population tested and is limited by the technology available in dealing with sensitivity, dynamic range and throughput, and is further limited by the specificity and availability of different sample types. In the study of pancreatic cancer, 2-DE-based and ICAT-based profiling technologies have so far been successfully applied to pancreatic tissue and body fluid such as pancreatic juice. These studies have led to the discovery of some valuable or promising biomarker candidates. SELDI-TOF MS has mainly been applied to the analysis of body fluid, such as serum and pancreatic juice, and has led to discovery of candidate biomarkers (70). To effectively handle complex biological and clinical samples, integration of multiple platform technologies will be necessary. A complementary strategy involving multiple technologies and different sample types will be beneficial for the discovery of clinically useful pancreatic cancer biomarkers.

PROTEIN PROFILING OF PANCREATIC TISSUE

There have been considerable efforts to study changes at the mRNA level in pancreatic cancers; however, reports using

a proteomic approach to analyze the protein pattern in pancreas tissue samples are limited. There are a few recent studies using 2-DE to study protein expression profile in pancreatic cancer. One recently published study by Shekouh and colleagues used 2-DE to study differentially regulated proteins in pancreatic cancer (12). The authors compared the 2-DE protein profile from nonmalignant and malignant microdissected ductal epithelial cells and discovered nine protein spots that were consistently differentially regulated. Five of these nine protein spots showed increased abundance in cancer cells, while the other four protein spots showed diminished abundance. To identify these proteins, the authors matched the position of the spot of interest from microdissected samples to the pattern obtained from large quantities of whole tissue sample, and sequenced the corresponding spot from whole-tissue samples. The limited amount of material directly from the 2-DE gel of microdissected sample was not sufficient for mass spectrometric identification. Only one protein spot was identified; the other eight differentially regulated protein spots remain unidentified at this time. The identified protein, S100A6, belongs to S100 protein family. Several members of S100 family have been reported to be overexpressed in pancreatic cancer both by mRNA and immunohistochemical analysis, and thus may be important in pancreatic cancer. Another study by Shen *et al.* identified 40 differentially expressed proteins using 2-DE to analyze whole pancreatic cancer tissue (13). A considerably higher number of differentially expressed proteins was identified in the study because it used whole cancer tissue, thus larger amount of sample was available for 2-DE analysis. Five of the 40 proteins identified had been previously associated with pancreatic disease in gene expression studies. Among the other proteins, annexin A4, cyclophilin A, cathepsin D, galectin-1, 14-3-3 ζ , α -enolase, peroxiredoxin I, TM2, and S100A8 were specifically overexpressed in tumors compared with normal and pancreatitis tissues (13).

In an effort to systematically study protein profiles in pancreatic cancer with the aim of identifying potential biomarkers, we used the ICAT technology to perform quantitative protein profiling of pancreatic cancer tissues and normal pancreas. Proteins extracted from pancreatic cancer tissue and matching normal pancreas were labeled with isotopically heavy and light ICAT reagents, respectively. The samples were then combined, tryptic digested, purified, and analyzed with LC ESI MS/MS. The obtained CID spectra of peptides were searched against the National Cancer Institute human sequence database using SEQUEST for peptide/protein sequence identification and statistically modeled for subsequent statistical validation (38, 71, 72). Relative quantification was accomplished by comparing the signal intensity of the peptides with identical sequence, but different stable isotope signatures (73).

Using this ICAT strategy, 656 unique proteins were identified (protein probability scores ≥ 0.9) and quantified.¹ These

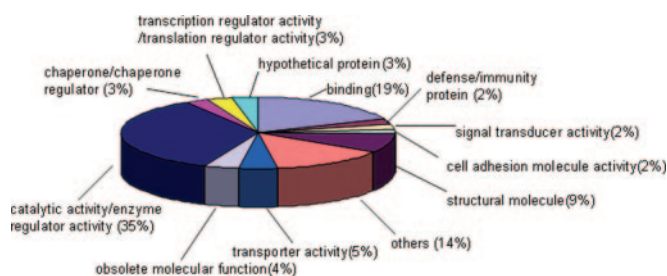


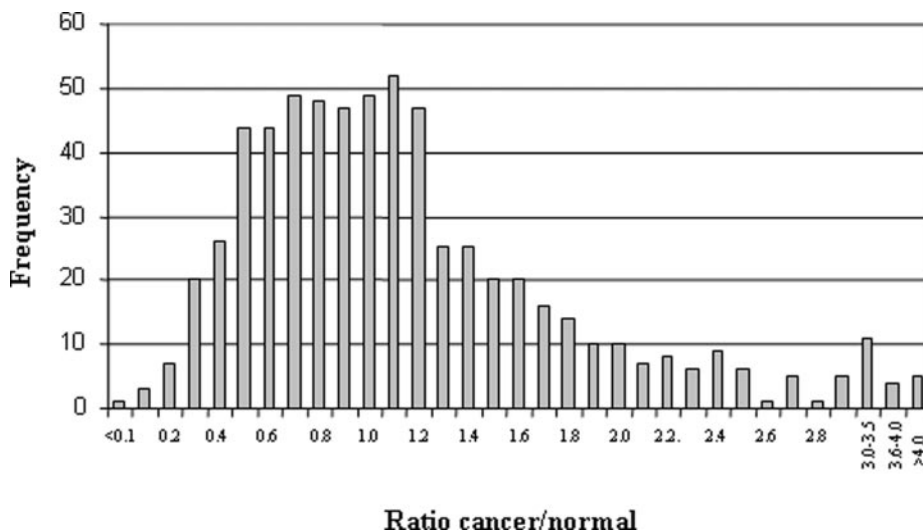
FIG. 2. **Protein categories identified in pancreatic cancer.** The 656 proteins identified and quantified in pancreatic cancer and normal pancreas were classified into many categories based on molecular function. The assignments were based on GO (Gene Ontology) Consortium.

proteins were classified into a variety of categories in molecular function (Fig. 2). Close to one-third of the proteins had catalytic activity or were regulated enzymes, which is not surprising due to the major digestion function of the pancreas. Some other significant molecular function groups included: binding proteins (19%), structural molecules (9%), transporter activity (5%), and obsolete molecular function (4%). The distribution of protein abundance ratios in pancreatic cancer for all these 656 proteins was centered around the ratio of 0.8–1.0 (Fig. 3), with 77% of the protein ratios being between 0.5 and 2.

There were 151 proteins for which the ratios were above 2.0 or below 0.5, representing at least 2-fold differential regulation in cancer samples. Over half of these proteins are active in metabolism and cellular physiological processes. In addition, many of these differentially regulated proteins are involved in signal transduction, cell growth, and/or maintenance. Moreover, many of the differentially expressed proteins play a role in the communication system through which epithelial and tumor cells interact with the extracellular matrix. These protein-driven interactions orchestrate tumor growth, migration, angiogenesis, invasion, metastasis, and immunologic escape. Several differentially regulated proteins identified by this ICAT approach were validated through Western blotting and immunohistochemical analysis on tissue array and might prove to be robust biomarker candidates.

There is currently a vast amount of information on the differential gene expression of pancreatic cancer at the mRNA level, including RNA expression arrays, DNA microarrays, differential display experiments, and SAGE analysis of pancreatic cancer (14, 25, 74–83). Comparing our ICAT proteomic data to the published gene expression studies, there were over 30 proteins/genes that have been detected as differentially regulated in pancreatic cancer by mRNA studies and the ICAT proteomic study. Significantly, 121 (80%) of the differentially regulated proteins identified by ICAT strategy in our study have not been detected by previous mRNA studies. The significant improvement in methods to identify differentially regulated proteins in cancer samples provides new targets and a greater pool for development of biomarker for early diagnosis and therapy.

FIG. 3. **Distribution of protein abundance ratios in pancreatic cancer.** Ratios were normalized to the median for the group. The ratios of a total of 656 proteins ratios were presented in this plot, ranging from 0.05 to 12.66. The distribution of the ratios was centered around the ratio of 0.8–1.0. The average ratio was 0.85. A number >1 means the protein was more abundant in cancer than normal.



Twenty-three out of the 40 differentially expressed proteins identified by 2-DE (13) were also identified in ICAT analysis; however, only 13 of them showed differential expression by at least 2-fold in cancer samples in our study. Some proteins (or genes), for example, annexin A1, glyceraldehyde 3-phosphate dehydrogenase, cofilin, galectin1, S100A protein family, and 14-3-3 protein σ have been repeatedly detected by various studies, including ours, to be differentially expressed in pancreatic cancer (10, 12, 15, 16, 18, 76, 82). Those proteins could be particularly important targets for future biomarker development.

CANCER SAMPLE PREPARATION

Pancreatic ductal adenocarcinoma consists of a solid mass with 30–90% of tumor cells interspersed with a fibroblastic stroma. What should be the starting material for protein profiling for the development of biomarkers of cancer? Should one start with purified cancer cells or with whole cancer tissue including the noncancerous components? Malignancy is a state that emerges from a tumor-host microenvironment (84, 85) in which the host participates in the induction, selection, and expansion of the neoplastic cells. Rather than being renegades, malignant tumor cells recruit vasculature and stroma through production and secretion of stimulatory growth factors and cytokines by both the tumor cells and the surrounding stroma (84, 86). The cross-talk between tumor cells and the surrounding stroma has previously been demonstrated as important in breast and melanoma tumor biology (84, 85). Our study, using ICAT profiling of pancreatic cancer of whole-tumor samples, verified this view. Many of the differentially regulated proteins underlying pancreatic tumorigenesis were involved in stroma-epithelial interactions.

On the other hand, it may still be of considerable interest to analyze pure tumor cells, provided these above factors were considered. The use of enriched cancer cells may facilitate the discovery of very-low-abundant biomarkers. Several approaches have been applied to isolate pure cancer cells from

pancreatic cancer tissue. The use of fixed tissue, such as that might be seen with microdissection of paraffin blocks, might not be suitable for proteomics; the proteins become cross-linked when fixed. Our strategy to isolate pure cancer cells used CELLection Epithelial Enrichment (DynaL Biotech, Oslo, Norway) to isolate epithelial cells. CELLection Epithelial Enrichment uses uniform, mono-dispersed, super-paramagnetic polymer spheres (Dynabeads) coated with monoclonal antibody BerEP4 (EpCAM Ab) against human epithelial cells. The beads bind to the target cells in a cellular slurry of whole pancreatic cancer. The bead:cell complexes are then separated from the sample with a magnetic particle concentrator. A DNA linker between the antibody and bead surface provides a cleavable site for cell detachment. Using a similar approach, another study has isolated pure epithelial cell from pancreatic tissue and demonstrated to be suitable for protein and mRNA studies (87).

A second method for isolating cells uses primary pancreatic ductal culture. This methodology is particularly useful for samples that are smaller in size and not amenable to the heavy cell losses associated with making the pancreatic cell slurry used for Dynabead separation. In our experience, the primary cultures can be passaged up to five times before senescing. Rigorous analysis of the cell cultures should be undertaken to ensure that the cells are epithelial in origin: cells are examined by hematoxylin and eosin, electron microscopy, organotypic culture with terminal differentiation, and tested for epithelial markers. This approach of short-term culturing cells, in effect, increases the quantity of small material. However, even short-term culture over three passages may result in unavoidable changes in protein expression (unpublished data).

A third method for isolating pure cancer cells is laser capture microdissection (LCM). LCM has been used to microdissect epithelial cells from cancers, effectively providing enriched populations of target cells. In a study using LCM to

enrich for both normal and malignant pancreatic ductal epithelial cells, investigators have managed to obtain enough material for 2-DE analysis; however, silver staining was required for detection and a number of technical challenges were encountered due to the limited number of cells captured from LCM (12). While it is effective in separating epithelial cells from other cell types, the LCM procedure requires a considerable commitment of time and is labor intensive.

Finally, an additional challenge in studying proteins from pancreas is the rich proteases produced in this organ. To prevent protein degradations, protease inhibitors are essential in sample preparation.

PROTEIN PROFILING OF PANCREATIC JUICE

Pancreatic juice was extensively studied in late 1970s and 1980s, primarily by early 2-DE analyses, which led to the discovery and description of several pancreatic enzymes (8, 88, 91). With recent progress in 2-DE and MS, several groups have started to employ a proteomics approach to study pancreatic juice. A detailed procedure for preparing pancreatic juice for 2-DE analysis was documented (92). In the study, the authors discovered that the addition of protease inhibitors resulted in well-resolved 2-D maps with a high reproducibility. Another proteomic approach for the analysis of pancreatic juice employed 1-DE and LC MS/MS (93). Pancreatic juice obtained from pancreatic cancer patients was first fractionated by 1-DE and subsequently analyzed by LC MS/MS for protein identification. A total of 170 unique proteins were identified, including several proteins that were previously found to be overexpressed in pancreatic cancer. The approach aimed at identifying the proteome constituent of pancreatic juice, and, thereby, was nonquantitative.

The SELDI-TOF approach has been applied to identify novel protein peaks in pancreatic juice to distinguish samples from cancer patients from normal controls (94). Using this methodology, Goggins and colleagues identified one peak present in the pancreatic juice from 10 of 15 of the patients with pancreatic cancer. This peak was further identified as hepatocarcinoma-intestine-pancreas/pancreatitis-associated protein (HIP/PAP) by a specialized protein chip and immunoassay. Because HIP/PAP is also present in pancreatitis and hepatoma cell lines, the limitations on sensitivity and specificity would preclude its use as a biomarker in the general population.

We have used the ICAT method to comprehensively study the proteome of pancreatic juice and quantitatively identify proteins that are differentially expressed in pancreatic cancer.¹ Three sets of comparison studies were performed: cancer to normal, pancreatitis to normal, and normal to normal. Because pancreatitis patients have the greatest risk for false-positive biomarker findings, we included juice from pancreatitis patients in our assessment to exclude any nonspecific crossover proteins. In the comparison of the cancer to normal pair, we identified and quantified a total of 78 proteins in

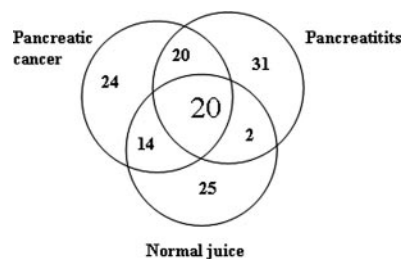


FIG. 4. Comparison of proteins identified in ICAT analysis of pancreatic juice from cancer sample, pancreatitis sample, and normal sample.

pancreatic juice of which 30 proteins displayed greater than 2-fold expression difference in cancer. These included 24 proteins that were up-regulated and six that were down-regulated in cancer sample. The pancreatitis and pooled normal pair analysis identified 73 proteins, of which 27 proteins showed greater than 2-fold expression difference in pancreatitis patients. Finally, the pair of normal to normal analysis identified 61 proteins. All together, in three studies we identified a total of 136 unique proteins from pancreatic juice with false-positive rate $\leq 0.9\%$ (11), including the samples from normal, pancreatitis, and cancer. Of the 136 proteins identified in pancreatic juice from normal, pancreatitis, and cancer, there were 20 proteins that were found in all three samples; 40 proteins identified in the pancreatitis samples were also identified in cancer samples; and 34 proteins identified in normal were also shared in cancer sample (Fig. 4). From the 30 proteins displaying greater than 2-fold expression difference in cancer, we subtracted the proteins that were also showing differential expression in pancreatitis and normal juice. Ultimately, we identified 21 proteins that were differentially expressed only in cancer juice.

It is important to note that in our study there was some nonspecific variability observed in protein profiles, even among the normal pancreatic juices. The variability from individual to individual can complicate the discovery of potential biomarkers. The ICAT study on a pair of normal to normal juice specimens indicated that 15 proteins displayed greater than 2-fold expression difference in two normal juices. This type of variability should be carefully considered for further biomarkers assessment.

PROTEIN PROFILING OF SERUM

Serum represents an ideal diagnostic specimen, due to its easy and inexpensive accessibility. Unfortunately, it is technically difficult to study low-abundance proteins in serum. One of the major difficulties associated with serum or plasma analysis is the fact that a small number of proteins such as albumin, $\alpha 2$ -macroglobulin, transferrin, and immunoglobulins, may represent as much as 80% of the total serum protein, while many other proteins are present in trace amounts. For example, important protein classes such as cytokines are at pg/ml level while serum albumin is at 35–55 mg/ml (95). The

large quantity of these abundant proteins makes it difficult to identify low-abundance proteins in serum using traditional 2-DE. Multi-dimensional chromatography and sample pre-fractionation prior to LC MS/MS have been used to reduce complexity and increase capacity for protein identification. However, the techniques remain problematic, mainly due to the great dynamic range in protein concentration. Lately, one study used ultra-high-efficiency strong cation exchange LC/reversed-phase LC/MS/MS separations to extend the dynamic range and coverage of human plasma proteins (96). The authors demonstrated that a protein identification dynamic range of $>10^8$ can be achieved using conventional ion trap MS/MS instrumentation. This approach resulted in the identification of >800 human plasma proteins from $\sim 365 \mu\text{g}$ of plasma without the need for depletion of high-abundant serum albumin or immunoglobulins. Even with this extensive effort, only a fraction of the total serum proteins have been identified.

Given the challenges in the direct proteomic analysis of serum with higher throughput, several alternative approaches have been developed to reduce the abundant proteins or the complexity of the sample. Different methods are now available to deplete albumin and other high-abundance proteins from serum, before 2-DE analysis, to allow the visualization of low-abundant proteins (97–101). One study used Affi-Gel Blue to treat serum samples before 2-DE analysis, and it resulted in the detection of 28 new spots, as well as enhanced staining intensity of different spots by several folds (97). While depletion of high-abundant proteins can increase the detection of low-abundant proteins, this approach is potentially problematic for quantitative measurements due to the variable and selective losses of other proteins along with the immunoglobulins and albumin. Careful experimental design and controls are needed to address these problems.

A third approach to reduce the complexity of serum proteins focuses on the in-depth analysis of sub-proteomes of rich biological context, thus minimizing the repeated analyses of abundantly expressed proteins. Recently, two new methods for the identification of *N*-linked glycopeptides in complex biological samples have been reported (102, 103). In general, these methods immobilize glycopeptides on a solid support, then the *N*-linked glycopeptides are released and collected for mass spectrometric analysis. By combining glycopeptide enrichment and stable isotope labeling, the method has been successfully applied to profiling glycoproteins in serum (103, 104).

Finally, the SELDI approach has been applied to profile serum in an attempt to identify protein peaks that can distinguish health from disease sample. A recent study from Goggins' group (70) analyzed serum samples from patients with and without pancreatic cancer using SELDI and identified two most discriminating protein peaks that could differentiate patients with pancreatic cancer from healthy controls with a sensitivity of 78% and specificity of 97%.

VALIDATION AND DEVELOPMENT OF BIOMARKER

The discovery of candidate biomarkers is the initial step toward the development of biomarkers for pancreatic cancer. It is crucial for a candidate biomarker to be validated at clinical level, using accessible biologic samples such as pancreatic juice or serum, in a larger cohorts of patients (Fig. 1).

Immunohistochemical analysis using tissue arrays is a widely accepted validation method for candidate biomarkers identified through tissue proteomic profiling. Tissue array technology allows the acceleration of studies by correlating protein expression with clinicopathologic information in large numbers of samples (105). Up to several hundred tissue sections can be processed in one slide for subsequent use. In our study of quantitative protein profiling of pancreatic cancer tissue, overexpression of annexin A2 was detected in both cancer samples by ICAT analysis. In the immunohistochemical analysis of a pancreas tissue array, annexin A2 was expressed negatively or mildly in all acinar cells, ductal cells, and islet cells of normal pancreas (12/12). However, 93% of the cancer samples (118/127) showed strong expression of annexin A2 in the cancer cells, while the majority of adjacent acinar cells, islet cells normal, and ductal cells from the same specimens showed mild or negative expression. Overexpression of annexin A2 has previously been observed in human pancreatic cancer and a pancreatic cancer-derived cell lines by mRNA measurement and immunohistochemical analysis (106, 107). However, previous efforts were limited to less than 10 samples, and no follow-up study has been performed. Interestingly, overexpression of annexin II (gene) has not been detected by recent mRNA profiling studies. We used tissue array analysis of over 100 pancreatic cancer samples to validate the overexpression of annexin A2 discovered by proteomic profiling. Our data detects overexpression in 93% of 127 pancreatic ductal adenocarcinoma samples, suggesting that annexin A2 is a good candidate for future biomarker development.

ELISA methodology is a tried and true method for detection of serum proteins and is useful for moderate numbers of sera samples. The ELISA test requires optimization prior to use and requires two antibodies for each protein that is being assayed. The two antibodies must be directed toward separate epitopes of the protein. This, in turn, makes the ELISA more difficult to construct for biomarker quantification. Alternatively, measurement of proteins in the body fluid can occur through a protein array, an emerging methodology. These arrays are constructed using antibodies that specifically recognize and capture proteins of interest. The antibody capture arrays can quantify sub-picomolar amounts protein abundance within body fluids or tissues of interest and can assay multiple targets simultaneously. However, this method, again, relies heavily on the availability of suitable antibodies.

Recently, a high-throughput proteome-screening technology (108) has been developed that provides an alternative

way for quantitative proteomics, specifically targeting certain proteins with biological significance such as biomarkers and prognostic proteins for identification and quantification. The highly selective platform consists of a LC/spotting system for peptide array preparation, a MALDI TOF/TOF tandem mass spectrometer, and uses synthesized stable isotope-labeled peptides as representative signatures for specific protein identification and quantification. The approach minimizes the interference of unwanted background proteins in a complex sample, therefore significantly improving the confidence in protein throughput and identification. Because the method is MS-based and directly focuses on selected peptides/proteins for identification and absolute quantification, it is applicable for profiling multiple proteins and well suited as a complementary tool for biomarker detection. This technology will be especially effective for detecting a panel of candidate biomarkers in large number of clinical samples.

The approaches for biomarker discovery in pancreatic cancer are summarized in Fig. 1. For discovery phase, protein profiling can be used to analyze pancreatic tissue, pancreatic juice, and serum to identify candidate biomarkers. Candidate biomarkers can be next validated by immunohistochemistry in large cohort of patient samples, such as in tissue arrays, and subsequently developed into diagnostic biomarkers in serum or pancreatic juice, using ELISA, protein arrays, and/or high-throughput proteome-screening technology.

CONCLUSION

MS-based proteomics has become an emerging and promising technology in understanding the process of tumorigenesis. Through the application of quantitative proteomics, such as the ICAT strategy, combined with other techniques such as tissue arrays, ELISA, protein arrays, and potentially high-throughput proteome-screening technology, we can now obtain informative and comprehensive information in identifying candidate biomarkers for pancreatic cancer. The differentially regulated proteins discovered in our studies and other studies brings to light a variety of cell functions and biological processes and can potentially reveal important information in understanding the progression of pancreatic cancer. The application of proteomics in studying pancreatic cancer is still in its early stage and remains challenging; however, as an emerging technology it has already provided fundamental information to improve the understanding of disease mechanism and potentially offer solutions for early detection of the cancer.

* This work was funded in part by the Gene and Mary Ann Walters Pancreatic Cancer Foundation and with Federal funds from the National Heart, Lung and Blood Institute, National Institutes of Health, under contract no. NOI-HV-28179. R. C. has been supported by GI training grant (NIH 5 T32 DK07742 08).

¶ To whom correspondence should be addressed: Department of Medicine, Division of Gastroenterology, University of Washington, Seattle, WA 98195. Tel.: 206-543-3280; Fax: 206-685-9478; E-mail: teribr@u.washington.edu.

REFERENCES

- Greenlee, R. T., Hill-Harmon, M. B., Murray, T., and Thun, M. (2001) Cancer statistics, 2001. *CA Cancer J. Clin.* **51**, 15–36
- Jemal, A., Thomas, A., Murray, T., and Thun, M. (2002) Cancer statistics, 2002. *CA Cancer J. Clin.* **52**, 23–47
- Hruban, R. H., Goggins, M., Parsons, J., and Kern, S. E. (2000) Progression model for pancreatic cancer. *Clin. Cancer Res.* **6**, 2969–2972
- Steinberg, W. (1990) The clinical utility of the CA 19–9 tumor-associated antigen. *Am. J. Gastroenterol.* **85**, 350–355
- Brand, R. (2001) The diagnosis of pancreatic cancer. *Cancer J.* **7**, 287–297
- Rosty, C., and Goggins, M. (2002) Early detection of pancreatic carcinoma. *Hematol. Oncol. Clin. North Am.* **16**, 37–52
- Hutter, G., and Sinha, P. (2001) Proteomics for studying cancer cells and the development of chemoresistance. *Proteomics* **1**, 1233–1248
- Lohr, M., and Faissner, R. (2004) Proteomics in pancreatic disease. *Pancreatol.* **4**, 67–75
- Wulfschlegel, J. D., Liotta, L. A., and Petricoin, E. F. (2003) Proteomic applications for the early detection of cancer. *Nat. Rev. Cancer* **3**, 267–275
- Deleted in proof
- Deleted in proof
- Shekhouh, A. R., Thompson, C. C., Prime, W., Campbell, F., Hamlett, J., Herrington, C. S., Lemoine, N. R., Crnogorac-Jurcevic, T., Buechler, M. W., Friess, H., Neoptolemos, J. P., Pennington, S. R., and Costello, E. (2003) Application of laser capture microdissection combined with two-dimensional electrophoresis for the discovery of differentially regulated proteins in pancreatic ductal adenocarcinoma. *Proteomics* **3**, 1988–2001
- Shen, J., Person, M. D., Zhu, J., Abbruzzese, J. L., and Li, D. (2004) Protein expression profiles in pancreatic adenocarcinoma compared with normal pancreatic tissue and tissue affected by pancreatitis as detected by two-dimensional gel electrophoresis and mass spectrometry. *Cancer Res.* **64**, 9018–9026
- Crnogorac-Jurcevic, T., Efthimiou, E., Nielsen, T., Loader, J., Terris, B., Stamp, G., Baron, A., Scarpa, A., and Lemoine, N. R. (2002) Expression profiling of microdissected pancreatic adenocarcinomas. *Oncogene* **21**, 4587–4594
- Han, H., Bearss, D. J., Browne, L. W., Calalupe, R., Nagle, R. B., and Von Hoff, D. D. (2002) Identification of differentially expressed genes in pancreatic cancer cells using cDNA microarray. *Cancer Res.* **62**, 2890–2896
- Iacobuzio-Donahue, C. A., Maitra, A., Shen-Ong, G. L., van Heek, T., Ashfaq, R., Meyer, R., Walter, K., Berg, K., Hollingsworth, M. A., Cameron, J. L., Yeo, C. J., Kern, S. E., Goggins, M., and Hruban, R. H. (2002) Discovery of novel tumor markers of pancreatic cancer using global gene expression technology. *Am. J. Pathol.* **160**, 1239–1249
- Iacobuzio-Donahue, C. A., Maitra, A., Olsen, M., Lowe, A. W., Van Heek, N. T., Rosty, C., Walter, K., Sato, N., Parker, A., Ashfaq, R., Jaffee, E., Ryu, B., Jones, J., Eshleman, J. R., Yeo, C. J., Cameron, J. L., Kern, S. E., Hruban, R. H., Brown, P. O., and Goggins, M. (2003) Exploration of global gene expression patterns in pancreatic adenocarcinoma using cDNA microarrays. *Am. J. Pathol.* **162**, 1151–1162
- Logsdon, C. D., Simeone, D. M., Binkley, C., Arumugam, T., Greenon, J. K., Giordano, T. J., Misek, D. E., and Hanash, S. (2003) Molecular profiling of pancreatic adenocarcinoma and chronic pancreatitis identifies multiple genes differentially regulated in pancreatic cancer. *Cancer Res.* **63**, 2649–2657
- Argani, P., Rosty, C., Reiter, R. E., Wilentz, R. E., Murugesan, S. R., Leach, S. D., Ryu, B., Skinner, H. G., Goggins, M., Jaffee, E. M., Yeo, C. J., Cameron, J. L., Kern, S. E., and Hruban, R. H. (2001) Discovery of new markers of cancer through serial analysis of gene expression: Prostate stem cell antigen is overexpressed in pancreatic adenocarcinoma. *Cancer Res.* **61**, 4320–4324
- Argani, P., Iacobuzio-Donahue, C., Ryu, B., Rosty, C., Goggins, M., Wilentz, R. E., Murugesan, S. R., Leach, S. D., Jaffee, E., Yeo, C. J., Cameron, J. L., Kern, S. E., and Hruban, R. H. (2001) Mesothelin is overexpressed in the vast majority of ductal adenocarcinomas of the pancreas: Identification of a new pancreatic cancer marker by serial analysis of gene expression (SAGE). *Clin. Cancer Res.* **7**, 3862–3868
- Koopmann, J., Fedarko, N. S., Jain, A., Maitra, A., Iacobuzio-Donahue, C., Rahman, A., Hruban, R. H., Yeo, C. J., and Goggins, M. (2004) Evaluation of osteopontin as biomarker for pancreatic adenocarcinoma. *Cancer Epidemiol. Biomarkers Prev.* **13**, 487–491
- Rosty, C., Ueki, T., Argani, P., Jansen, M., Yeo, C. J., Cameron, J. L.,

- Hruban, R. H., and Goggins, M. (2002) Overexpression of S100A4 in pancreatic ductal adenocarcinomas is associated with poor differentiation and DNA hypomethylation. *Am. J. Pathol.* **160**, 45–50
23. Ryu, B., Jones, J., Blades, N. J., Parmigiani, G., Hollingsworth, M. A., Hruban, R. H., and Kern, S. E. (2002) Relationships and differentially expressed genes among pancreatic cancers examined by large-scale serial analysis of gene expression. *Cancer Res.* **62**, 819–826
 24. Welsh, J. B., Sapinoso, L. M., Kern, S. G., Brown, D. A., Liu, T., Bauskin, A. R., Ward, R. L., Hawkins, N. J., Quinn, D. I., Russell, P. J., Sutherland, R. L., Breit, S. N., Moskaluk, C. A., Frierson, H. F., Jr., and Hampton, G. M. (2003) Large-scale delineation of secreted protein biomarkers overexpressed in cancer tissue and serum. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 3410–3415
 25. Zhou, W., Sokoll, L. J., Bruzek, D. J., Zhang, L., Velculescu, V. E., Goldin, S. B., Hruban, R. H., Kern, S. E., Hamilton, S. R., Chan, D. W., Vogelstein, B., and Kinzler, K. W. (1998) Identifying markers for pancreatic cancer by gene expression analysis. *Cancer Epidemiol. Biomarkers Prev.* **7**, 109–112
 26. Seki, K., Suda, T., Aoyagi, Y., Sugawara, S., Natsui, M., Motoyama, H., Shirai, Y., Sekine, T., Kawai, H., Mita, Y., Waguri, N., Kuroiwa, T., Igarashi, M., and Asakura, H. (2001) Diagnosis of pancreatic adenocarcinoma by detection of human telomerase reverse transcriptase messenger RNA in pancreatic juice with sample qualification. *Clin. Cancer Res.* **7**, 1976–1981
 27. Gygi, S. P., Rochon, Y., Franza, B. R., and Aebersold, R. (1999) Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* **19**, 1720–1730
 28. Tian, Q., Stepaniants, S. B., Mao, M., Weng, L., Feetham, M. C., Doyle, M. J., Yi, E. C., Dai, H., Thorsson, V., Eng, J., Goodlett, D., Berger, J. P., Gunter, B., Linseley, P. S., Stoughton, R. B., Aebersold, R., Collins, S. J., Hanlon, W. A., and Hood, L. E. (2004) Integrated genomic and proteomic analyses of gene expression in mammalian cells. *Mol. Cell Proteomics* **3**, 960–969
 29. Gygi, S. P., Corthals, G. L., Zhang, Y., Rochon, Y., and Aebersold, R. (2000) Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 9390–9395
 30. Rabilloud, T. (2002) Two-dimensional gel electrophoresis in proteomics: Old, old fashioned, but it still climbs up the mountains. *Proteomics* **2**, 3–10
 31. Marko-Varga, G., and Fehninger, T. E. (2004) Proteomics and disease—The challenges for technology and discovery. *J. Proteome Res.* **3**, 167–178
 32. Aebersold, R., and Goodlett, D. R. (2001) Mass spectrometry in proteomics. *Chem. Rev.* **101**, 269–295
 33. Klose, J. (1975) Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik* **26**, 231–243
 34. O'Farrell, P. H. (1975) High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **250**, 4007–4021
 35. Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422**, 198–207
 36. Gauss, C., Kalkum, M., Lowe, M., Lehrach, H., and Klose, J. (1999) Analysis of the mouse proteome. (I) Brain proteins: Separation by two-dimensional electrophoresis and identification by mass spectrometry and genetic variation. *Electrophoresis* **20**, 575–600
 37. Unlu, M., Morgan, M. E., and Minden, J. S. (1997) Difference gel electrophoresis: A single gel method for detecting changes in protein extracts. *Electrophoresis* **18**, 2071–2077
 38. Eng, J., McCormack, A. L., and Yates, J. R. (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5**, 976–989
 39. Mann, M., and Wilm, M. (1994) Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal. Chem.* **66**, 4390–4399
 40. Perkins, D. N., Pappin, D. J., Creasy, D. M., and Cottrell, J. S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567
 41. Corthals, G. L., Wasinger, V. C., Hochstrasser, D. F., and Sanchez, J. C. (2000) The dynamic range of protein expression: A challenge for proteomic research. *Electrophoresis* **21**, 1104–1115
 42. Shevchenko, A., Jensen, O. N., Podtelejnikov, A. V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Shevchenko, A., Boucherie, H., and Mann, M. (1996) Linking genome and proteome by mass spectrometry: Large-scale identification of yeast proteins from two dimensional gels. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 14440–14445
 43. Patterson, S. D., and Aebersold, R. H. (2003) Proteomics: The first decade and beyond. *Nat. Genet.* **33**, (suppl.) 311–323
 44. Bouwmeester, T., Bauch, A., Ruffner, H., Angrand, P. O., Bergamini, G., Crougton, K., Cruciat, C., Eberhard, D., Gagneur, J., Ghidelli, S., Hopf, C., Huhse, B., Mangano, R., Michon, A. M., Schirle, M., Schlegl, J., Schwab, M., Stein, M. A., Bauer, A., Casari, G., Drewes, G., Gavin, A. C., Jackson, D. B., Joberty, G., Neubauer, G., Rick, J., Kuster, B., and Superti-Furga, G. (2004) A physical and functional map of the human TNF- α /NF- κ B signal transduction pathway. *Nat. Cell Biol.* **6**, 97–105
 45. Durr, E., Yu, J., Krasinska, K. M., Carver, L. A., Yates, J. R., Testa, J. E., Oh, P., and Schnitzer, J. E. (2004) Direct proteomic mapping of the lung microvascular endothelial cell surface in vivo and in cell culture. *Nat. Biotechnol.* **22**, 985–992
 46. Everley, P. A., Krijgsveld, J., Zetter, B. R., and Gygi, S. P. (2004) Quantitative cancer proteomics: Stable isotope labeling with amino acids in cell culture (SILAC) as a tool for prostate cancer research. *Mol. Cell. Proteomics* **3**, 729–735
 47. Guina, T., Purvine, S. O., Yi, E. C., Eng, J., Goodlett, D. R., Aebersold, R., and Miller, S. I. (2003) Quantitative proteomic analysis indicates increased synthesis of a quinolone by *Pseudomonas aeruginosa* isolates from cystic fibrosis airways. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 2771–2776
 48. Shio, Y., Donohoe, S., Yi, E. C., Goodlett, D. R., Aebersold, R., and Eisenman, R. N. (2002) Quantitative proteomic analysis of Myc oncoprotein function. *EMBO J.* **21**, 5088–5096
 49. Wright, M. E., Eng, J., Sherman, J., Hockenbery, D. M., Nelson, P. S., Galitski, T., and Aebersold, R. (2003) Identification of androgen-coregulated protein networks from the microsomes of human prostate cancer cells. *Genome Biol.* **5**, R4
 50. Andersen, J. S., Wilkinson, C. J., Mayor, T., Mortensen, P., Nigg, E. A., and Mann, M. (2003) Proteomic characterization of the human centrosome by protein correlation profiling. *Nature* **426**, 570–574
 51. Blagoev, B., Kratchmarova, I., Ong, S. E., Nielsen, M., Foster, L. J., and Mann, M. (2003) A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat. Biotechnol.* **21**, 315–318
 52. Brand, M., Ranish, J. A., Kummer, N. T., Hamilton, J., Igarashi, K., Francastel, C., Chi, T. H., Crabtree, G. R., Aebersold, R., and Groudine, M. (2004) Dynamic changes in transcription factor complexes during erythroid differentiation revealed by quantitative proteomics. *Nat. Struct. Mol. Biol.* **11**, 73–80
 53. Ranish, J. A., Yi, E. C., Leslie, D. M., Purvine, S. O., Goodlett, D. R., Eng, J., and Aebersold, R. (2003) The study of macromolecular complexes by quantitative proteomics. *Nat. Genet.* **33**, 349–355
 54. Ranish, J. A., Hahn, S., Lu, Y., Yi, E. C., Li, X. J., Eng, J., and Aebersold, R. (2004) Identification of TFB5, a new component of general transcription and DNA repair factor IIH. *Nat. Genet.* **36**, 707–713
 55. Rout, M. P., Aitchison, J. D., Suprpto, A., Hjertaas, K., Zhao, Y., and Chait, B. T. (2000) The yeast nuclear pore complex: Composition, architecture, and transport mechanism. *J. Cell Biol.* **148**, 635–651
 56. Pusch, W., Flocco, M. T., Leung, S. M., Thiele, H., and Kostrzewa, M. (2003) Mass spectrometry-based clinical proteomics. *Pharmacogenomics* **4**, 463–476
 57. Cagney, G., and Emili, A. (2002) De novo peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abundance tagging. *Nat. Biotechnol.* **20**, 163–170
 58. Chakraborty, A., and Regnier, F. E. (2002) Global internal standard technology for comparative proteomics. *J. Chromatogr. A* **949**, 173–184
 59. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**, 994–999
 60. Zhou, H., Ranish, J. A., Watts, J. D., and Aebersold, R. (2002) Quantitative proteome analysis by solid-phase isotope tagging and mass spectrometry. *Nat. Biotechnol.* **20**, 512–515
 61. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **1**, 376–386

62. Veenstra, T. D., Martinovic, S., Anderson, G. A., Pasa-Tolic, L., and Smith, R. D. (2000) Proteome analysis using selective incorporation of isotopically labeled amino acids. *J. Am. Soc. Mass Spectrom.* **11**, 78–82
63. Mirgorodskaya, O. A., Kozmin, Y. P., Titov, M. I., Korner, R., Sonksen, C. P., and Roepstorff, P. (2000) Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using ¹⁸O-labeled internal standards. *Rapid Commun. Mass Spectrom.* **14**, 1226–1232
64. Yao, X., Freas, A., Ramirez, J., Demirev, P. A., and Fenselau, C. (2001) Proteolytic ¹⁸O labeling for comparative proteomics: Model studies with two serotypes of adenovirus. *Anal. Chem.* **73**, 2836–2842
65. Zhou, H., Watts, J. D., and Aebersold, R. (2001) A systematic approach to the analysis of protein phosphorylation. *Nat. Biotechnol.* **19**, 375–378
66. Petricoin, E. F., Ardekani, A. M., Hitt, B. A., Levine, P. J., Fusaro, V. A., Steinberg, S. M., Mills, G. B., Simone, C., Fishman, D. A., Kohn, E. C., and Liotta, L. A. (2002) Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* **359**, 572–577
67. Cottingham, K. (2003) Clinical proteomics: Are we there yet? *Anal. Chem.* **75**, 472A–476A
68. Diamandis, E. P. (2004) Mass spectrometry as a diagnostic and a cancer biomarker discovery tool: Opportunities and potential limitations. *Mol. Cell. Proteomics* **3**, 367–378
69. Diamandis, E. P. (2004) Analysis of serum proteomic patterns for early cancer diagnosis: Drawing attention to potential problems. *J. Natl. Cancer Inst.* **96**, 353–356
70. Koopmann, J., Zhang, Z., White, N., Rosenzweig, J., Fedarko, N., Jagannath, S., Canto, M. I., Yeo, C. J., Chan, D. W., and Goggins, M. (2004) Serum diagnosis of pancreatic adenocarcinoma using surface-enhanced laser desorption and ionization mass spectrometry. *Clin. Cancer Res.* **10**, 860–868
71. Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74**, 5383–5392
72. Nesvizhskii, A. I., Keller, A., Kolker, E., and Aebersold, R. (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75**, 4646–4658
73. Li, X. J., Zhang, H., Ranish, J. A., and Aebersold, R. (2003) Automated statistical analysis of protein abundance ratios from data generated by stable-isotope dilution and tandem mass spectrometry. *Anal. Chem.* **75**, 6648–6657
74. Aguirre, A. J., Brennan, C., Bailey, G., Sinha, R., Feng, B., Leo, C., Zhang, Y., Zhang, J., Gans, J. D., Bardeesy, N., Cauwels, C., Cordon-Cardo, C., Redston, M. S., DePinho, R. A., and Chin, L. (2004) High-resolution characterization of the pancreatic adenocarcinoma genome. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 9067–9072
75. Crnogorac-Jurcevic, T., Missiaglia, E., Blaveri, E., Gangeswaran, R., Jones, M., Terris, B., Costello, E., Neoptolemos, J. P., and Lemoine, N. R. (2003) Molecular alterations in pancreatic carcinoma: Expression profiling shows that dysregulated expression of S100 genes is highly prevalent. *J. Pathol.* **201**, 63–74
76. Grutzmann, R., Foerder, M., Aldinger, I., Staub, E., Brummendorf, T., Ropcke, S., Li, X., Kristiansen, G., Jesnowski, R., Sipos, B., Lohr, M., Luttgies, J., Ockert, D., Kloppel, G., Saeger, H. D., and Pilarsky, C. (2003) Gene expression profiles of microdissected pancreatic ductal adenocarcinoma. *Virchows Arch.* **443**, 508–517
77. Holzmann, K., Kohlhammer, H., Schwaenen, C., Wessendorf, S., Kestler, H. A., Schwoerer, A., Rau, B., Radlwimmer, B., Dohner, H., Lichter, P., Gress, T., and Bentz, M. (2004) Genomic DNA-chip hybridization reveals a higher incidence of genomic amplifications in pancreatic cancer than conventional comparative genomic hybridization and leads to the identification of novel candidate genes. *Cancer Res.* **64**, 4428–4433
78. Hustinx, S. R., Cao, D., Maitra, A., Sato, N., Martin, S. T., Sudhir, D., Iacobuzio-Donahue, C., Cameron, J. L., Yeo, C. J., Kern, S. E., Goggins, M., Mollenhauer, J., Pandey, A., and Hruban, R. H. (2004) Differentially expressed genes in pancreatic ductal adenocarcinomas identified through serial analysis of gene expression. *Cancer Biol. Ther.* **3**, in press
79. Iacobuzio-Donahue, C. A., Ashfaq, R., Maitra, A., Adsay, N. V., Shen-Ong, G. L., Berg, K., Hollingsworth, M. A., Cameron, J. L., Yeo, C. J., Kern, S. E., Goggins, M., and Hruban, R. H. (2003) Highly expressed genes in pancreatic ductal adenocarcinomas: A comprehensive characterization and comparison of the transcription profiles obtained from three major technologies. *Cancer Res.* **63**, 8614–8622
80. Kohlhammer, H., Schwaenen, C., Wessendorf, S., Holzmann, K., Kestler, H. A., Kienle, D., Barth, T. F., Moller, P., Ott, G., Kalla, J., Radlwimmer, B., Pscherer, A., Stilgenbauer, S., Dohner, H., Lichter, P., and Bentz, M. (2004) Genomic DNA-chip hybridization in t(11;14)-positive mantle cell lymphomas shows a high frequency of aberrations and allows a refined characterization of consensus regions. *Blood* **104**, 795–801
81. Missiaglia, E., Blaveri, E., Terris, B., Wang, Y.H., Costello, E., Neoptolemos, J. P., Crnogorac-Jurcevic, T., and Lemoine, N. R. (2004) Analysis of gene expression in cancer cell lines identifies candidate markers for pancreatic tumorigenesis and metastasis. *Int. J. Cancer* **112**, 100–112
82. Nakamura, T., Furukawa, Y., Nakagawa, H., Tsunoda, T., Ohigashi, H., Murata, K., Ishikawa, O., Ohgaki, K., Kashimura, N., Miyamoto, M., Hirano, S., Kondo, S., Katoh, H., Nakamura, Y., and Katagiri, T. (2004) Genome-wide cDNA microarray analysis of gene expression profiles in pancreatic cancers using populations of tumor cells and normal ductal epithelial cells selected for purity by laser microdissection. *Oncogene* **23**, 2385–2400
83. Sato, N., Fukushima, N., Maitra, A., Iacobuzio-Donahue, C. A., Van Heek, N. T., Cameron, J. L., Yeo, C. J., Hruban, R. H., and Goggins, M. (2004) Gene expression profiling identifies genes associated with invasive intraductal papillary mucinous neoplasms of the pancreas. *Am. J. Pathol.* **164**, 903–914
84. Liotta, L. A., and Kohn, E. C. (2001) The microenvironment of the tumour-host interface. *Nature* **411**, 375–379
85. Wernert, N. (1997) The multiple roles of tumour stroma. *Virchows Arch.* **430**, 433–443
86. Brown, L. F., Guidi, A. J., Schnitt, S. J., Van De, W. L., Iruela-Arispe, M. L., Yeo, T. K., Tognazzi, K., and Dvorak, H. F. (1999) Vascular stroma formation in carcinoma in situ, invasive carcinoma, and metastatic carcinoma of the breast. *Clin. Cancer Res.* **5**, 1041–1056
87. Kellner, U., Steinert, R., Seibert, V., Heim, S., Kellner, A., Schulz, H. U., Roessner, A., Kruger, S., and Reymond, M. (2004) Epithelial cell preparation for proteomic and transcriptomic analysis in human pancreatic tissue. *Pathol. Res. Pract.* **200**, 155–163
88. Goke, B., Keim, V., Dagorn, J. C., Arnold, R., and Adler, G. (1990) Resolution of human exocrine pancreatic juice proteins by reversed-phase high performance liquid chromatography (HPLC). *Pancreas* **5**, 261–266
89. Keim, V., Iovanna, J. L., Rohr, G., Usadel, K. H., and Dagorn, J. C. (1991) Characterization of a rat pancreatic secretory protein associated with pancreatitis. *Gastroenterology* **100**, 775–782
90. Scheele, G. A. (1975) Two-dimensional gel analysis of soluble proteins. Characterization of guinea pig exocrine pancreatic proteins. *J. Biol. Chem.* **250**, 5375–5385
91. Scheele, G. A., and Palade, G. E. (1975) Studies on the guinea pig pancreas. Parallel discharge of exocrine enzyme activities. *J. Biol. Chem.* **250**, 2660–2670
92. Wandschneider, S., Fehring, V., Jacobs-Emeis, S., Thiesen, H. J., and Lohr, M. (2001) Autoimmune pancreatic disease: Preparation of pancreatic juice for proteome analysis. *Electrophoresis* **22**, 4383–4390
93. Gronborg, M., Bunkenborg, J., Kristiansen, T. Z., Jensen, O. N., Yeo, C. J., Hruban, R. H., Maitra, A., Goggins, M. G., and Pandey, A. (2004) Comprehensive proteomic analysis of human pancreatic juice. *J. Proteome Res.* **3**, 1042–1055
94. Rosty, C., Christa, L., Kuzdzal, S., Baldwin, W. M., Zahurak, M. L., Carnot, F., Chan, D. W., Canto, M., Lillemoe, K. D., Cameron, J. L., Yeo, C. J., Hruban, R. H., and Goggins, M. (2002) Identification of hepatocarcinoma-intestine-pancreas/pancreatitis-associated protein I as a biomarker for pancreatic ductal adenocarcinoma by protein biochip technology. *Cancer Res.* **62**, 1868–1875
95. Anderson, N. L., and Anderson, N. G. (2002) The human plasma proteome: History, character, and diagnostic prospects. *Mol. Cell. Proteomics* **1**, 845–867
96. Shen, Y., Jacobs, J. M., Camp, D. G., Fang, R., Moore, R. J., Smith, R. D., Xiao, W., Davis, R. W., and Tompkins, R. G. (2004) Ultra-high-efficiency strong cation exchange LC/RPLC/MS/MS for high dynamic range characterization of the human plasma proteome. *Anal. Chem.* **76**, 1134–1144
97. Ahmed, N., Barker, G., Oliva, K., Garfin, D., Talmadge, K., Georgiou, H., Quinn, M., and Rice, G. (2003) An approach to remove albumin for the proteomic analysis of low abundance biomarkers in human serum. *Proteomics* **3**, 1980–1987

98. Brzeski, H., Katenhusen, R. A., Sullivan, A. G., Russell, S., George, A., Somiari, R. I., and Shriver, C. (2003) Albumin depletion method for improved plasma glycoprotein analysis by two-dimensional difference gel electrophoresis. *BioTechniques* **35**, 1128–1132
99. Govorukhina, N. I., Keizer-Gunnink, A., van der Zee, A. G., de Jong, S., de Bruijn, H. W., and Bischoff, R. (2003) Sample preparation of human serum for the analysis of tumor markers. Comparison of different approaches for albumin and gamma-globulin depletion. *J. Chromatogr. A* **1009**, 171–178
100. Greenough, C., Jenkins, R. E., Kitteringham, N. R., Pirmohamed, M., Park, B. K., and Pennington, S. R. (2004) A method for the rapid depletion of albumin and immunoglobulin from human plasma. *Proteomics* **4**, 3107–3111
101. Rothmund, D. L., Locke, V. L., Liew, A., Thomas, T. M., Wasinger, V., and Rylatt, D. B. (2003) Depletion of the highly abundant protein albumin from human plasma using the Gradiflow. *Proteomics* **3**, 279–287
102. Kaji, H., Saito, H., Yamauchi, Y., Shinkawa, T., Taoka, M., Hirabayashi, J., Kasai, K., Takahashi, N., and Isobe, T. (2003) Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify *N*-linked glycoproteins. *Nat. Biotechnol.* **21**, 667–672
103. Zhang, H., Li, X. J., Martin, D. B., and Aebersold, R. (2003) Identification and quantification of *N*-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.* **21**, 660–666
104. Zhang, H., Yi, E. C., Li, X. J., Mallick, P., Kelly-Spratt, K. S., Masselon, C. D., Camp, D. G., 2nd, Smith, R. D., Kemp, C. J., and Aebersold, R. (2004). High throughput quantitative analysis of serum proteins using glycopeptide capture and liquid chromatography mass spectrometry. *Mol. Cell. Proteomics* **4**, 144–155
105. Kononen, J., Bubendorf, L., Kallioniemi, A., Barlund, M., Schraml, P., Leighton, S., Torhorst, J., Mihatsch, M. J., Sauter, G., and Kallioniemi, O. P. (1998) Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat. Med.* **4**, 844–847
106. Kumble, K. D., Hirota, M., Pour, P. M., and Vishwanatha, J. K. (1992) Enhanced levels of annexins in pancreatic carcinoma cells of Syrian hamsters and their intrapancreatic allografts. *Cancer Res.* **52**, 163–167
107. Vishwanatha, J. K., Chiang, Y., Kumble, K. D., Hollingsworth, M. A., and Pour, P. M. (1993) Enhanced expression of annexin II in human pancreatic carcinoma cells and primary pancreatic cancers. *Carcinogenesis* **14**, 2575–2579
108. Pan, S., Zhang, H., Rush, J., Eng, J., Zhang, N., Patterson, D., Comb, M. J., and Aebersold, R. H. (2005). High-throughput proteome-screening for biomarker detection. *Mol. Cell. Proteomics* **4**, 182–190