[CANCER RESEARCH 63, 8614–8622, December 15, 2003]

# Highly Expressed Genes in Pancreatic Ductal Adenocarcinomas: A Comprehensive Characterization and Comparison of the Transcription Profiles Obtained from Three Major Technologies

**Christine A. Iacobuzio-Donahue,[1,3] Raheela Ashfaq,[5] Anirban Maitra,[1] N. Volkan Adsay,[6] Grace L. Shen-Ong,[7] Karin Berg,[1] Michael A. Hollingsworth,[8] John L. Cameron,[2] Charles J. Yeo,[2] Scott E. Kern,[1,3] Michael Goggins,[1,4] and Ralph H. Hruban[1,3]**

*Departments of [1]Pathology, [2]Surgery, [3]Oncology, and [4]Medicine, The Johns Hopkins Medical Institutions, Baltimore, Maryland; [5]Department of Pathology, University of Texas Southwestern Medical Center, Dallas, Texas; [6]Department of Pathology, Wayne State University, Detroit, Michigan; [7]Gene Logic, Inc., Gaithersburg, Maryland; and [8]The Eppley Institute, University of Nebraska, Omaha, Nebraska*

## ABSTRACT

**When using gene expression profiling to understand human tumors, one is often confronted with long lists of genes that need to be further categorized into meaningful data. We performed a comprehensive evaluation and comparison of gene expression profiles obtained from pancreatic cancers to determine those genes most differentially expressed and thus with the most promise for translation into clinically useful targets. cDNA was prepared from 50 samples of normal pancreas or duodenal mucosal tissues, 7 samples of chronic pancreatitis, and 39 samples of pancreas cancer tissues or cancer cell lines and hybridized to the complete Affymetrix Human Genome U133 GeneChip set (arrays U133A and U133B) for simultaneous analysis of 45,000 fragments corresponding to 33,000 known genes and 6,000 expressed sequence tags. Genes expressed at levels at least 3-fold greater in the pancreatic cancers as compared with nonneoplastic tissues were identified. Three hundred seventy-seven Affymetrix fragments were identified as having ≥3-fold expression levels in pancreas cancer specimens as compared with nonneoplastic tissues, corresponding to 234 known genes. Serial analysis of gene expression libraries (http://www.ncbi.nlm.nih.gov/SAGE/) of two normal pancreatic ductal cell cultures (HX and H126) were used to exclude 17 genes with high expression levels in the normal duct epithelium (more than five tags/library). Of the remaining 217 known genes, 75 have been previously reported as highly expressed in pancreatic cancers, while the remaining 142 genes are novel. We used principal components analysis (PCA) to identify the genes among these 217 identified as the most differentially expressed and specific to pancreatic cancer tissues or cell lines. Among the most differentially expressed genes identified by PCA were Mesothelin, Muc4, Muc5A/C, Kallikrein 10, Transglutaminase 2, Fascin, TMPRSS3 and stratifin. The differential expression identified by PCA for these genes indicates they are among the more attractive targets for novel therapeutic targets, tumor markers, or as a means of screening pancreatic cancer samples for information regarding tumor classification or potential therapeutic responses. Our findings were also compared in detail to the previously reported findings of highly expressed genes in other studies of global gene expression in pancreatic cancers. We found that robust changes in gene expression were most often identified by more than one gene expression platform. Forty genes were identified by more than one method (U133 oligonucleotide arrays, cDNA arrays or serial analysis of gene expression), and 6 of these genes were identified by all three methods. Our findings identify a novel set of genes as highly expressed in pancreatic cancer, validate the differential expression of previously reported genes, and provide additional support for those genes most differentially expressed to be translated into clinically useful targets.**

## INTRODUCTION

Gene expression technologies are increasingly being used to understand the complexities of transcriptional regulation and dysregulation in normal and cancerous tissues. One important application of these methods is for the rapid identification of differentially expressed genes in neoplastic tissues as compared with their normal counterparts. These differentially expressed genes may form the basis of the development of novel diagnostic markers or therapeutic targets.

The recent mapping of the human genome has greatly facilitated the success of global gene expression methodologies in the study of human cancers because these technologies are based on coding sequences derived from the human genome project (1, 2). Recently, Affymetrix, Inc., has produced the UG-133 GeneChip based on Build 133 of the human genome.[9] One particular advantage of the UG-133 chip is that many transcripts previously identified only as expressed sequence tags (ESTs) have now been characterized, allowing for a more comprehensive view of gene expression in samples studied.

More than 100 novel overexpressed markers of pancreatic cancer have now been identified using a variety of global gene expression methodologies, including representational difference analysis and subtractive hybridization (3–5), differential display (6), serial analysis of gene expression (SAGE; 7–9), cDNA microarrays (10–13), and oligonucleotide arrays (14). A major issue in the use of gene expression profiling is how these different platforms compare with each other with respect to various parameters such as sensitivity and specificity in detecting highly expressed genes. These and other factors may affect not only one's choice of methodology used in designing an experiment, but also how the derived data should be interpreted.

In this study, we have analyzed a large set of pancreatic adenocarcinomas, pancreas cancer cell lines, chronic pancreatitis, normal pancreas, and normal duodenal tissues using the UG-133 GeneChip in an effort to (*a*) further characterize the gene expression signature of pancreatic cancers; (*b*) identify novel markers of pancreatic cancer not recognized using prior methods, including our previous study using UG-95 GeneChips; (*c*) identify those highly expressed genes with the most promise for development into novel diagnostic or therapeutic markers; and (*d*) compare our findings with previously reported studies of gene expression in pancreatic cancers.

## MATERIALS AND METHODS

**Tissues.** Samples (0.5 g) of normal pancreas ($n = 25$), normal duodenal mucosa ($n = 25$), chronic pancreatitis ($n = 7$), and infiltrating pancreatic adenocarcinoma ($n = 26$) were collected from surgical resection specimens from patients at The Johns Hopkins Hospital. Eleven of these samples of normal pancreas and 14 of these samples of infiltrating pancreatic carcinoma

Received 1/24/03; revised 8/1/03; accepted 8/27/03.

**Grant support:** Jeff Zgonina Fund and the Michael Rolfe Fund for Pancreatic Cancer Research and the NIH Specialized Programs of Research Excellence in Gastrointestinal Cancer Grant CA62924.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

**Requests for reprints:** Christine A. Iacobuzio-Donahue, The Johns Hopkins Hospital, Division of Gastrointestinal/Liver Pathology, 720 Rutland Street, Ross Building Room 632, Baltimore, MD 21205. Phone: (410) 955-3511; Fax: (410) 614-0671; E-mail: ciacobu@jhmi.edu.

[9] Internet address: http://www.affymetrix.com.

were previously analyzed by U95 GeneChips and reported in our previous study (13). In each case, the specimens were harvested within 10 min of resection from the patient and snap frozen in liquid nitrogen before storage at −80°C. H&E-stained sections of adjacent sections of the tissue were prepared before snap freezing to confirm the diagnosis. The normal duodenal mucosa samples were included in the analyses to facilitate the identification of markers of pancreatic cancer that would be useful in screening secondary sources such as in duodenal fluid samples.

**Cell Lines.** Human pancreatic cancer cell lines AsPc1, CAPAN1, CAPAN2, CFPAC1, COLO357, Hs766T, MiaPaCa2, and Panc-1 were obtained from the American Type Culture Collection (Manassas, VA). Panc 2.13, Panc 4.14, Panc 6.03, Panc 8.13, Panc 3.27, and PL4 are low-passage pancreatic carcinoma cell lines kindly provided by Dr. Elizabeth Jaffee (15). Panc 2.13, Panc 4.14, Panc 6.03, Panc 8.13, and Panc 3.27 were recently submitted to the American Type Culture Collection and are now more generally available online.[10] All cell lines were cultured in DMEM supplemented with 10% FBS and antibiotics (100 units/ml penicillin and 100 $\mu$g/ml streptomycin) with the exception of the CAPAN1 and CAPAN2 cell lines cultured in RPMI 1640 (Life Technologies, Inc., Gaithersburg, MD) supplemented with 10% FBS and antibiotics (100 units/ml penicillin and 100 $\mu$g/ml streptomycin). The use of different media minimized the variance in growth rates that would otherwise be exaggerated with a single medium. Cells were incubated at 37°C in a humidified atmosphere of 5% $CO_2$ in air.

**mRNA Extractions and Affymetrix GeneChip Hybridization.** Sample preparation and processing procedure was performed as described in detail in the Affymetrix GeneChip Expression Analysis Manual (Santa Clara, CA) and in our prior publication (14) with the exception that the labeled cRNA samples were hybridized to the complete human U133 GeneChip set (HG_U133 A and HG_U133 B).

**Statistical Data Analysis.** The GeneExpress Software System Fold Change Analysis tool was used to identify all present genes expressed at least 3-fold greater in the pancreatic cancers compared with normal tissues. For each gene fragment, the ratio of the geometric means of the expression intensities in the normal control tissues and the pancreatic cancer samples was calculated, and the fold change was then calculated on a per fragment basis. Confidence limits were calculated using a two-sided Welch modified *t* test on the difference of the means of the logs of the intensities.

**Biocomputational Tools.** The Cluster and TreeView computer programs were obtained from an online resource[11] and used for principal component analysis (PCA) and visualization of the data (16). Raw expression data for all gene fragments corresponding to known genes were imported into the Cluster program, log-transformed, and PCA performed with visualization of the tree diagrams generated using the TreeView program.

**SAGE.** SAGE library data from the short-term cultures of non-neoplastic pancreatic duct epithelial cells available on the Cancer Genome Anatomy Project web site[11] as part of the SAGEmap database were used to identify genes expressed in normal ductal cells. The preparation and validation of these short-term cultures of nonneoplastic pancreatic ductal epithelial cells (HX and H126) were described previously (17).

**Construction of Tissue Microarrays and Immunohistochemical Labeling.** Tissue microarrays containing 68 different surgically resected infiltrating ductal pancreatic adenocarcinomas and a variety of normal tissues were constructed as described previously (18). Each carcinoma was represented four times in each tissue microarray to account for potential tumor heterogeneity. Unstained 4-$\mu$m sections were cut from each tissue microarray and deparaffinized by routine techniques before placing in 200 ml of Target Retrieval Solution (pH 6.0, Envision Plus Detection kit; Dako, Carpenteria, CA) for 20 min at 100°C. After cooling for 20 min, slides were quenched with 3% $H_2O_2$ for 5 min before incubating with the appropriate dilution of each primary antibody (a 1:1000 dilution of mouse monoclonal antihuman Muc 5A/C or a 1:1000 dilution of polyclonal antihuman ESDN; Orbigen, San Diego, CA) for 30 min using the Dako Autostainer. Labeling was detected with the Dako Envision system following the manufacturers protocol. Labeling was detected by adding biotinylated secondary antibodies, avidin-biotin complex, and 3,3′-diaminobenzidine. All sections were counterstained with hematoxylin.

**Comparison of Gene Expression Patterns Identified by Different Methodologies.** We have previously reported the gene expression profiles of pancreatic cancer identified by U95 oligonucleotide arrays (14) by cDNA microarrays (10) and by SAGE (7–9). The data obtained from these studies was compiled into one Excel spreadsheet for a comparison of gene expression profiling of pancreatic cancers, with all otherwise nonannotated (orphan) ESTs excluded because of the inability to cross validate the identities of many of these ESTs between different gene expression platforms.

## RESULTS

**Data Filtering.** cRNA samples were hybridized to the complete Affymetrix Human Genome U133 GeneChip set (arrays U133 A and B) for simultaneous analysis of 45,000 fragments corresponding to 33,000 known genes and 6,000 ESTs. Genes with a 3-fold greater increase in expression in the 39 pancreatic adenocarcinoma tissues or cell lines compared with 50 normal tissues were identified. Thresholds of expression were set to include all genes that were expressed in at least 50% of the malignant samples (cell lines and tumor tissues) and in no more than 5% of the normal tissues. We identified 377 fragments expressed at least 3-fold greater in pancreatic cancer samples as compared with normal tissues, of which 279 corresponded to known genes and 92 were expressed >5-fold. The level of significance for each gene fragment ranged from less than $P = 0.00001$ to $P = 0.01$ (modified Welch *t* test).

**Identification of Highly Expressed Genes in Pancreatic Cancers.** Among these 279 fragments identified corresponding to known genes, 38 genes were represented by more than one Affymetrix fragment (range, 2–4). When these fragments were accounted for, 234 individual genes were identified. Those expressed ≥4-fold are listed in Table 1. Eighty-one of these 234 known genes have been previously reported as overexpressed in pancreatic cancer (Table 2).

In gene expression analyses of bulk normal pancreas, the normal pancreatic duct epithelium is underrepresented because of the predominance of acinar cells, stroma, and islets of Langerhans. Therefore, the 234 candidate genes identified were additionally refined to exclude genes highly expressed in short-term cultures of normal pancreatic ductal epithelial cells. For each gene identified as differentially expressed by Affymetrix GeneChip, the corresponding SAGE tag was identified, and the total number of SAGE tags present in the SAGEmap database[12] of normal pancreas duct epithelium libraries HX and H126 was determined. Any gene having more than five tags in at least one of these two SAGE libraries was then excluded from additional analysis. Using this approach, 17 genes were identified as having high levels of expression in normal pancreatic ductal epithelium. These genes were excluded, leaving 217 genes identified as highly expressed in pancreatic cancer. A PubMed search for each of these genes indicated that 143 are now reported for the first time in association with pancreatic cancer.

**PCA of Highly Expressed Genes.** In the evaluation of highly expressed genes in pancreatic cancer, it would be useful to identify those genes that may be differentially expressed between the normal and neoplastic epithelium (as opposed to those genes generally upregulated in the presence of chronic pancreatitis) and therefore more specific tumor markers and therapeutic targets. We used PCA to identify those highly expressed genes that are most differentially expressed among all 217 known genes identified. PCA can provide a global overview of the relatedness of gene expression profiles among samples while better avoiding the deterministic and rather arbitrary nature of hierarchical clustering. This delineation allows not only for a better understanding of the role of these genes in pancreas cancers but also those most promising for development into novel screening tests and therapeutic targets.

PCA of the expression data derived from 96 samples representing

[10] Internet address: http://www.atcc.org.
[11] Internet address: http://www.microarrays.org/software.html.
[12] Internet address: http://www.ncbi.nlm.nih.gov/SAGE/.

Table 1 *Novel highly expressed genes in pancreatic cancer identified by U133 oligonucleotide arrays*[a]

| Known gene name | Cluster ID | Fold change[b] |
|---|---|---|
| Neuromedin U | Hs.2841 | 18.11 |
| small inducible cytokine subfamily B (Cys-X-Cys), member 5 (epithelial-derived neutrophil-activating peptide 78) | Hs.89714 | 16.36 |
| melanoma inhibitory activity | Hs.279651 | 16.3 |
| small inducible cytokine subfamily B (Cys-X-Cys), member 5 (epithelial-derived neutrophil-activating peptide 78) | Hs.89714 | 15.78 |
| kallikrein 10 | Hs.69423 | 12.37 |
| sciellin | Hs.115166 | 9.77 |
| solute carrier family 21 (organic anion transporter), member 11 | Hs.14805 | 9.34 |
| FOS-like antigen 1 | Hs.283565 | 9.08 |
| phosphorylase, glycogen; brain | Hs.75658 | 6.37 |
| Ras-induced senescence 1 | Hs.35861 | 6.36 |
| phorbol-12-myristate-13-acetate-induced protein 1 | Hs.96 | 6.21 |
| cystatin SN | Hs.123114 | 6.15 |
| gap junction protein, $\beta$ 3, 31kD (connexin 31) | Hs.98485 | 6.12 |
| Zic family member 2 (odd-paired homolog, Drosophila) | Hs.132863 | 6.09 |
| S100 calcium binding protein A2 | Hs.38991 | 5.84 |
| cartilage oligomeric matrix protein (pseudoachondroplasia, epiphyseal dysplasia 1, multiple) | Hs.1584 | 5.78 |
| sushi-repeat protein | Hs.126782 | 5.62 |
| tripartite motif-containing 29 | Hs.82237 | 5.55 |
| matrilin 3 | Hs.278461 | 5.46 |
| phorbol-12-myristate-13-acetate-induced protein 1 | Hs.96 | 5.45 |
| collagen, type X, $\alpha$ 1 (Schmid metaphyseal chondrodysplasia) | Hs.179729 | 5.28 |
| collagen, type X, $\alpha$ 1 (Schmid metaphyseal chondrodysplasia) | Hs.179729 | 5.25 |
| endothelial and smooth muscle cell-derived neuropilin-like protein | Hs.173374 | 5.19 |
| dickkopf homologue 1 (*Xenopus laevis*) | Hs.40499 | 5.13 |
| RAB34, member RAS oncogene family | Hs.301853 | 5.12 |
| RAS protein activator like 2 | Hs.227806 | 5.11 |
| MAD1 mitotic arrest deficient-like 1 (yeast) | Hs.7345 | 5.06 |
| Homer, neuronal immediate early gene, 1B | Hs.337737 | 4.97 |
| ubiquitin carrier protein | Hs.174070 | 4.94 |
| protease inhibitor 3, skin-derived (SKALP) | Hs.112341 | 4.92 |
| cadherin, EGF LAG seven-pass G-type receptor 1 (flamingo homolog, Drosophila) | Hs.252387 | 4.84 |
| BH-protocadherin (brain-heart) | Hs.34073 | 4.82 |
| TONDU | Hs.9030 | 4.79 |
| protein tyrosine phosphatase, nonreceptor type 12 | Hs.62 | 4.77 |
| claudin 18 | Hs.278966 | 4.75 |
| microfibrillar-associated protein 2 | Hs.83551 | 4.75 |
| HSPC150 protein similar to ubiquitin-conjugating enzyme | Hs.5199 | 4.69 |
| BCL2-interacting killer (apoptosis-inducing) | Hs.155419 | 4.68 |
| NAD(P)H dehydrogenase, quinone 1 | Hs.80706 | 4.67 |
| transducin ($\beta$)-like 1 | Hs.76536 | 4.59 |
| anillin, actin binding protein (scraps homolog, Drosophila) | Hs.62180 | 4.57 |
| KIAA0233 gene product | Hs.79077 | 4.54 |
| cystatin E/M | Hs.83393 | 4.42 |
| caspase recruitment domain family, member 11 | Hs.293867 | 4.41 |
| acetyl-Coenzyme A carboxylase $\alpha$ | Hs.7232 | 4.4 |
| BCL2-related ovarian killer | Hs.293753 | 4.39 |
| epidermal growth factor receptor pathway substrate 8 related protein 1 | Hs.28907 | 4.37 |
| homeo box A10 | Hs.110637 | 4.34 |
| chloride intracellular channel 3 | Hs.64746 | 4.32 |

[a] Shown are the novel genes with the greatest fold change value in pancreas cancer tissues or pancreas cancer cell lines as compared with normal pancreas or normal duodenal mucosa.

[b] All genes had a $P$ of at least 0.1 to $<0.0001$ as described in "Materials and Methods" section.

the 50 normal, 7 chronic pancreatitis, and 39 pancreatic cancers confirmed the overall high levels of expression of these genes in cancer tissue and cell line samples in comparison to normal duodenal or pancreas tissues (Fig. 1). However, three patterns were visualized based on this analysis. First, we noted that some of these genes were only expressed within the resected cancer tissues (*e.g., collagen type X*$\alpha$1) or the cancer cell lines (*e.g., pleckstrin homology-like domain* and *family A member 1*) as compared with normal tissues but not in both the resected cancers and the cancer cell lines. Those genes expressed among the resected pancreas cancer tissues and not the cell lines may reflect either genes expressed by the cellular components of the host stromal response or genes expressed by the neoplastic epithelium as a result of its interactions with the juxtatumoral stroma (19, 20). In contrast, genes expressed by the pancreas cancer cell lines but not the resected cancer tissues may reflect the enhanced proliferation rate commonly observed for cells growing in artificial conditions. For example, the *Ki67* proliferation antigen was highly expressed in the pancreas cancer cell lines and normal duodenal mucosa (a highly proliferative tissue) but not the resected cancers.

In the second pattern revealed by PCA, some genes were expressed in all sample types (neoplastic, chronic pancreatitis, and normal) but were relatively overexpressed within the neoplastic samples. Genes with this pattern identified by PCA included some that have been previously reported as highly expressed in pancreas cancer (*e.g., Trop 2* and *lipocalin 2*), whereas others are newly identified such as *ESDN* (endothelial and smooth muscle cell-derived neuropilin like protein). Relatively higher expression in pancreas cancers than that detected in normal pancreas suggests an increased transcription of these genes by the neoplastic cells, although an increased representation of nonneoplastic cells expressing these genes in tumor tissues is also a possibility (see section on "Verification and Localization of Selected Candidate Tumor Markers").

In the third pattern revealed by PCA, many genes were noted to be specifically expressed among both the resected cancer tissues and cancer cell lines but were not detected in normal pancreas, normal duodenum, or chronic pancreatitis tissues. These genes appear to be the most specifically expressed within pancreas cancers among all of the highly expressed genes identified. Some of these genes have previously been reported as highly expressed in pancreatic cancer (*e.g., mesothelin, TMPRSS3, Muc5A/C,* and *Fascin*), whereas several others identified with this pattern are novel (*e.g., antigen p97, BCL-2 interacting killer, cathepsin L2, kallikrein 10,* and *matrilin 3*).

Table 2 *Genes previously reported as highly expressed in pancreas cancer and identified in the current study*

| Known gene name | Cluster ID | Reference[a] |
|---|---|---|
| ADP-ribosylation factor-like 7 | Hs.111554 | 10,14 |
| Annexin A1 | Hs.78225 | 11,14,10 |
| antigen identified by monoclonal antibody Ki-67 | Hs.80976 | 11 |
| aspartate beta-hydroxylase | Hs.283664 | 14 |
| cadherin 3, type 1, P-cadherin (placental) | Hs.2877 | 42 |
| capping protein (actin filament), gelsolin-like | Hs.82422 | 10 |
| carcinoembryonic antigen-related cell adhesion molecule 5 | Hs.220529 | 7,10 |
| carcinoembryonic antigen-related cell adhesion molecule 6 (nonspecific cross-reacting antigen) | Hs.73848 | 6,7 |
| cathepsin L2 | Hs.87417 | 43 |
| caveolin 2 | Hs.139851 | 14,10 |
| cell division cycle 2, $G_1$–S and $G_2$–M | Hs.334562 | 44 |
| cell division cycle 2-like 1 (PITSLRE proteins) | Hs.214291 | 45 |
| cyclin B1 | Hs.23960 | 44 |
| cyclin D1 (PRAD1: parathyroid adenomatosis 1) | Hs.82932 | 46 |
| drebrin 1 | Hs.89434 | 14 |
| epiregulin | Hs.115263 | 47 |
| eukaryotic translation initiation factor 2C, 2 | Hs.193053 | 14 |
| fer-1-like 3, myoferlin (C. elegans) | Hs.234680 | 10,14,11 |
| FOS-like antigen 1 | Hs.283565 | 11 |
| gap junction protein, $\beta$ 2, 26 kD (connexin 26) | Hs.323733 | 14 |
| heparan sulfate (glucosamine) 3-O-sulfotransferase 1 | Hs.40968 | 10 |
| homeo box B6 | Hs.98428 | 10 |
| IGF-II mRNA-binding protein 3 | Hs.79440 | 14 |
| immortalization-up-regulated protein | Hs.348553 | 10 |
| inhibin, beta A (activin A, activin AB $\alpha$ polypeptide) | Hs.727 | 48 |
| insulin-like growth factor binding protein 3 | Hs.77326 | 14 |
| integrin, $\beta$-like 1 (with EGF-like repeat domains) | Hs.82582 | 14 |
| interleukin 8 | Hs.624 | 14,49 |
| keratin 17 | Hs.2785 | 10,50 |
| keratin 19 | Hs.182265 | 7,37,10 |
| keratin 7 | Hs.23881 | 14,50 |
| KIAA1199 protein | Hs.50081 | 14 |
| KIAA1265 protein | Hs.355925 | 14 |
| KIAA1363 protein | Hs.22941 | 14 |
| laminin, $\gamma$ 2 [nicein (100kD), kalinin (105kD), BM600 (100kD), Herlitz junctional epidermolysis bullosa)] | Hs.54451 | 14,10 |
| lectin, galactoside-binding, soluble, 1 (galectin 1) | Hs.227751 | 51 |
| lipocalin 2 (oncogene 24p3) | Hs.204238 | 10,35 |
| matrix metalloproteinase 11 (stromelysin 3) | Hs.155324 | 52 |
| matrix metalloproteinase 7 (matrilysin, uterine) | Hs.2256 | 53 |
| mesothelin | Hs.155981 | 7,9 |
| met proto-oncogene (hepatocyte growth factor receptor) | Hs.285754 | 54 |
| mitogen-activated protein kinase kinase kinase kinase 4 | Hs.3628 | 14 |
| mucin 4, tracheobronchial | Hs.198267 | 22,21 |
| mucin 5, subtype B, tracheobronchial | Hs.102482 | 21 |
| mucin 5, subtypes A and C, tracheobronchial/gastric | Hs.103707 | 21 |
| nicotinamide *N*-methyltransferase | Hs.76669 | 14 |
| plasminogen activator, urokinase | Hs.77274 | 14,55 |
| plasminogen activator, urokinase receptor | Hs.179657 | 11,14,55,56 |
| pleckstrin homology-like domain, family A, member 1 | Hs.82101 | 14 |
| plectin 1, intermediate filament binding protein, 500kD | Hs.79706 | 14 |
| prostate stem cell antigen | Hs.20166 | 8 |
| RAP2B, member of RAS oncogene family | Hs.239527 | 10 |
| retinoic acid induced 3 | Hs.194691 | 7,10,14 |
| runt-related transcription factor 1 (acute myeloid leukemia 1; aml1 oncogene) | Hs.129914 | 14 |
| RuvB-like 1 (E. coli) | Hs.272822 | 10 |
| S100 calcium binding protein A11 (calgizzarin) | Hs.256290 | 10,11 |
| S100 calcium binding protein A4 (calcium protein, calvasculin, metastasin, murine placental homologue) | Hs.81256 | 57 |
| S100 calcium binding protein P | Hs.2962 | 14,10,13 |
| secretory leukocyte protease inhibitor (antileukoproteinase) | Hs.251754 | 10,7 |
| serine (or cysteine) proteinase inhibitor, clade H (heat shock protein 47), member 1, (collagen binding protein 1), serine (or cysteine) proteinase inhibitor, clade H (heat shock protein 47), member 2 | Hs.241579 | 14 |
| serum-inducible kinase | Hs.3838 | 14 |
| singed-like (fascin homolog, sea urchin) (Drosophila) | Hs.118400 | 14 |
| solute carrier family 16 (monocarboxylic acid transporters), member 3 | Hs.85838 | 14 |
| solute carrier family 21 (organic anion transporter), member 12 | Hs.235782 | 14 |
| stratifin | Hs.184510 | 38,7,10 |
| tetraspan 1 | Hs.38972 | 7,10 |
| thrombospondin 2 | Hs.108623 | 14 |
| topoisomerase (DNA) II $\alpha$ (170kD) | Hs.156346 | 14 |
| transcription factor BMAL2 | Hs.222024 | 14 |
| transmembrane protease, serine 3 | Hs.298241 | 33 |
| transmembrane, prostate androgen-induced RNA | Hs.83883 | 14 |
| trefoil factor 2 (spasmolytic protein 1) | Hs.2979 | 8 |
| tumor-associated calcium signal transducer 2 | Hs.23582 | 7,34 |
| type I intermediate filament cytokeratin | Hs.9029 | 58 |

[a] On the basis of previous identification as a highly expressed gene in our own gene expression studies (7–10, 14) or those of other investigators identified through PubMed.

Of note, one sample of normal pancreas tissue (np17) was found to have gene expression levels similar to those seen in some of the pancreatic cancers. The pathology report of this sample was reviewed and indicated that foci of pancreatic intraepithelial neoplasia, but not infiltrating carcinoma, were present within the sample.

**Verification and Localization of Selected Candidate Tumor Markers.** A number of highly expressed genes identified in this study have previously been validated as overexpressed in pancreas cancers, including *mesothelin* (9), *PSCA* (8), *Muc4* (21), *Fascin* (18), and *topoisomerase IIα* (14), consistent with our findings using PCA. Two

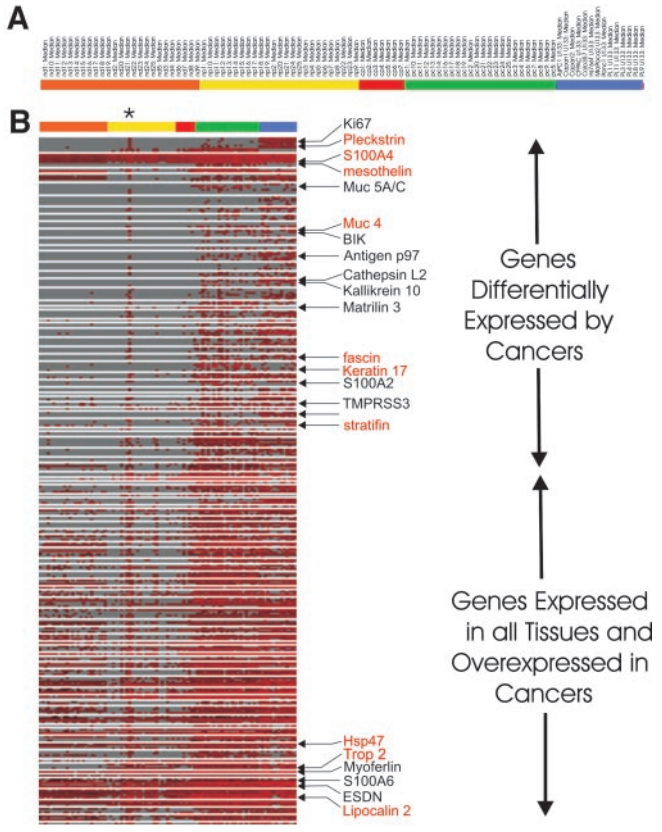


Fig. 1. Principal components analysis of highly expressed genes in pancreatic cancer. *A*, shown are 96 samples used for principal components analysis. The *multicolored horizontal bar* corresponds to 25 samples of normal duodenal mucosa (*orange bar*), 25 samples of normal pancreas (*yellow bar*), 7 samples of chronic pancreatitis (*red bar*), 25 samples of resected pancreas cancer tissue (*green bar*), and 14 pancreatic cancer cell lines (*blue bar*). *B*, principal components analysis of the 279 gene fragments identified as expressed at least 5-fold greater in pancreas cancer tissues of cell lines as compared with normal pancreas or duodenal mucosa tissues. The *colored horizontal bar* indicates the identity of the samples within the diagram. Genes previously validated as expressed in pancreas cancer are indicating by red text. Gene expression is seen to segregate into those with differential expression and those genes with overexpression as compared with normal pancreas or duodenal tissues and chronic pancreatitis. The *asterisk* indicates the presence of np17, a normal pancreas tissue specimen subsequently found to contain pancreatic intraepithelial neoplasia (PanIN).

additional genes (*Muc5A/C* and *ESDN*) were selected for verification of the expression patterns identified by PCA in normal pancreas and resected pancreatic cancer tissues using immunohistochemical labeling (Fig. 2).

*Muc5A/C* was expressed in pancreas cancer tissues and cell lines but not normal pancreas or normal duodenal tissues. Similar to a prior study (22), immunohistochemical studies for Muc5A/C protein showed labeling of the neoplastic cells in 60 of 66 (91%) pancreatic cancers. Of the 60 cancers that labeled, 40 were diffusely positive and 20 were focally positive. In all positive cases, the neoplastic epithelium showed granular cytoplasmic labeling with membranous accentuation (Fig. 2, *A* and *B*). Normal duct epithelium was negative for Muc 5A/C, with exception of scattered ducts in areas of atrophy, which showed only weak labeling.

*ESDN* was one of the genes that, although consistently expressed in normal pancreas or duodenal tissues, were expressed at higher levels in the chronic pancreatitis and pancreatic cancer samples. Immunohistochemical labeling for ESDN protein was consistent with this pattern. ESDN protein was detected in all 66 paraffin-embedded pancreas cancers (100%) analyzed, and it was expressed in both the neoplastic epithelium and desmoplastic stroma. The neoplastic epithelium labeled in a cytoplasmic distribution and the stroma labeling was noted both within fibroblasts and inflammatory cells. ESDN labeling was also noted within all samples of normal pancreas or duodenal mucosa. However, the intensity of labeling within both the neoplastic epithelium and the intervening stroma was significantly greater than that observed within the normal pancreas tissues (Fig. 2, *C* and *D*).

**Comparison of Gene Expression Patterns Identified among Different Technologies.** Of major interest would be the direct comparison of competing expression profiling technologies. We previously reported the gene expression profiles of pancreatic cancer identified by oligonucleotide arrays (14), cDNA microarrays (10), and SAGE (7–9). We chose to compare the previously generated cDNA microarray and SAGE data we have to the current findings to determine the similarities of each method in detecting robust increases in gene expression.

Among the three methods, a total of 422 unique genes or hypothetical proteins were identified (Fig. 3). A total of 263 overexpressed genes or hypothetical proteins was identified in this current analysis, 55 by our previous SAGE analysis, and 150 by our previous cDNA microarrays analysis in which we used Statistical Analysis of Microarray. Forty of these 422 genes were identified by more than one method (Table 3). Six of these 40 genes were commonly identified by all three methods, 26 of 40 were identified by both oligonucleotide arrays and cDNA microarrays only, 5 of 40 genes by both oligonucleotide arrays and SAGE only, and 3 of 40 genes by SAGE and cDNA microarrays only. Fifteen of these 40 genes have previously been reported as highly expressed in pancreas cancer (Table 3). These 15 genes comprised 3 of 6 genes common to all three methods of detection, 9 of 26 additional genes common to the U133 arrays and cDNA microarrays, 1 of 3 additional genes common to SAGE and cDNA microarrays, and 2 of 5 additional genes common to SAGE and U133 arrays. These numbers could be expected to vary, differing on the cutoff criteria applied to individual technologies and the effort devoted to additional validation efforts. Nevertheless, overexpressed genes can be identified by more than one method, and genes identifiable by more than one method were preferentially enriched in the prior literature.

## DISCUSSION

The identification of genes differentially expressed in pancreatic cancer is critical to the development of novel strategies to detect and treat this highly lethal cancer. We used the Gene Logic Inc. BioExpress platform and the most current version of Affymetrix GeneChip arrays to perform a comprehensive study of gene expression in pancreatic ductal adenocarcinomas. Gene expression levels in 39 cancers were compared with those of two forms of normal samples. We identified 143 novel overexpressed genes in pancreatic cancer. Our findings also support the overexpression of a number of genes previously reported as highly expressed in pancreatic cancer, and we attempted to stratify these differentially expressed genes to highlight into those promising for translation into clinical targets. These differentially expressed genes may form the basis for the development of diagnostic markers or screening methods to detect this disease at an earlier, potentially curable stage or may serve as novel targets for drug development or imaging.

When using gene expression profiling to understand human neoplasms, long lists of genes are produced that require further categorization into meaningful groups. In this study, we used PCA to identify the patterns of differential gene expression and therefore most promising for translation into useful clinical targets. Among the most differentially expressed genes identified by PCA were Muc4, Muc5A/C, Fascin, epiregulin, TMPRSS3, mesothelin, and topoi-

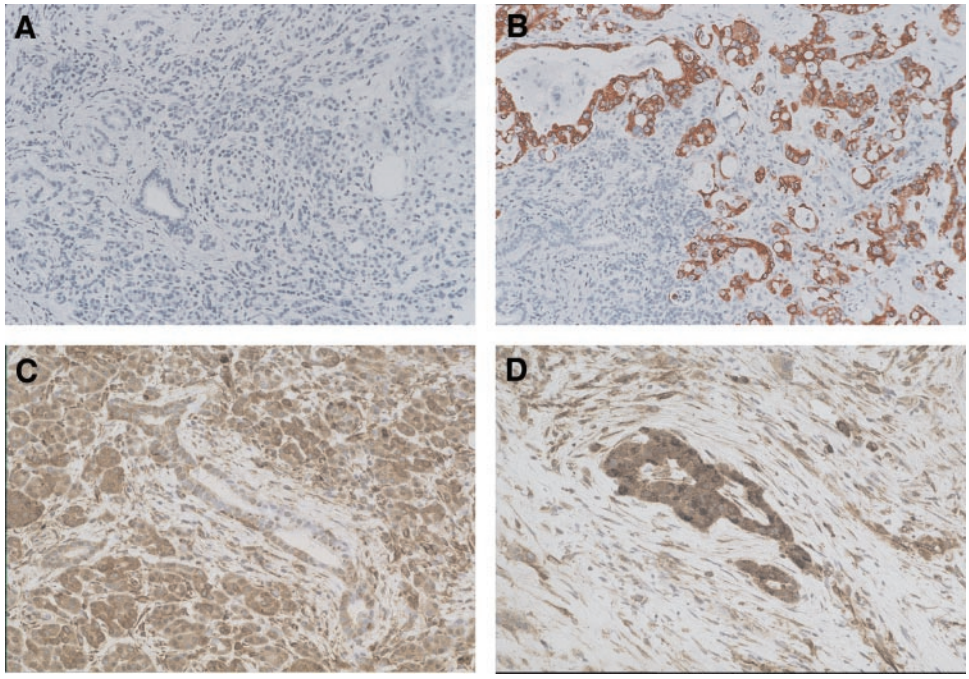


Fig. 2. Immunohistochemical verification and localization of protein expression in pancreatic cancer tissues. *A*, immunolabeling for Muc 5A/C protein in normal pancreatic duct epithelium. The pancreatic duct epithelium is negative for Muc5A/C, including the surrounding pancreatic acinar cells. *B*, Muc 5A/C protein in pancreatic duct adenocarcinomas. The infiltrating carcinoma is strongly positive (seen in the *top right*) in contrast to the normal pancreas within the same section. *C*, ESDN protein expression in normal pancreatic duct epithelium. Positive immunolabeling of the duct epithelium is seen. *D*, ESDN protein in infiltrating duct adenocarcinoma. Immunolabeling for ESDN protein shows intense labeling within the neoplastic epithelium. The surrounding stroma is also positive, although it does not label as intensely as the neoplastic epithelium.

somerase II$\alpha$. These genes have been previously identified as highly expressed in pancreatic cancers (Table 2), and there differential expression identified by PCA indicates they are among the more attractive targets for novel therapeutic strategies. Several other genes that have not been previously reported also showed differential expression among pancreatic cancer specimens, including, *tumor antigen p97*, *cadherin EGF LAG seven-pass G-type receptor 1*, *cathepsin L2*, *matrilin3*, *and kallikrein 10.*

*Tumor antigen 97* and *cadherin EGF LAG seven-pass G-type receptor 1* are both membrane-bound proteins. Tumor antigen p97, also known as melanotransferrin, is a membrane-bound transferrin homologue with several characteristics in common with serum transferrin. Tumor antigen p97 is found at high levels in melanoma cells and binds ferritin (23). *Cadherin EGF LAG seven-pass G-type receptor 1* (also known as *flamingo homologue, Drosophilia*) is a G protein-coupled receptor thought to function in contact-mediated communication between cells. Within its external domain are several cadherin-like repeats and seven EGF-like repeats, both known mediators of protein-protein interactions (24).

Cathepsin L2, matrilin 3, and kallikrein 10 are not membrane-bound. Cathepsin L2 is one of several cysteine proteinases that are involved in a variety of normal cellular processes and up-regulated in breast and colorectal cancers (25). Cysteine proteinases are synthesized as preproenzymes, which are processed to the corresponding proenzymes. The proenzymes are either targeted to the lysosome or continue along the cellular secretory route where they may become activated in the extracellular space. Kallikrein 10 and matrilin 3 also function within the extracellular space. Kallikrein 10 is a member of the kallikrein family of secreted serine proteases that include 15 proteins, including the prostate-specific antigen, also known as hK3 (26, 27). Matrilin 3 is one of a family of matrilin proteins that are components of the extracellular matrix. Matrilin 3 contains a von Willebrand factor type A-like domain thought to be necessary for interactions with collagen (28).

Overexpressed genes that are not expressed in normal tissues and that encode for membrane bound proteins (mesothelin, TMPRSS3, Muc4, Muc5A/C, or tumor antigen p97) hold the most promise for use as specific targets for antibody or cell-mediated vaccines for pancreatic cancer. For example, cell-mediated immunotherapy can be both safe and effective in generating immune responsiveness in patients with pancreatic cancer (15). Thus, each of these differentially expressed membrane bound genes represents a potential target for the development of a cell-mediated vaccine or antibody-based immunotherapy against pancreatic cancers expressing these antigens (29–32). Membrane-bound proteins such as TMPRSS3 are additionally interesting because their proteolytic activity can be exploited to activate prodrugs at the cell surface (33). In contrast, differentially expressed genes that encode for cytoplasmic or lysosomal enzymes may be more ideal for targeting by novel chemotherapeutic agents or radioimaging such as the lysosomal enzyme cathepsin L2 (25).

Seventy-five of the overexpressed genes identified in this study have been previously reported as highly expressed in pancreatic cancer (Table 2). These 75 genes validate our approach in gene marker discovery within pancreatic cancers, and these 75 genes, in

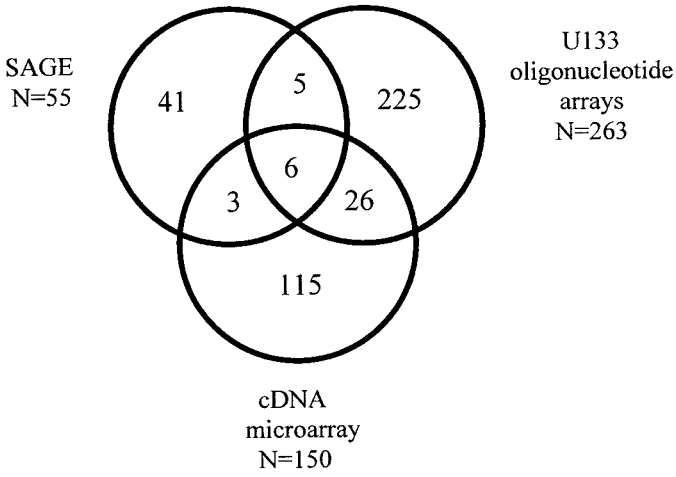


Fig. 3. Comparison of gene expression profiles obtained by various methods. Shown is a Venn diagram indicating the numbers of genes identified and reported as highly expressed by serial analysis of gene expression, cDNA microarrays, and U133 oligonucleotide arrays. Forty genes were identified by more than one technique, 6 of which were identified by all three techniques (see Table 3 for identities of these genes).

Table 3 *Highly expressed genes identified by U133 oligonucleotide arrays, cDNA microarrays, or serial analysis of gene expression (SAGE)*

| Known gene name | Also identified by other investigators | Reference |
|---|---|---|
| Genes identified by U133 arrays, cDNA microarrays, and SAGE | | |
| Keratin 19 | +[a] | 37 |
| Retinoic acid induced 3 | − | |
| Secretory leukocyte protease inhibitor | − | |
| Stratifin | + | 38 |
| Tetraspan 1 | − | |
| Transglutaminase 2 | + | 39 |
| Genes identified by both U133 arrays and cDNA microarrays | | |
| ADP-ribosylation factor-like 7 | − | |
| Annexin A1 | + | 11 |
| Capping protein, (actin filament) gelsolin like | − | |
| CEACAM5 | − | |
| Caveolin 2 | − | |
| Cdc2, $G_1$–S and $G_2$–M | + | 44 |
| Collagen, type XVII, alpha 1 | − | |
| Fer-1 (C. elegans)-like 3 (myoferlin) | + | 11 |
| Heparin sulfate (glucosamine 3-*O*-sulfotransferase 1 | − | |
| Homeobox B6 | − | |
| Immortalization-up-regulated protein | − | |
| Insulin-like growth factor binding protein 3 | − | |
| Keratin 15 | − | |
| Keratin 17 | + | 50 |
| Laminin, $\gamma$ 2 | − | |
| Lipocalin 2 (oncogene 24p3) | + | 35 |
| MCM4 | − | |
| Proteosome subunit, $\alpha$ type, 7 | + | 59 |
| RAP2B | − | |
| RuvB-like 1 | − | |
| S100A11 | + | 11 |
| S100P | − | 13 |
| sciellin | − | |
| Solute carrier family 2, member 1 | − | |
| TIMP1[b] | + | 12 |
| Ubiquitin carrier protein | − | |
| Genes identified by both SAGE and cDNA microarrays | | |
| Anterior gradient 2 | − | |
| Claudin 4 | + | 36 |
| FXYD domain-containing ion transport regulator 3 | − | |
| Genes identified by both SAGE and U133 arrays | | |
| CEACAM6 | + | 6 |
| mesothelin | − | |
| Prostate stem cell antigen | − | |
| S100A4 | − | |
| Tumor-associated calcium signal transducer 2 | + | 34 |

[a] In addition to being identified as overexpressed using one or more of the three global analyses of gene expression, these genes were also previously reported as overexpressed in pancreatic cancer using other methodologies.

[b] This gene contained greater than five tags in at least one of the two normal pancreatic duct epithelial SAGE libraries (HX and H126).

turn, are also validated as potentially important due to their repeated identification as overexpressed in pancreatic cancer by multiple platforms. We also identified 142 additional genes not previously reported as overexpressed in pancreatic cancer. The large number of genes identified as highly expressed in this study may indicate a greater sensitivity of the UG-133 GeneChip as compared with other studies but may also be accounted for by the greater number and type of specimens used or the relatively low threshold criterion (fold-change) used. Some of the genes identified in this analysis are likely attributable to expression by entrapped residual pancreatic acini or islet cells within the resected pancreatic cancer samples (*i.e.*, biliverdin reductase A or glucosidase $\beta$). Thus, it is likely that not all of the 143 genes identified will ultimately represent clinically useful targets for therapeutic development, and additional studies to evaluate the use of these genes are warranted.

The redundancy observed in published reports for many of the highly expressed genes identified in pancreatic cancers lends support to these genes playing a role in the neoplastic process in pancreatic cancers. We also found that a considerable overlap exists in the gene expression detected using three of the major global gene expression methodologies currently used in transcription profiling. Moreover, the genes commonly detected by these methods are among those that have been identified as highly expressed by a variety of independent investigators such as Trop-2 (identified by SAGE and U133 arrays; Ref. 34), myoferlin and lipocalin 2 (identified by U133 arrays and cDNA arrays; Refs. 11, 35), or claudin 4 (identified by SAGE and cDNA arrays; Ref. 36). Six genes were identified by all three methodologies (keratin 19, retinoic acid-induced 3, secretory leukocyte protease inhibitor, stratifin, tetraspan 1, and transglutaminase 2), three of which have been previously reported as highly expressed in pancreatic cancers (37–39).

One sample of normal pancreas was found to have high levels of expression of a number of genes also highly expressed in pancreatic cancers. This normal pancreas was found upon review to contain pancreatic intraepithelial neoplasia, the presumed precursor to invasive pancreatic cancer, suggesting that some of the genes identified may be up-regulated in expression early in the neoplastic progression to invasive carcinoma (40). Indeed, Muc4 was one of the genes expressed by this normal pancreas sample. Muc4 has been shown by Swartz *et al.* (21) to be expressed at high levels in pancreatic intraepithelial neoplasias and in invasive pancreatic cancers, indicating that some of these genes may have use as markers for detecting preinvasive duct lesions in addition to invasive cancers. Similarly, genes highly expressed within pancreatic intraepithelial neoplasias may provide targets for imaging of preinvasive neoplastic lesions within the pancreatic duct system. The identification of early, preinvasive lesions is of particular interest for patients with a strong family history of pancreatic cancer whose lifetime risk for developing pancreatic cancer may be increased (41).

The benefits to be derived from transcription profiling of pancreatic cancer will only be realized with the pursuit of those targets most meaningful for translation into useful clinical modalities for treating this lethal disease. Our comprehensive characterization of those genes most highly expressed within pancreas ductal adenocarcinomas provides validation of a number of genes identified as highly expressed and also provides clues as to which genes are most promising for additional evaluation.

## REFERENCES

1. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J., Gabor Miklos, G. L., Nelson, C., Broder, S., Clark, A. G., Nadeau, J., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di, F., V, Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R. R., Ke, Z., Ketchum, K. A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zhu, S., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A., Cheng, M. L., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S., Dodson, K., Doup, L., Ferriera, S., Garg, N., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hostin, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H., Reardon, M., Rodriguez, R., Rogers, Y. H., Romblad, D., Ruhfel, B., Scott, R., Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint, N. N., Tse, S., Vech, C., Wang, G., Wetter, J., Williams, S., Williams, M., Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J. F., Guigo, R., Campbell, M. J., Sjolander, K. V., Karlak, B., Kejariwal, A., Mi, H., Lazareva, B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S., Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yooseph, S.,

Allen, D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y. H., Coyne, M., Dahlke, C., Mays, A., Dombroski, M., Donnelly, M., Ely, D., Esparham, S., Fosler, C., Gire, H., Glanowski, S., Glasser, K., Glodek, A., Gorokhov, M., Graham, K., Gropman, B., Harris, M., Heil, J., Henderson, S., Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C., Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D., Majoros, W., McDaniel, J., Murphy, S., Newman, M., Nguyen, T., Nguyen, N., and Nodell, M. The sequence of the human genome. Science, *291:* 1304–1351, 2001.

2. Olivier, M., Aggarwal, A., Allen, J., Almendras, A. A., Bajorek, E. S., Beasley, E. M., Brady, S. D., Bushard, J. M., Bustos, V. I., Chu, A., Chung, T. R., De Witte, A., Denys, M. E., Dominguez, R., Fang, N. Y., Foster, B. D., Freudenberg, R. W., Hadley, D., Hamilton, L. R., Jeffrey, T. J., Kelly, L., Lazzeroni, L., Levy, M. R., Lewis, S. C., Liu, X., Lopez, F. J., Louie, B., Marquis, J. P., Martinez, R. A., Matsuura, M. K., Misherghi, N. S., Norton, J. A., Olshen, A., Perkins, S. M., Perou, A. J., Piercy, C., Piercy, M., Qin, F., Reif, T., Sheppard, K., Shokoohi, V., Smick, G. A., Sun, W. L., Stewart, E. A., Fernando, J., Tejeda, Tran, N. M., Trejo, T., Vo, N. T., Yan, S. C., Zierten, D. L., Zhao, S., Sachidanandam, R., Trask, B. J., Myers, R. M., and Cox, D. R. A high-resolution radiation hybrid map of the human genome draft sequence. Science, *291:* 1298–1302, 2001.
3. Wallrapp, C., and Gress, T. M. Isolation of differentially expressed genes by representational difference analysis. Methods Mol. Biol., *175:* 279–294, 2001.
4. Gress, T. M., Wallrapp, C., Frohme, M., Muller-Pillasch, F., Lacher, U., Friess, H., Buchler, M., Adler, G., and Hoheisel, J. D. Identification of genes with specific expression in pancreatic cancer by cDNA representational difference analysis. Genes Chromosomes Cancer, *19:* 97–103, 1997.
5. Paciucci, R., Berrozpe, G., Tora, M., Navarro, E., Garcia, d. H., and Real, F. X. Isolation of tissue-type plasminogen activator, cathepsin H, and non-specific cross-reacting antigen from SK-PC-1 pancreas cancer cells using subtractive hybridization. FEBS Lett., *385:* 72–76, 1996.
6. Gardner-Thorpe, J., Ito, H., Ashley, S. W., and Whang, E. E. Differential display of expressed genes in pancreatic cancer cells. Biochem. Biophys. Res Commun., *293:* 391–395, 2002.
7. Ryu, B., Jones, J., Blades, N. J., Parmigiani, G., Hollingsworth, M. A., Hruban, R. H., and Kern, S. E. Relationships and differentially expressed genes among pancreatic cancers examined by large-scale serial analysis of gene expression. Cancer Res., *62:* 819–826, 2002.
8. 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. 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, 2001.
9. 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. 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, 2001.
10. 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. Exploration of global gene expression patterns in pancreatic adenocarcinoma using cDNA microarrays. Am. J. Pathol., *162:* 1151–1162, 2003.
11. Han, H., Bearss, D. J., Browne, L. W., Calaluce, R., Nagle, R. B., and Von Hoff, D. D. Identification of differentially expressed genes in pancreatic cancer cells using cDNA microarray. Cancer Res., *62:* 2890–2896, 2002.
12. Crnogorac-Jurcevic, T., Efthimiou, E., Nielsen, T., Loader, J., Terris, B., Stamp, G., Baron, A., Scarpa, A., and Lemoine, N. R. Expression profiling of microdissected pancreatic adenocarcinomas. Oncogene, *21:* 4587–4594, 2002.
13. Crnogorac-Jurcevic, T., Missiaglia, E., Blaveri, E., Gangeswaran, R., Jones, M., Terris, B., Costello, E., Neoptolemos, J. P., and Lemoine, N. R. Molecular alterations in pancreatic carcinoma: expression profiling shows that dysregulated expression of S100 genes is highly prevalent. J. Pathol., *201:* 63–74, 2003.
14. 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. Discovery of novel tumor markers of pancreatic cancer using global gene expression technology. Am. J. Pathol, *160:* 1239–1249, 2002.
15. Jaffee, E. M., Hruban, R. H., Biedrzycki, B., Laheru, D., Schepers, K., Sauter, P. R., Goemann, M., Coleman, J., Grochow, L., Donehower, R. C., Lillemoe, K. D., O'Reilly, S., Abrams, R. A., Pardoll, D. M., Cameron, J. L., and Yeo, C. J. Novel allogeneic granulocyte-macrophage colony-stimulating factor-secreting tumor vaccine for pancreatic cancer: a Phase I trial of safety and immune activation. J. Clin. Oncol., *19:* 145–156, 2001.
16. 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. USA, *95:* 14863–14868, 1998.
17. Kolar, C., Caffrey, T., Hollingsworth, M., Scheetz, M., Sutherlin, M., Weide, L., and Lawson, T. Duct epithelial cells cultured from human pancreas processed for transplantation retain differentiated ductal characteristics. Pancreas, *15:* 265–271, 1997.
18. Maitra, A., Iacobuzio-Donahue, C., Rahman, A., Sohn, T. A., Argani, P., Meyer, R., Yeo, C. J., Cameron, J. L., Goggins, M., Kern, S. E., Ashfaq, R., Hruban, R. H., and Wilentz, R. E. Immunohistochemical validation of a novel epithelial and a novel stromal marker of pancreatic ductal adenocarcinoma identified by global expression microarrays: sea urchin fascin homolog and heat shock protein 47. Am. J. Clin. Pathol., *118:* 52–59, 2002.
19. Iacobuzio-Donahue, C. A., Ryu, B., Hruban, R. H., and Kern, S. E. Exploring the host desmoplastic response to pancreatic carcinoma: gene expression of stromal and neoplastic cells at the site of primary invasion. Am. J. Pathol., *160:* 91–99, 2002.
20. Iacobuzio-Donahue, C. A., Argani Pedram, Hempen, P., Jones, J., and Kern, S. E. The desmoplastic response to infiltrating breast carcinoma: gene expression at the site of primary invasion with implications for comparisons between tumor types. Cancer Res., *62:* 5351–5357, 2002.
21. Swartz, M. J., Batra, S. K., Varshney, G. C., Hollingsworth, M. A., Yeo, C. J., Cameron, J. L., Wilentz, R. E., Hruban, R. H., and Argani, P. MUC4 expression increases progressively in pancreatic intraepithelial neoplasia. Am. J. Clin. Pathol., *117:* 791–796, 2002.
22. Balague, C., Gambus, G., Carrato, C., Porchet, N., Aubert, J. P., Kim, Y. S., and Real, F. X. Altered expression of MUC2, MUC4, and MUC5 mucin genes in pancreas tissues and cancer cell lines. Gastroenterology, *106:* 1054–1061, 1994.
23. Richardson, D. R. The role of the membrane-bound tumour antigen, melanotransferrin (p97), in iron uptake by the human malignant melanoma cell. Eur. J. Biochem., *267:* 1290–1298, 2000.
24. Wu, Q., and Maniatis, T. A striking organization of a large family of human neural cadherin-like cell adhesion genes. Cell, *97:* 779–790, 1999.
25. Santamaria, I., Velasco, G., Cazorla, M., Fueyo, A., Campo, E., and Lopez-Otin, C. Cathepsin L2, a novel human cysteine proteinase produced by breast and colorectal carcinomas. Cancer Res., *58:* 1624–1630, 1998.
26. Luo, L. Y., Katsaros, D., Scorilas, A., Fracchioli, S., Bellino, R., van Gramberen, M., de Bruijn, H., Henrik, A., Stenman, U. H., Massobrio, M., van der Zee, A. G., Vergote, I., and Diamandis, E. P. The serum concentration of human kallikrein 10 represents a novel biomarker for ovarian cancer diagnosis and prognosis. Cancer Res., *63:* 807–811, 2003.
27. Petraki, C. D., Karavana, V. N., Luo, L. Y., and Diamandis, E. P. Human kallikrein 10 expression in normal tissues by immunohistochemistry. J. Histochem. Cytochem., *50:* 1247–1261, 2002.
28. Klatt, A. R., Nitsche, D. P., Kobbe, B., Morgelin, M., Paulsson, M., and Wagener, R. Molecular structure and tissue distribution of matrilin-3, a filament-forming extracellular matrix protein expressed during skeletal development. J. Biol. Chem., *275:* 3999–4006, 2000.
29. Kerr, D. E., Vrudhula, V. M., Svensson, H. P., Siemers, N. O., and Senter, P. D. Comparison of recombinant and synthetically formed monoclonal antibody-$\beta$-lactamase conjugates for anticancer prodrug activation. Bioconjug. Chem., *10:* 1084–1089, 1999.
30. Hassan, R., Viner, J. L., Wang, Q. C., Margulies, I., Kreitman, R. J., and Pastan, I. Anti-tumor activity of K1-LysPE38QQR, an immunotoxin targeting mesothelin, a cell-surface antigen overexpressed in ovarian cancer and malignant mesothelioma. J. Immunother., *23:* 473–479, 2000.
31. Saffran, D. C., Raitano, A. B., Hubert, R. S., Witte, O. N., Reiter, R. E., and Jakobovits, A. Anti-PSCA mAbs inhibit tumor growth and metastasis formation and prolong the survival of mice bearing human prostate cancer xenografts. Proc. Natl. Acad. Sci. USA, *98:* 2658–2663, 2001.
32. Ross, S., Spencer, S. D., Strickland, L., Lasky, L. A., and Koeppen, H. Prostate stem cell antigen (PSCA) is a potential therapy target for pancreatic cancer. U. S. Cancer Acad. Pathol., 2002.
33. Wallrapp, C., Hahnel, S., Muller-Pillasch, F., Burghardt, B., Iwamura, T., Ruthenburger, M., Lerch, M. M., Adler, G., and Gress, T. M. A novel transmembrane serine protease (TMPRSS3) overexpressed in pancreatic cancer. Cancer Res., *60:* 2602–2606, 2000.
34. Linnenbach, A. J., Wojcierowski, J., Wu, S. A., Pyrc, J. J., Ross, A. H., Dietzschold, B., Speicher, D., and Koprowski, H. Sequence investigation of the major gastrointestinal tumor-associated antigen gene family, GA733. Proc. Natl. Acad. Sci. USA, *86:* 27–31, 1989.
35. Furutani, M., Arii, S., Mizumoto, M., Kato, M., and Imamura, M. Identification of a neutrophil gelatinase-associated lipocalin mRNA in human pancreatic cancers using a modified signal sequence trap method. Cancer Lett., *122:* 209–214, 1998.
36. Michl, P., Buchholz, M., Rolke, M., Kunsch, S., Lohr, M., McClane, B., Tsukita, S., Leder, G., Adler, G., and Gress, T. M. Claudin-4: a new target for pancreatic cancer treatment using Clostridium perfringens enterotoxin. Gastroenterology, *121:* 678–684, 2001.
37. Real, F. X., Villa, M. R., Skoudy, A., Ramaekers, F. C. S., and Corominas, J. M. Intermediate filaments as differentiation markers of exocrine pancreas. II. Expression of cytokeratins of complex and stratified epithelia in normal pancreas and in pancreas cancer. Int. J. Cancer, *54:* 720–727, 1993.
38. Sinha, P., Hutter, G., Kottgen, E., Dietel, M., Schadendorf, D., and Lage, H. Increased expression of epidermal fatty acid binding protein, cofilin, and 14-3-3-$\sigma$ (stratifin) detected by two-dimensional gel electrophoresis, mass spectrometry and microsequencing of drug-resistant human adenocarcinoma of the pancreas. Electrophoresis, *20:* 2952–2960, 1999.
39. Elsasser, H. P., MacDonald, R., Dienst, M., and Kern, H. F. Characterization of a transglutaminase expressed in human pancreatic adenocarcinoma cells. Eur. J. Cell Biol., *61:* 321–328, 1993.
40. Maitra, A., Adsay, N. V., Argani, P., Iacobuzio-Donahue, C., De Marzo, A., Cameron, J. L., Yeo, C. J., and Hruban, R. H. Multicomponent analysis of the pancreatic adenocarcinoma progression model using a pancreatic intraepithelial neoplasia tissue microarray. Mod. Pathol., *16:* 902–912, 2003.
41. Tersmette, A. C., Petersen, G. M., Offerhaus, G. J., Falatko, F. C., Brune, K. A., Goggins, M., Rozenblum, E., Wilentz, R. E., Yeo, C. J., Cameron, J. L., Kern, S. E., and Hruban, R. H. Increased risk of incident pancreatic cancer among first-degree relatives of patients with familial pancreatic cancer. Clin. Cancer Res., *7:* 738–744, 2001.

42. Joo, Y. E., Rew, J. S., Park, C. S., and Kim, S. J. Expression of E-cadherin, α- and β-catenins in patients with pancreatic adenocarcinoma. Pancreatology, *2:* 129–137, 2002.
43. Yamaguchi, N., Chung, S. M., Shiroeda, O., Koyama, K., and Imanishi, J. Characterization of a cathepsin L-like enzyme secreted from human pancreatic cancer cell line HPC-YP. Cancer Res., *50:* 658–663, 1990.
44. Ito, Y., Takeda, T., Wakasa, K., Tsujimoto, M., Okada, M., and Matsuura, N. Expression of the $G_2$-M modulators in pancreatic adenocarcinoma. Pancreatology, *2:* 138–145, 2002.
45. Yang, T., and Chen, J. Y. Identification and cellular localization of human PFTAIRE1. Gene (Amst.), *267:* 165–172, 2001.
46. Poch, B., Gansauge, F., Schwarz, A., Seufferlein, T., Schnelldorfer, T., Ramadani, M., Beger, H. G., and Gansauge, S. Epidermal growth factor induces cyclin D1 in human pancreatic carcinoma: evidence for a cyclin D1-dependent cell cycle progression. Pancreas, *23:* 280–287, 2001.
47. Zhu, Z., Kleeff, J., Friess, H., Wang, L., Zimmermann, A., Yarden, Y., Buchler, M. W., and Korc, M. Epiregulin is up-regulated in pancreatic cancer and stimulates pancreatic cancer cell growth. Biochem. Biophys. Res. Commun., *273:* 1019–1024, 2000.
48. Kleeff, J., Ishiwata, T., Friess, H., Buchler, M. W., and Korc, M. Concomitant over-expression of activin/inhibin β subunits and their receptors in human pancreatic cancer. Int. J. Cancer, *77:* 860–868, 1998.
49. Shi, Q., Abbruzzese, J. L., Huang, S., Fidler, I. J., Xiong, Q., and Xie, K. Constitutive and inducible interleukin 8 expression by hypoxia and acidosis renders human pancreatic cancer cells more tumorigenic and metastatic. Clin. Cancer Res., *5:* 3711–3721, 1999.
50. Goldstein, N. S., and Bassi, D. Cytokeratins 7, 17, and 20 reactivity in pancreatic and ampulla of vater adenocarcinomas. Percentage of positivity and distribution is affected by the cut-point threshold. Am. J. Clin. Pathol., *115:* 695–702, 2001.
51. Berberat, P. O., Friess, H., Wang, L., Zhu, Z., Bley, T., Frigeri, L., Zimmermann, A., and Buchler, M. W. Comparative analysis of galectins in primary tumors and tumor metastasis in human pancreatic cancer. J. Histochem. Cytochem., *49:* 539–549, 2001.
52. von Marschall, Z., Riecken, E. O., and Rosewicz, S. Stromelysin 3 is overexpressed in human pancreatic carcinoma and regulated by retinoic acid in pancreatic carcinoma cell lines. Gut, *43:* 692–698, 1998.
53. Yamamoto, H., Itoh, F., Iku, S., Adachi, Y., Fukushima, H., Sasaki, S., Mukaiya, M., Hirata, K., and Imai, K. Expression of matrix metalloproteinases and tissue inhibitors of metalloproteinases in human pancreatic adenocarcinomas: clinicopathologic and prognostic significance of matrilysin expression. J. Clin. Oncol., *19:* 1118–1127, 2001.
54. Tomioka, D., Maehara, N., Kuba, K., Mizumoto, K., Tanaka, M., Matsumoto, K., and Nakamura, T. Inhibition of growth, invasion, and metastasis of human pancreatic carcinoma cells by NK4 in an orthotopic mouse model. Cancer Res., *61:* 7518–7524, 2001.
55. Ellenrieder, V., Hendler, S. F., Ruhland, C., Boeck, W., Adler, G., and Gress, T. M. TGF-β-induced invasiveness of pancreatic cancer cells is mediated by matrix metalloproteinase-2 and the urokinase plasminogen activator system. Int. J. Cancer, *93:* 204–211, 2001.
56. Paciucci, R., Vila, M. R., Adell, T., Diaz, V. M., Tora, M., Nakamura, T., and Real, F. X. Activation of the urokinase plasminogen activator/urokinase plasminogen activator receptor system and redistribution of E-cadherin are associated with hepatocyte growth factor-induced motility of pancreas tumor cells overexpressing Met. Am. J. Pathol., *153:* 201–212, 1998.
57. Rosty, C., Ueki, T., Argani, P., Jansen, M., Yeo, C. J., Cameron, J. L., Hruban, R. H., and Goggins, M. Overexpression of S100A4 in pancreatic ductal adenocarcinomas is associated with poor differentiation and DNA hypomethylation. Am. J. Pathol., *160:* 45–50, 2002.
58. Zhang, J. S., Wang, L., Huang, H., Nelson, M., and Smith, D. I. Keratin 23 (K23), a novel acidic keratin, is highly induced by histone deacetylase inhibitors during differentiation of pancreatic cancer cells. Genes Chromosomes Cancer, *30:* 123–135, 2001.
59. Ohnami, S., Matsumoto, N., Nakano, M., Aoki, K., Nagasaki, K., Sugimura, T., Terada, M., and Yoshida, T. Identification of genes showing differential expression in antisense K-ras-transduced pancreatic cancer cells with suppressed tumorigenicity. Cancer Res., *59:* 5565–5571, 1999. and Hruban, R. H. Increased risk of incident pancreatic cancer among first-degree relatives of patients with familial pancreatic cancer. Clin. Cancer Res., *7:* 738–744, 2001.



42. Joo, Y. E., Rew, J. S., Park, C. S., and Kim, S. J. Expression of E-cadherin, α- and β-catenins in patients with pancreatic adenocarcinoma. Pancreatology, *2:* 129–137, 2002.
43. Yamaguchi, N., Chung, S. M., Shiroeda, O., Koyama, K., and Imanishi, J. Characterization of a cathepsin L-like enzyme secreted from human pancreatic cancer cell line HPC-YP. Cancer Res., *50:* 658–663, 1990.
44. Ito, Y., Takeda, T., Wakasa, K., Tsujimoto, M., Okada, M., and Matsuura, N. Expression of the $G_2$-M modulators in pancreatic adenocarcinoma. Pancreatology, *2:* 138–145, 2002.
45. Yang, T., and Chen, J. Y. Identification and cellular localization of human PFTAIRE1. Gene (Amst.), *267:* 165–172, 2001.
46. Poch, B., Gansauge, F., Schwarz, A., Seufferlein, T., Schnelldorfer, T., Ramadani, M., Beger, H. G., and Gansauge, S. Epidermal growth factor induces cyclin D1 in human pancreatic carcinoma: evidence for a cyclin D1-dependent cell cycle progression. Pancreas, *23:* 280–287, 2001.
47. Zhu, Z., Kleeff, J., Friess, H., Wang, L., Zimmermann, A., Yarden, Y., Buchler, M. W., and Korc, M. Epiregulin is up-regulated in pancreatic cancer and stimulates pancreatic cancer cell growth. Biochem. Biophys. Res. Commun., *273:* 1019–1024, 2000.
48. Kleeff, J., Ishiwata, T., Friess, H., Buchler, M. W., and Korc, M. Concomitant over-expression of activin/inhibin β subunits and their receptors in human pancreatic cancer. Int. J. Cancer, *77:* 860–868, 1998.
49. Shi, Q., Abbruzzese, J. L., Huang, S., Fidler, I. J., Xiong, Q., and Xie, K. Constitutive and inducible interleukin 8 expression by hypoxia and acidosis renders human pancreatic cancer cells more tumorigenic and metastatic. Clin. Cancer Res., *5:* 3711–3721, 1999.
50. Goldstein, N. S., and Bassi, D. Cytokeratins 7, 17, and 20 reactivity in pancreatic and ampulla of vater adenocarcinomas. Percentage of positivity and distribution is affected by the cut-point threshold. Am. J. Clin. Pathol., *115:* 695–702, 2001.
51. Berberat, P. O., Friess, H., Wang, L., Zhu, Z., Bley, T., Frigeri, L., Zimmermann, A., and Buchler, M. W. Comparative analysis of galectins in primary tumors and tumor metastasis in human pancreatic cancer. J. Histochem. Cytochem., *49:* 539–549, 2001.
52. von Marschall, Z., Riecken, E. O., and Rosewicz, S. Stromelysin 3 is overexpressed in human pancreatic carcinoma and regulated by retinoic acid in pancreatic carcinoma cell lines. Gut, *43:* 692–698, 1998.
53. Yamamoto, H., Itoh, F., Iku, S., Adachi, Y., Fukushima, H., Sasaki, S., Mukaiya, M., Hirata, K., and Imai, K. Expression of matrix metalloproteinases and tissue inhibitors of metalloproteinases in human pancreatic adenocarcinomas: clinicopathologic and prognostic significance of matrilysin expression. J. Clin. Oncol., *19:* 1118–1127, 2001.
54. Tomioka, D., Maehara, N., Kuba, K., Mizumoto, K., Tanaka, M., Matsumoto, K., and Nakamura, T. Inhibition of growth, invasion, and metastasis of human pancreatic carcinoma cells by NK4 in an orthotopic mouse model. Cancer Res., *61:* 7518–7524, 2001.
55. Ellenrieder, V., Hendler, S. F., Ruhland, C., Boeck, W., Adler, G., and Gress, T. M. TGF-β-induced invasiveness of pancreatic cancer cells is mediated by matrix metalloproteinase-2 and the urokinase plasminogen activator system. Int. J. Cancer, *93:* 204–211, 2001.
56. Paciucci, R., Vila, M. R., Adell, T., Diaz, V. M., Tora, M., Nakamura, T., and Real, F. X. Activation of the urokinase plasminogen activator/urokinase plasminogen activator receptor system and redistribution of E-cadherin are associated with hepatocyte growth factor-induced motility of pancreas tumor cells overexpressing Met. Am. J. Pathol., *153:* 201–212, 1998.
57. Rosty, C., Ueki, T., Argani, P., Jansen, M., Yeo, C. J., Cameron, J. L., Hruban, R. H., and Goggins, M. Overexpression of S100A4 in pancreatic ductal adenocarcinomas is associated with poor differentiation and DNA hypomethylation. Am. J. Pathol., *160:* 45–50, 2002.
58. Zhang, J. S., Wang, L., Huang, H., Nelson, M., and Smith, D. I. Keratin 23 (K23), a novel acidic keratin, is highly induced by histone deacetylase inhibitors during differentiation of pancreatic cancer cells. Genes Chromosomes Cancer, *30:* 123–135, 2001.
59. Ohnami, S., Matsumoto, N., Nakano, M., Aoki, K., Nagasaki, K., Sugimura, T., Terada, M., and Yoshida, T. Identification of genes showing differential expression in antisense K-ras-transduced pancreatic cancer cells with suppressed tumorigenicity. Cancer Res., *59:* 5565–5571, 1999.