Published OnlineFirst December 8, 2010; DOI:10.1158/1055-9965.EPI-10-0397

Liquid chromatography/mass spectrometry

The HFM-treated samples were digested with sequencing grade-modified trypsin (Promega) and analyzed in duplicate using a nano-flow high-performance liquid chromatography (HPLC; NanoFrontier nLC, Hitachi High-technologies) connected to an electrospray ionization quadrupole time-of-flight (ESI-Q-TOF) mass spectrometer (Q-Tof Ultima, Waters).

MS peaks were detected, normalized, and quantified using the in-house 2DICAL software package, as described previously (17). A serial identification (ID) number was applied to each of the MS peaks detected (1 to 53,009). The stability of LC-MS was monitored by calculating the correlation coefficient (CC) and coefficient of variance (CV) of every measurement. The mean CC ± SD and CV ± SD for all 53,009 peaks observed in the 45 duplicate runs were as high as 0.946 ± 0.042 and as low as 0.053 ± 0.010, respectively.

Protein identification by tandem MS (MS/MS)

Peak lists were generated using the Mass Navigator software package (version 1.2; Mitsui Knowledge Industry) and searched against the SwissProt database (downloaded on April 22, 2009) using the Mascot software package (version 2.2.1; Matrix Science). The search parameters used were as follows. A database of human proteins was selected. Trypsin was designated as the enzyme, and up to 1 missed cleavage was allowed. Mass tolerances for precursor and fragment ions were ± 0.6 Da and ± 0.2 Da, respectively. The score threshold was set to $P < 0.05$ based on the size of the database used in the search. If a peptide was matched to multiple proteins, the protein name with the highest Mascot score was selected.

Western blot analysis

Primary antibodies used were a rabbit polyclonal antibody against platelet basic protein (PBP) precursor (Sigma) and a mouse monoclonal antibody against human complement C3b-α (PROGEN). The anti-PBP antibody recognizes all the known cleaved forms of PBP including CTAP-III and NAP-2. Six microliters of 1:10 diluted plasma sample was separated by SDS-PAGE and electroblotted onto a polyvinylidene difluoride (PVDF) membrane. The membrane was then incubated with the primary antibody and subsequently with the relevant horseradish peroxidase (HRP)-conjugated anti-rabbit or anti-mouse IgG, as described previously (26, 27). Blots were developed using an enhanced chemiluminescence (ECL) detection system (GE Healthcare).

Reverse-phase protein microarray

The plasma samples were passed through IgY microbeads (Seppro-IgY12, Sigma-Aldrich) using an automated Magtration System SA-1 (Precision System Science) in accordance with the manufacturer's instructions to reduce the 12 most abundant plasma proteins. The flow-through portion was serially diluted 1:50, 1:100, 1:200, and 1:400 using a Biomek 2000 Laboratory Automation Robot (Beckman Coulter) and randomly plotted onto ProteoChip glass slides (Proteogen) in quadruplicate in a 6,144-spot/slide format using a Protein Microarrayer Robot (Kaken Geneqs). The spotted slides were incubated overnight with the anti-PBP precursor antibody and then with biotinylated anti-rabbit IgG (Vector Laboratories) and subsequently with streptavidin-HRP conjugate (GE Healthcare). The peroxidase activity was detected using the Tyramide Signal Amplification (TSA) Cyanine 5 System (PerkinElmer). The slides were counterstained with Alexa Fluor 546-labeled goat anti-human IgG (Invitrogen; spotting control).

The stained slides were scanned on a microarray scanner (InnoScan 700AL; Innopsys). Fluorescence intensity, determined as the mean value of quadruplicate samples, was determined using the Mapix software package (Innopsys). All determined intensity values were transformed into logarithmic variables.

The reproducibility of the reverse-phase protein microarray assay was determined by repeating the same experiment, as reported previously (28). A plasma sample after reduction of the 12 most abundant plasma proteins was serially diluted within a range of 25- to 6,400-fold. Each diluted sample was spotted in quadruplicate onto glass slides and blotted with the anti-PBP antibody. In a representative quality control experiment, the CC value was 0.980 between days and the median CV was 0.047 among the quadruplicates.

Multiplex assay

The levels of CXCL7 in plasma samples were measured using a Milliplex Human Cytokine/Chemokine panel III kit (Millipore) in accordance with the manufacturer's instructions.

Statistical analysis

Statistical significance of intergroup differences was assessed with the Wilcoxon test, Mann–Whitney U test, Welch's t test, or Fisher's exact test, as appropriate. The area under the curve (AUC) of the receiver-operating characteristic (ROC) was calculated for each marker to evaluate its diagnostic significance. A composite index of 2 markers was generated using the results of multivariate logistic regression analysis, which also enabled the calculation of sensitivity, specificity, and the ROC curve. Statistical analyses were performed using an open-source statistical language R (version 2.7.0) with the optional module Design package.

Results

Plasma proteins associated with pancreatic cancer

A plasma sample from 1 healthy volunteer was processed 3 times using the HFM filtration technique. The concentration of β2-microglobulin before and after HFM treatment was measured. The recovery rates were

Published OnlineFirst December 8, 2010; DOI:10.1158/1055-9965.EPI-10-0397

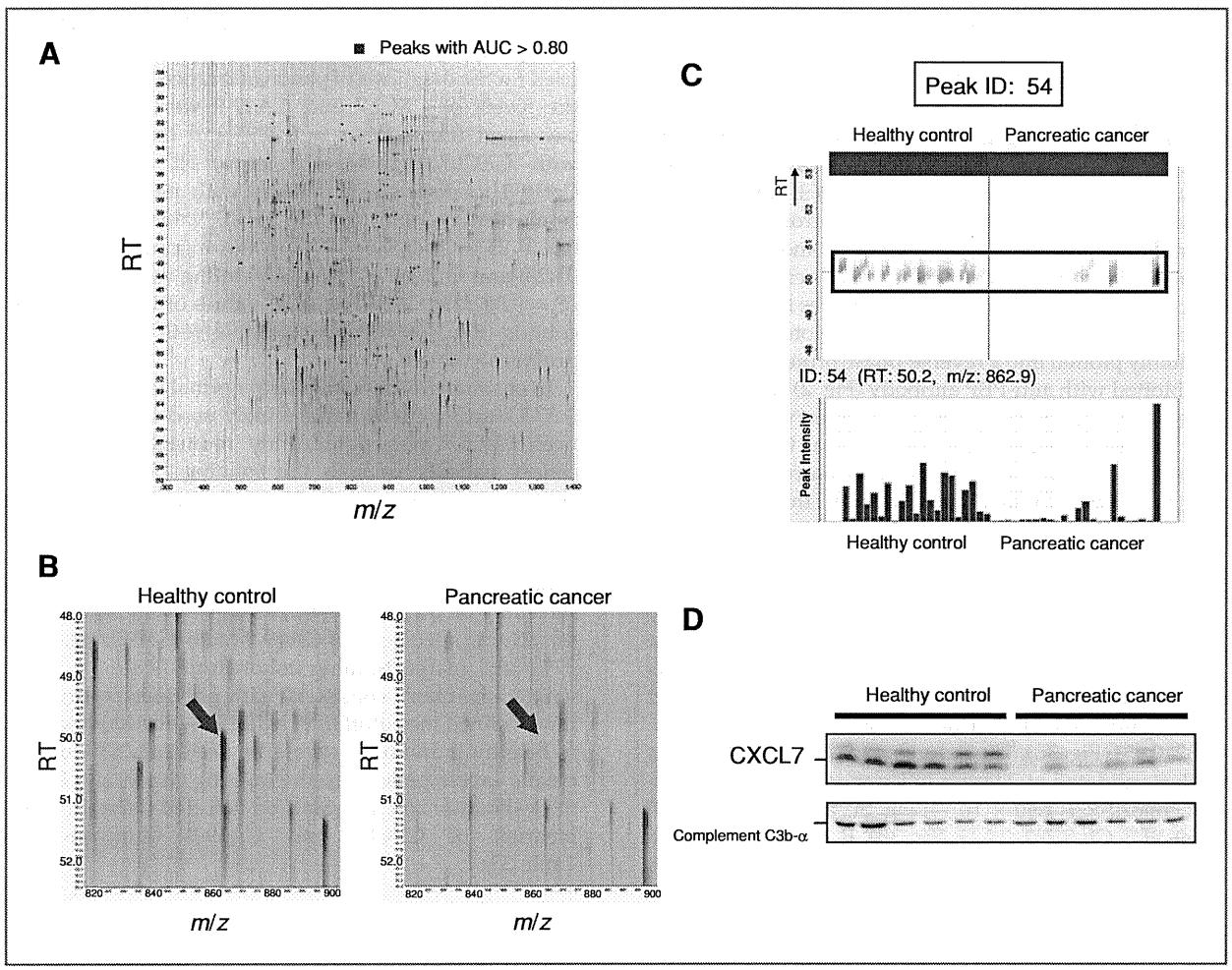


Figure 1. A, 2-dimensional display of all (>53,000) the MS peaks with *m/z* values along the *x*-axis and RT of LC along the *y*-axis. The peaks are displayed with a bin size of 1.0 *m/z*. The 140 MS peaks whose mean intensity of duplicates that distinguished pancreatic cancer patients from healthy controls with AUC values of greater than 0.800 are highlighted in red. B, CXCL7-derived MS peak (ID 54, at 863 *m/z* and 50.2 minutes) in representative patients from the cancer and control groups. C, CXCL7-derived MS peak (ID 54) in 45 duplicate LC-MS runs [patients with pancreatic cancer (red) and healthy controls (blue)] aligned along the RT of LC. Columns represent the mean intensity of duplicates (bottom). D, detection of CXCL7 and complement C3b-α (loading control) by Western blotting. Multiple bands for CXCL7 indicate the presence of proteolytic products.

25.11%, 25.73%, and 29.16%, respectively. Although the rates were seemingly low, the HFM treatment was highly reproducible with a CV of 0.081 and the amount of β2-microglobulin relative to total protein was increased 150 to 200-fold after HFM treatment.

To identify a diagnostic biomarker for pancreatic cancer, we compared the plasma LMW proteome between 24 patients with pancreatic cancer and 21 healthy controls (training cohort) using 2DICAL. Among a total of 53,009 independent MS peaks detected within the range 250 to 1,600 m/z and within a time range of 20 to 70 minutes, we found that 140 peaks had discriminatory ability with a AUC of above 0.800. Figure 1A is a representative 2-dimensional view of all the MS peaks displayed with m/z along the x-axis and the retention time (RT) of LC along the y-axis.

Twenty-five MS/MS spectra acquired from those 140 peaks were recurrently matched to 10 proteins in the database with a Mascot score of greater than 30 (Supplementary Table S1). Notably, one MS peak (ID 54) matched the amino acid sequence of the *CXCL7* gene product (Swiss-Prot_P02775) with the highest score of 99.6 (Supplementary Fig. S2). Figure 1B shows the CXCL7-derived MS peak (ID 54, at 863 *m/z* and 50.2 minutes) that appeared in a representative pancreatic cancer patient and a healthy individual. Figure 1C demonstrates the distribution of the MS peak (ID 54) in patients with pancreatic cancer (red) and healthy controls (blue) in the training cohort (AUC = 0.839; $P = 4.54 \times 10^{-5}$; Mann–Whitney U test). The differential expression and identification of CXCL7 was confirmed by immunoblotting (Fig. 1D).

Published OnlineFirst December 8, 2010; DOI:10.1158/1055-9965.EPI-10-0397

Validation of reduced CXCL7 in pancreatic cancer patients

The level of plasma CXCL7 was quantified in 12 patients with pancreatic cancer and 12 healthy individuals in the training cohort using multiplex assay. Consistent with 2DICAL, CXCL7 was found to be significantly decreased in patients with pancreatic cancer (mean ± SD, 744 ± 182 ng/mL) in comparison with healthy controls (1,355 ± 386 ng/mL; $P = 0.0003$). To further verify and validate the reduction of plasma CXCL7 in pancreatic cancer patients, 280 plasma samples [including 43 samples from the training cohort and new 237 samples (validation cohort)] were randomly plotted into a reverse-phase protein microarray and blotted with anti-PBP antibody (Fig. 2). Two samples from healthy controls in the training cohort were excluded due to an insufficient sample volume. Quadruplicate spots for representative cases and controls with high and low levels of CXCL7 are shown in the right panels of Figure 2.

The results of reverse-phase protein microarray were well correlated with those of multiplex assay (Pearson's $r = 0.65$; $P = 0.0006$; Supplementary Fig. S3). Microarray analysis also showed a significant reduction of the plasma CXCL7 level in the pancreatic cancer patients of the training cohort ($P = 5.96 \times 10^{-5}$; Welch's t test; Fig. 3A and Table 1) with an AUC value of 0.872 (95% CI: 0.732–0.951; Fig. 3B). The reduction of plasma CXCL7 was further validated in a larger independent cohort (validation cohort; $P = 1.40 \times 10^{-16}$; Fig. 3C and AUC value of 0.850, 95% CI: 0.792–0.895; Fig. 3B). Because there was a difference in age distribution between the cancer patients and healthy controls of the validation cohort (Table 1), we performed a subgroup analysis of 79 pancreatic cancer patients (median age, 61) and 20 healthy controls (median age, 60) aged 50 to 70 years. The reduction of plasma CXCL7 in patients with pancreatic cancer was statistically significant even in this subgroup ($P = 0.0001$), indicating that the decrease of the CXCL7 level was not merely due to the difference of age distribution between the pancreatic cancer patients and controls.

CXCL7 was significantly reduced in patients with any stage of pancreatic cancer (Table 2), including those with stage I (<0.001) and II (<0.001) disease. The significant alteration evident in early-stage patients indicated that the reduction of plasma CXCL7 is an early event in pancreatic carcinogenesis and may precede the development of cancer. The persistent presence of inflammation is known to promote carcinogenesis in various organs, and chronic pancreatitis is suspected to be one a precancerous condition for pancreatic cancer, although opinions on this issue vary. We measured the plasma level of CXCL7 in a small number of patients diagnosed as having chronic pancreatitis ($n = 10$) using the reverse-phase protein microarray (Table 1). The CXCL7 levels in patients with chronic pancreatitis were significantly lower than those in healthy controls ($P = 0.0002$), but slightly higher than those in patients with pancreatic cancer ($P = 0.095$; Fig. 3C).

Complementation of CA19-9 by CXCL7

CA19-9 is an established biomarker that has long been used for the diagnosis of pancreatic cancer. We found that the levels of CXCL7 and CA19-9 were not mutually correlated (Pearson's $r = 0.289$) and that combination with CXCL7 significantly improved the ability of CA19-9 to distinguish patients with pancreatic cancer from healthy controls: the AUC value improved to 0.965 (95% CI: 0.865–0.994) in the training cohort ($P = 0.026$) and to 0.961 (0.932–0.979) in the validation cohort ($P = 0.002$; Fig. 3D). The AUC values of CA19-9 in the 2 cohorts (Fig. 3D) were comparable with those reported previously (29–31).

Even among individuals with normal levels of CA19-9 (<37 U/mL; a cutoff value widely used in clinical practice), CXCL7 was significantly reduced in pancreatic cancer patients in both the training [$P = 0.014$ and AUC = 0.853 (95% CI: 0.650–0.957; Fig. 4A and B)] and validation [$P < 0.0001$ and AUC = 0.834 (95% CI: 0.747–0.899; Fig. 4B and C)] cohorts.

Because of the low prevalence of pancreatic cancer, any screening biomarker must have high specificity (32). The sensitivity/specificity of CA19-9 (cutoff: 37 U/mL) were 79%/89% in the training cohort and 79%/95% in the validation cohort, consistent with previous reports (32). If we defined the cutoff for CXCL7 as a level at which 95% of healthy individuals would be excluded, 83% of pancreatic cancer patients in the training cohort and 84% in the validation cohort would be detected using the combination of CXCL7 and CA19-9 (Supplementary Table S2).

Discussion

Early detection and subsequent radical surgical resection would most likely provide a chance of cure for patients with pancreatic cancer (7). However, patients with early-stage pancreatic cancer are generally asymptomatic and have little opportunity to undergo imaging and/or other diagnostic procedures until their disease becomes advanced. If a sensitive, but minimally invasive and cost-effective, plasma/serum test were available, it would be effective for alerting patients with early pancreatic cancer and offer them a chance to receive prompt and effective medical attention. In the present study, we compared the plasma LMW proteome between patients with pancreatic cancer and healthy controls using a new proteome platform, 2DICAL (Fig. 1), and found a significant decrease of the plasma CXCL7 level in patients with pancreatic cancer (Fig. 1B and C). The result of quantitative LC-MS was then verified using 3 different methods: immunoblotting (Fig. 1D), multiplex, and reverse-phase protein microarray (Figs. 2 and 3) assays. We further validated the significant decrease of CXCL7 in a larger independent cohort (validation cohort). The level of plasma CXCL7 was confirmed to be decreased reproducibly in patients with pancreatic cancer including those with Stage I and II disease (Table 2). CXCL7 did not

Published OnlineFirst December 8, 2010; DOI:10.1158/1055-9965.EPI-10-0397

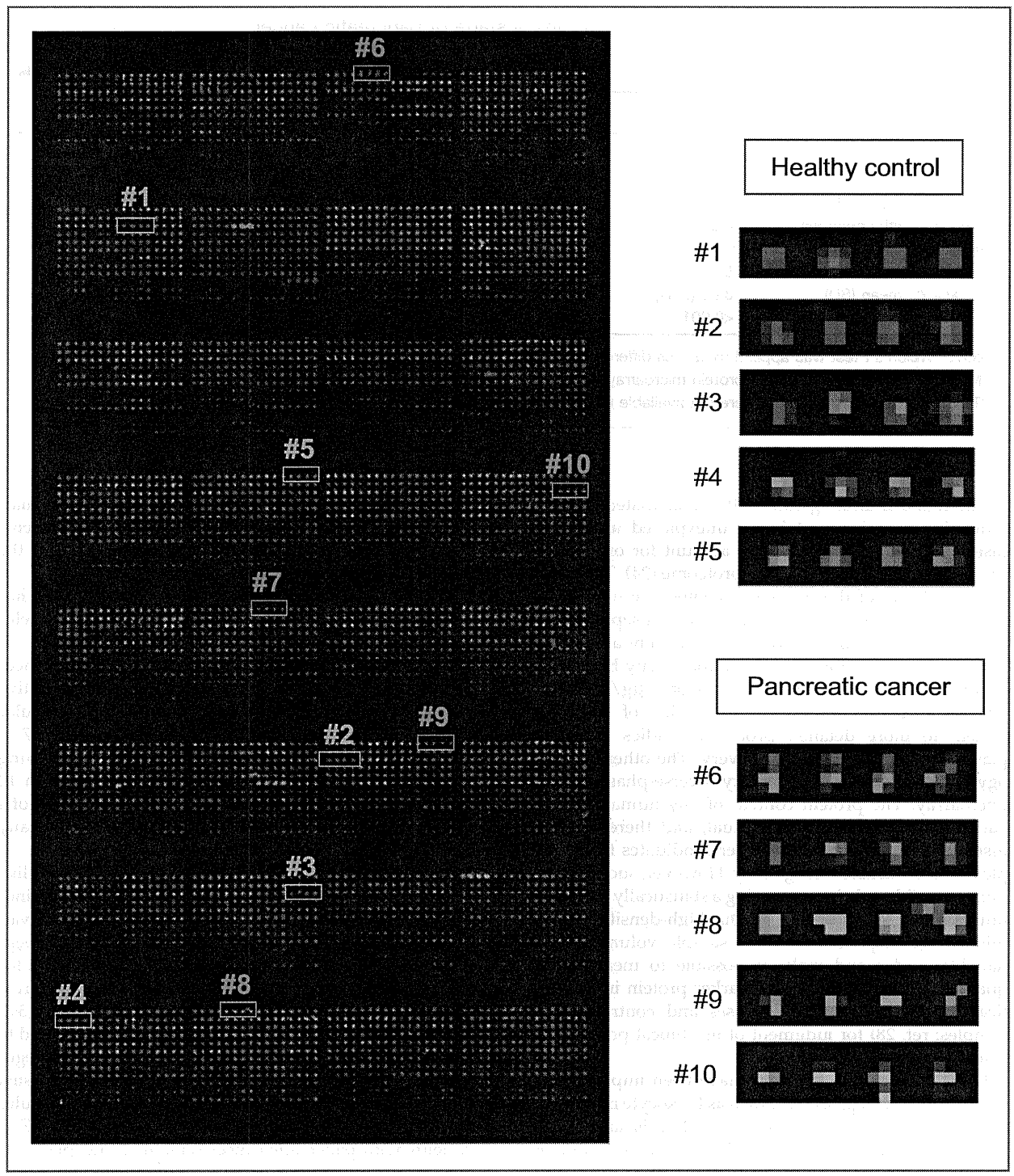


Figure 2. Image of a representative reverse-phase protein microarray slide stained with anti-PBP antibody (left). Samples were randomly assigned, and quadruplicate spots from representative patients with high and low levels of CXCL7 were extracted (right).

surpass the sensitivity of CA19-9, but was able to supplement it. Combination with CXCL7 significantly improved the sensitivity of CA19-9 (Fig. 3D and Supplementary Table S2).

In addition to 2DICAL, we utilized 2 state-of-the-art proteome technologies. The proteome analysis of plasma/serum samples has been hampered by the prominence of a handful of abundant proteins such as

Published OnlineFirst December 8, 2010; DOI:10.1158/1055-9965.EPI-10-0397

Table 2. Plasma CXCL7 level according to clinical stage of pancreatic cancer

	Pancreatic cancer patients				Healthy controls
	Stage I	Stage II	Stage III	Stage IV	
Training cohort					
No. of cases	1	6	4	13	19[b]
CXCL7[a], mean (SD)	3.67 (-)	3.93 (0.24)	3.75 (0.17)	3.82 (0.33)	4.14 (0.18)
P (vs. healthy controls)	0.01	0.01	<0.001	<0.001	
Validation cohort					
No. of cases	5	25	40	70	87
CXCL7[a], mean (SD)	3.89 (0.34)	3.96 (0.25)	4.02 (0.18)	3.86 (0.32)	4.18 (0.14)
P (vs. healthy controls)	<0.001	<0.001	<0.001	<0.001	

NOTE. Welch's *t* test was applied to assess differences in values.

[a]Measured using reverse-phase protein microarrays.

[b]Two patients whose samples were not available for reverse-phase protein microarrays were excluded.

albumin and immunoglobulin. It is anticipated that the remaining proteins contain an unexplored archive of disease-driven information, but account for only about 1% of the entire human plasma proteome (24). To reduce the complexity of the plasma proteome, we used HFM filtration technology. Our HFM devise can separate and concentrate LMW plasma proteins in a fully automated manner (22) and allows identification of any biomarker candidate that is present at a level of 1 μg/mL. This discovery justifies the future application of the HFM system to more detailed proteome studies aimed at plasma/serum biomarker discovery. The other technology we employed is high-density reverse-phase protein microarray. The protein content of any human sample varies according to the individual, and therefore it is essential to distinguish biomarker candidates from simple interindividual heterogeneity. However, such distinction is possible only by comparing a statistically sufficient number of cases and controls. Our high-density protein microarrays require a minimal sample volume of the nanoliter order and make it possible to measure the quantity of any candidate biomarker protein in a statistically sufficient number of cases and controls (>300 samples; ref. 28) for judgment of its clinical potential in a single experiment.

LMW chemotactic cytokines have been implicated in various biological processes, such as leukocyte migration, angiogenesis, hematopoiesis, atherosclerosis, and cancer migration and metastasis. CXCL7, also known as PBP, is one of the members of the angiogenic ELR^+ CXC chemokine family (33). It is reportedly produced and stored in platelets, monocytes, neutrophils, and megakaryocytes. Secreted CXCL7 binds to CXC chemokine receptor 2 (CXCR2) on endothelium and mediates angiogenesis through activation of the Ras/Raf/mitogen-activated protein kinase (MAPK) and PI3K/AKT/mTOR signaling pathways (33, 34). The histology of pancreatic ductal adenocarcinoma is often characterized by hypovascularization. The reduction of circulating CXCL7 in patients with pancreatic cancer may play a certain role in the suppression of angiogenesis.

Recently, reduction in the level of serum CXCL7 has been reported to be a biomarker for advanced myelodysplastic syndrome (35). In contrast, CXCL7 is increased in the pulmonary venous blood of lung cancer patients and is significantly decreased after curative surgical resection of the lung lesions. Of particular interest is the fact that the increment of CXCL7 is detectable several months before diagnosis of lung cancer (36). We observed a reduction of CXCL7 in 10 patients with chronic pancreatitis; but, examination of a larger number of patients will be needed before any definite conclusion can be reached.

CXCL7 is N-terminally truncated by cathepsin G-like enzymes and converted to other types of chemokines with distinct functions such as connective tissue-activating peptide III (CTAP-III) and neutrophil-activating peptide 2 (NAP-2; refs. 37, 38). One possible explanation for the reduction of plasma CXCL7 in patients with pancreatic cancer is degradation by certain exoproteases (39). Matrix metalloproteinase-9 (MMP9) has been reported to degrade CXC chemokines (40). MMP9 is often upregulated in pancreatic cancer cells and secreted into plasma (41). However, in this study, the precise molecular mechanisms behind the reduction of plasma CXCL7 in patients with pancreatic cancer remained unexplained.

Because the process of pancreatic carcinogenesis is probably mediated by various molecular pathways (42), the diagnosis of pancreatic cancer using a single biomarker may not be realistic, and a combination of different biomarkers with distinct spectra would appear to be a more realistic alternative. CA19-9 is the most widely used serum biomarker for pancreatic cancer; but, its sensitivity and specificity have been recognized

Published OnlineFirst December 8, 2010; DOI:10.1158/1055-9965.EPI-10-0397

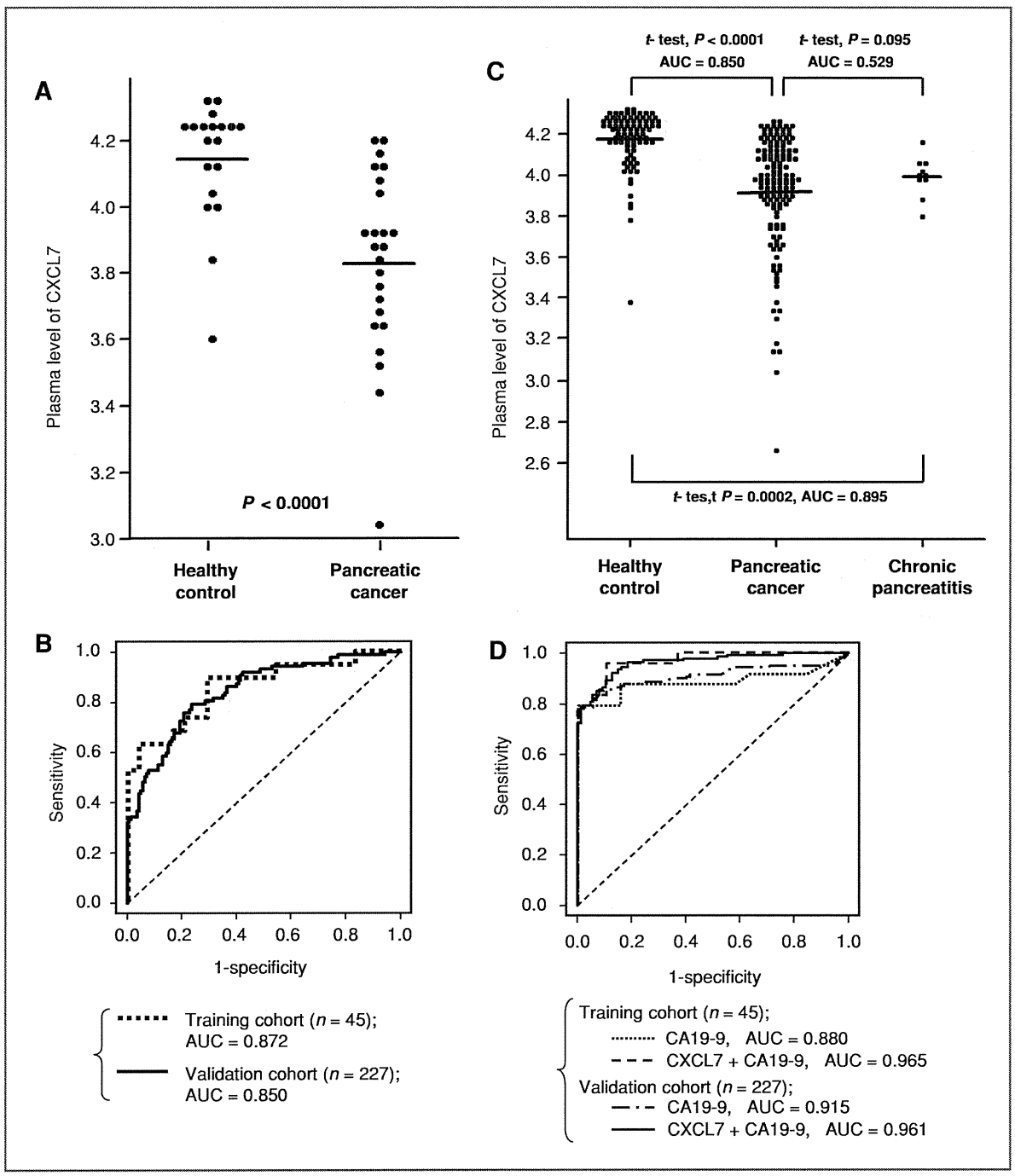


Figure 3. A and C, plasma levels (in arbitrary units) of CXCL7 in healthy controls, patients with pancreatic cancer, and patients with chronic pancreatitis in the training (A) and validation (C) cohorts. Horizontal lines represent the average levels of CXCL7. B, ROC analyses for the discriminatory value of CXCL7 in the training (dotted line) and validation (solid line) cohorts. D, ROC analyses for the discriminatory value of CA19-9 and the composite index of CA19-9 and CXCL7 in the training and validation cohorts.

to be unsatisfactory for pancreatic cancer screening (7, 12). We demonstrated that CXCL7 significantly improved the discriminatory ability of CA19-9, and this improvement was reproducibly validated in a large multi-institutional cohort. However, further independent validation by other investigators is still mandatory before its clinical application can be warranted (15, 29–31, 43).

Published OnlineFirst December 8, 2010; DOI:10.1158/1055-9965.EPI-10-0397

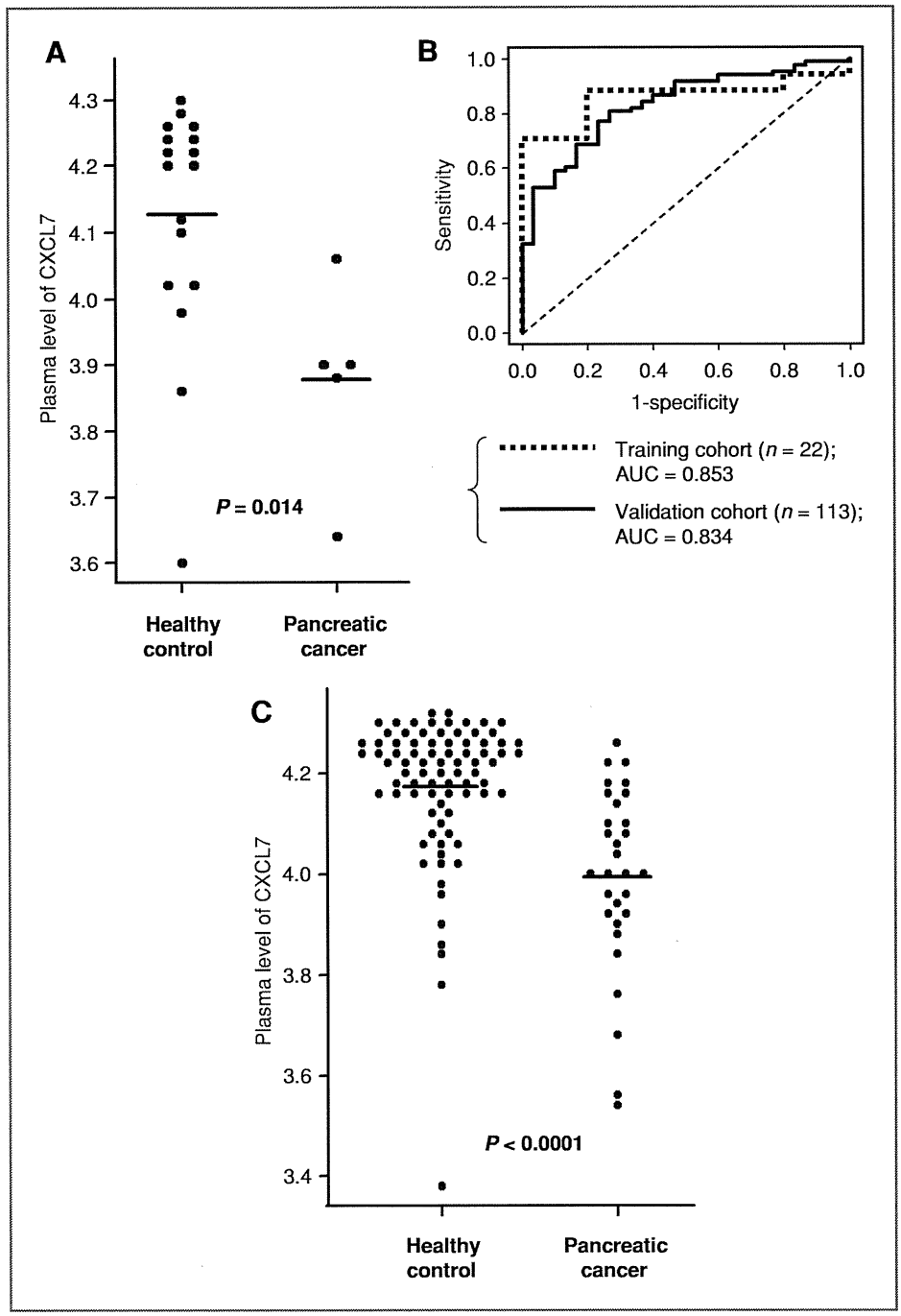


Figure 4. A and C, plasma levels of CXCL7 in healthy controls (left) and patients with pancreatic cancer (right) with the normal levels of CA19-9 (<37 U/mL) in the training (A) and validation (C) cohorts. Horizontal lines represent the average levels of CXCL7. B, ROC analyses of CXCL7 for discrimination between pancreatic cancer patients and healthy controls with normal levels of CA19-9 (<37 U/mL) in the training (dotted line) and validation (solid line) cohorts.

The primary goal of the present study was to discover new biomarkers useful for the early detection of pancreatic cancer in an asymptomatic population. Aberrations of circulating CXCL7 have also been reported in other premalignant conditions. The present study has not only explored the utility of CXCL7 as a biomarker, but also provided a novel insight into the chemokine-mediated reactions that occur during early carcinogenesis.

Disclosure of Potential Conflicts of Interest

These sponsors had no role in the design of the study, the collection of the data, the analysis and interpretation of the data, the decision to submit the manuscript for publication, or the writing of the manuscript.

Acknowledgment

We thank Ms. Ayako Igarashi, Ms. Tomoko Umaki, and Ms. Yuka Nakamura for their technical assistance.

Published OnlineFirst December 8, 2010; DOI:10.1158/1055-9965.EPI-10-0397

Grant Support

Program for Promotion of Fundamental Studies in Health Sciences conducted by the National Institute of Biomedical Innovation of Japan, the Third-Term Comprehensive Control Research for Cancer and Research on Biological Markers for New Drug Development conducted by the Ministry of Health and Labor of Japan.

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.

Received April 15, 2010; revised October 22, 2010; accepted November 28, 2010; published OnlineFirst December 8, 2010.

References

1. Jemal A, Murray T, Samuels A, Ghafoor A, Ward E, Thun MJ. Cancer statistics, 2003. CA Cancer J Clin 2003;53:5–26.
2. Michl P, Pauls S, Gress TM. Evidence-based diagnosis and staging of pancreatic cancer. Best Pract Res Clin Gastroenterol 2006;20:227–51.
3. Ministry of Health, Labour and Welfare. Japanese Government: vital statistics of Japan2009. Available from: http://ganjoho.ncc.go.jp/professional/statistics/odjrh3000000hwsa-att/cancer_mortality(1958-2008).xls.
4. American Cancer Society. Cancer Facts and Figures2007. Atlanta, GA: American Cancer Society; 2007. p. 16–7.
5. Nitecki SS, Sarr MG, Colby TV, van Heerden JA. Long-term survival after resection for ductal adenocarcinoma of the pancreas. Is it really improving?Ann Surg 1995;221:59–66.
6. Tsuchiya R, Noda T, Harada N, Miyamoto T, Tomioka T, Yamamoto K, et al. Collective review of small carcinomas of the pancreas. Ann Surg 1986;203:77–81.
7. DiMagno EP, Reber HA, Tempero MA. AGA technical review on the epidemiology, diagnosis, and treatment of pancreatic ductal adenocarcinoma. American Gastroenterological Association. Gastroenterology 1999;117:1464–84.
8. Nieto J, Grossbard ML, Kozuch P. Metastatic pancreatic cancer 2008: is the glass less empty?Oncologist 2008;13:562–76.
9. Burris HA 3rd, Moore MJ, Andersen J, Green MR, Rothenberg ML, Modiano MR, et al. Improvements in survival and clinical benefit with gemcitabine as first-line therapy for patients with advanced pancreas cancer: a randomized trial. J Clin Oncol 1997;15:2403–13.
10. Poplin E, Feng Y, Berlin J, Rothenberg ML, Hochster H, Mitchell E, et al. Phase III, randomized study of gemcitabine and oxaliplatin versus gemcitabine (fixed-dose rate infusion) compared with gemcitabine (30-minute infusion) in patients with pancreatic carcinoma E6201: a trial of the Eastern Cooperative Oncology Group. J Clin Oncol 2009;27:3778–85.
11. NCCN Clinical Practice Guidelines in Oncology: Pancreatic Adenocarcinoma V.1. National Comprehensive Cancer Network; 2009.
12. Goggins M. Molecular markers of early pancreatic cancer. J Clin Oncol 2005;23:4524–31.
13. Goggins M, Canto M, Hruban R. Can we screen high-risk individuals to detect early pancreatic carcinoma?J Surg Oncol 2000;74:243–8.
14. Hanash S. Disease proteomics. Nature 2003;422:226–32.
15. Grantzdorffer I, Carl-McGrath S, Ebert MP, Rocken C. Proteomics of pancreatic cancer. Pancreas 2008;36:329–36.
16. Yamaguchi U, Nakayama R, Honda K, Ichikawa H, Hasegawa T, Shitashige M, et al. Distinct gene expression-defined classes of gastrointestinal stromal tumor. J Clin Oncol 2008;26:4100–8.
17. Ono M, Shitashige M, Honda K, Isobe T, Kuwabara H, Matsuzuki H, et al. Label-free quantitative proteomics using large peptide data sets generated by nanoflow liquid chromatography and mass spectrometry. Mol Cell Proteomics 2006;5:1338–47.
18. Matsubara J, Ono M, Negishi A, Ueno H, Okusaka T, Furuse J, et al. Identification of a predictive biomarker for hematologic toxicities of gemcitabine. J Clin Oncol 2009;27:2261–8.
19. Negishi A, Ono M, Handa Y, Kato H, Yamashita K, Honda K, et al. Large-scale quantitative clinical proteomics by label-free liquid chromatography and mass spectrometry. Cancer Sci 2009;100:514–9.
20. Ono M, Matsubara J, Honda K, Sakuma T, Hashiguchi T, Nose H, et al. Prolyl 4-hydroxylation of alpha-fibrinogen: a novel protein modification revealed by plasma proteomics. J Biol Chem 2009;284:29041–9.
21. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics 2002;1:845–67.
22. Tanaka Y, Akiyama H, Kuroda T, Jung G, Tanahashi K, Sugaya H, et al. A novel approach and protocol for discovering extremely low-abundance proteins in serum. Proteomics 2006;6:4845–55.
23. Kennedy S. The role of proteomics in toxicology: identification of biomarkers of toxicity by protein expression analysis. Biomarkers 2002;7:269–90.
24. Tirumalai RS, Chan KC, Prieto DA, Issaq HJ, Conrads TP, Veenstra TD. Characterization of the low molecular weight human serum proteome. Mol Cell Proteomics 2003;2:1096–103.
25. Available from: http://www.fhcrc.org/science/international_biomarker/.
26. Honda K, Yamada T, Hayashida Y, Idogawa M, Sato S, Hasegawa F, et al. Actinin-4 increases cell motility and promotes lymph node metastasis of colorectal cancer. Gastroenterology 2005;128:51–62.
27. Idogawa M, Yamada T, Honda K, Sato S, Imai K, Hirohashi S. Poly (ADP-ribose) polymerase-1 is a component of the oncogenic T-cell factor-4/beta-catenin complex. Gastroenterology 2005;128:1919–36.
28. Matsubara J, Ono M, Honda K, Negishi A, Ueno H, Okusaka T, et al. Survival prediction for pancreatic cancer patients receiving gemcitabine treatment. Mol Cell Proteomics 2010;9:695–704.
29. Fiedler GM, Leichtle AB, Kase J, Baumann S, Ceglarek U, Felix K, et al. Serum peptidome profiling revealed platelet factor 4 as a potential discriminating Peptide associated with pancreatic cancer. Clin Cancer Res 2009;15:3812–9.
30. Koopmann J, Zhang Z, White N, Rosenzweig J, Fedarko N, Jagannath S, et al. Serum diagnosis of pancreatic adenocarcinoma using surface-enhanced laser desorption and ionization mass spectrometry. Clin Cancer Res 2004;10:860–8.
31. Koopmann J, Buckhaults P, Brown DA, Zahurak ML, Sato N, Fukushima N, et al. Serum macrophage inhibitory cytokine 1 as a marker of pancreatic and other periampullary cancers. Clin Cancer Res 2004;10:2386–92.
32. Rhodes JM. Usefulness of novel tumour markers. Ann Oncol 1999;10Suppl 4:118–21.
33. Strieter RM, Burdick MD, Mestas J, Gomperts B, Keane MP, Belperio JA. Cancer CXC chemokine networks and tumour angiogenesis. Eur J Cancer 2006;42:768–78.
34. Heidemann J, Ogawa H, Dwinell MB, Rafiee P, Maaser C, Gockel HR, et al. Angiogenic effects of interleukin 8 (CXCL8) in human intestinal microvascular endothelial cells are mediated by CXCR2. J Biol Chem 2003;278:8508–15.
35. Aivado M, Spentzos D, Germing U, Alterovitz G, Meng XY, Grall F, et al. Serum proteome profiling detects myelodysplastic syndromes and identifies CXC chemokine ligands 4 and 7 as markers for advanced disease. Proc Natl Acad Sci USA 2007;104:1307–12.
36. Yee J, Sadar MD, Sin DD, Kuzyk M, Xing L, Kondra J, et al. Connective tissue-activating peptide III: a novel blood biomarker for early lung cancer detection. J Clin Oncol 2009;27:2787–92.
37. Ehlert JE, Ludwig A, Grimm TA, Lindner B, Flad HD, Brandt E. Down-regulation of neutrophil functions by the ELR(+) CXC chemokine platelet basic protein. Blood 2000;96:2965–72.
38. Pillai MM, Iwata M, Awaya N, Graf L, Torok-Storb B. Monocyte-derived CXCL7 peptides in the marrow microenvironment. Blood 2006;107:3520–6.

Published OnlineFirst December 8, 2010; DOI:10.1158/1055-9965.EPI-10-0397

39. Villanueva J, Shaffer DR, Philip J, Chaparro CA, Erdjument-Bromage H, Olshen AB, et al. Differential exoprotease activities confer tumor-specific serum peptidome patterns. J Clin Invest 2006;116:271–84.
40. Van Den Steen PE, Proost P, Wuyts A, Van Damme J, Opdenakker G. Neutrophil gelatinase B potentiates interleukin-8 tenfold by aminoterminal processing, whereas it degrades CTAP-III, PF-4, and GRO-alpha and leaves RANTES and MCP-2 intact. Blood 2000;96:2673–81.
41. Tian M, Cui YZ, Song GH, Zong MJ, Zhou XY, Chen Y, et al. Proteomic analysis identifies MMP-9, DJ-1 and A1BG as overexpressed proteins in pancreatic juice from pancreatic ductal adenocarcinoma patients. BMC Cancer 2008;8:241.
42. Gansauge S, Gansauge F, Beger HG. Molecular oncology in pancreatic cancer. J Mol Med 1996;74:313–20.
43. Honda K, Hayashida Y, Umaki T, Okusaka T, Kosuge T, Kikuchi S, et al. Possible detection of pancreatic cancer by plasma protein profiling. Cancer Res 2005;65:10613–22.

LEGENDS FOR SUPPLEMENTARY FIGURES

Supplementary Figure S1

Ten randomly selected plasma samples from the Training Cohort were filtered through hollow-fiber membranes (HFM), loaded onto 5-20% SDS gel under reducing conditions, and visualized by silver staining.

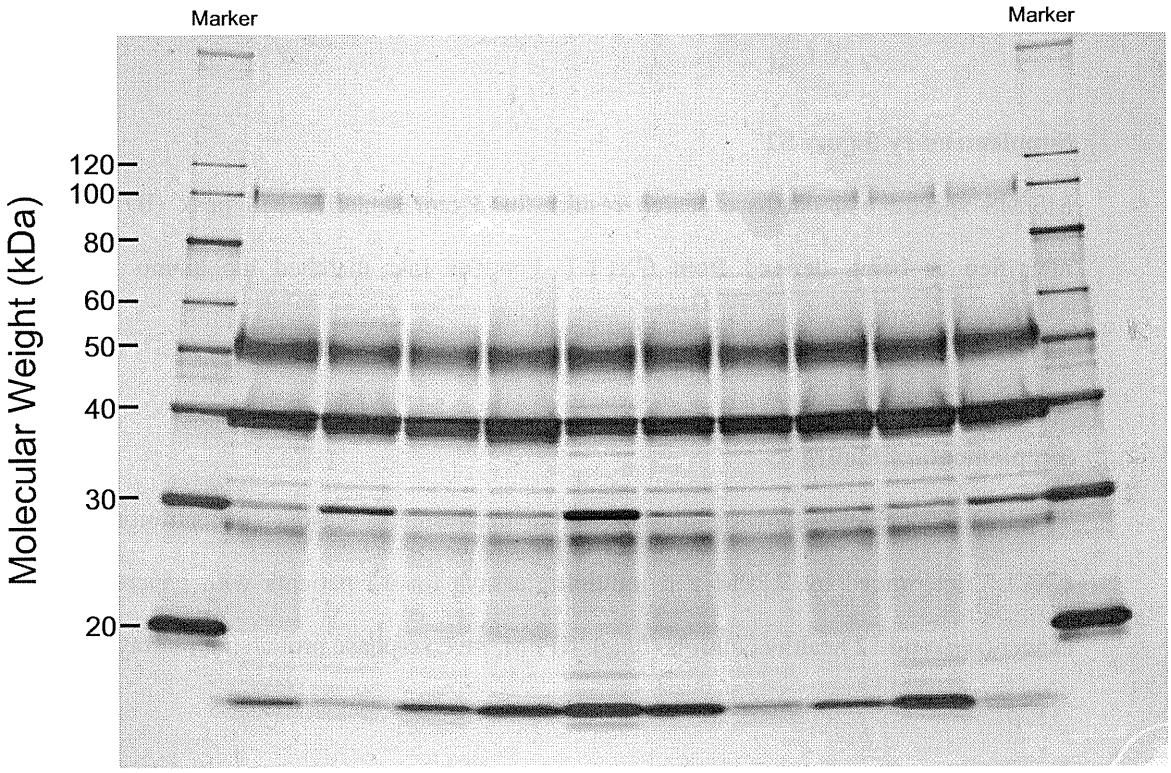
Supplementary Figure S2

MS/MS spectra and database search results for a representative MS peak (ID 54) identified as being derived from CXCL7. Peptides that matched the amino acid sequences in the database are highlighted in red.

Supplementary Figure S3

Two-dimensional plot showing the correlation between the plasma concentration of CXCL7 determined by RPPM and multiplex assays for 12 patients with pancreatic cancer (*red*) and 12 healthy controls (*blue*). RPPM, reverse-phase protein microarrays.

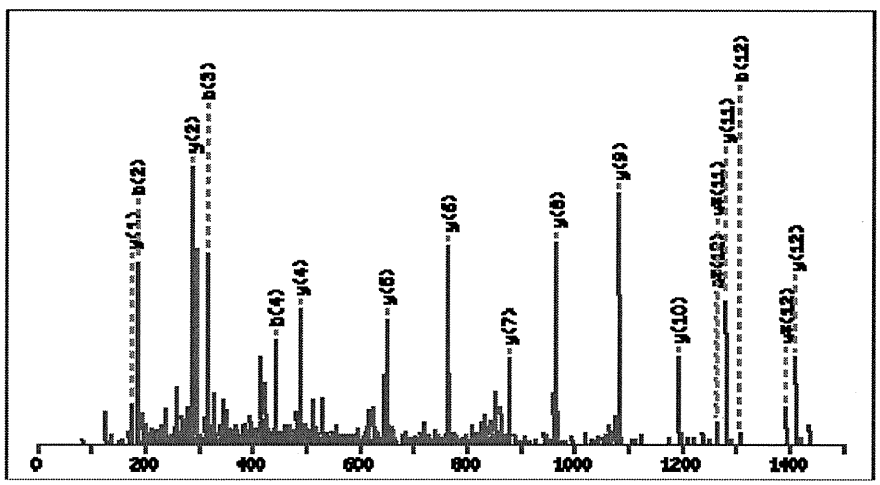
Supplementary Figure S1



Supplementary Figure S2

Peak ID: 54

MS/MS Fragmentation of **GKEESLDSDLYAELR**
Found in **SCYB7_HUMAN**, Platelet basic protein precursor (PBP) (CXCL7)
[Contains: Connective tissue-activating peptide III (CTAP-III)]
Match to Query 4: 1723.752730 from(862.883641,2+)



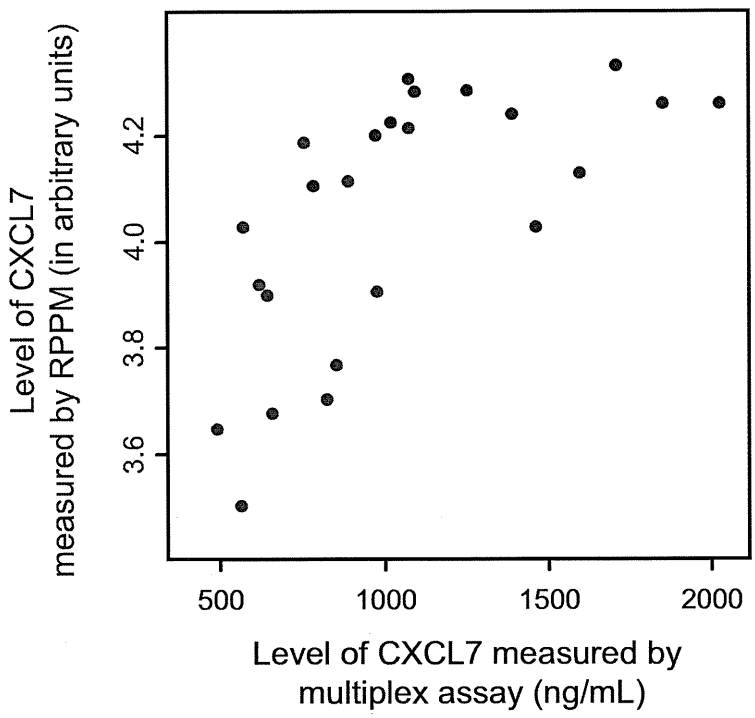
Monoisotopic mass of neutral peptide Mr(calc): 1723.8264
Ions Score: 100 Expect: 6.2e-008
Matches (Bold): 18/164 fragment ions using 29 most intense peaks

#	a	a^{++}	a*	a^{*++}	b	b^{++}	b*	b^{*++}	Seq.	y	y^{++}	y*	y^{*++}	#
1	30.0338	15.5206			58.0287	29.5180			G					15
2	158.1288	79.5680	141.1022	71.0548	**186.1237**	93.5655	169.0972	85.0522	K	1667.8123	834.4098	1650.7857	825.8965	14
3	287.1714	144.0893	270.1448	135.5761	**315.1663**	158.0868	298.1397	149.5735	E	1539.7173	770.3623	1522.6908	761.8490	13
4	416.2140	208.6106	399.1874	200.0974	**444.2089**	222.6081	427.1823	214.0948	E	**1410.6747**	705.8410	**1393.6482**	697.3277	12
5	503.2460	252.1266	486.2195	243.6134	531.2409	266.1241	514.2144	257.6108	S	**1281.6321**	641.3197	**1264.6056**	632.8064	11
6	616.3301	308.6687	599.3035	300.1554	644.3250	322.6661	627.2984	314.1529	L	**1194.6001**	597.8037	1177.5735	589.2904	10
7	731.3570	366.1821	714.3305	357.6689	759.3519	380.1796	742.3254	371.6663	D	**1081.5160**	541.2617	1064.4895	532.7484	9
8	818.3890	409.6982	801.3625	401.1849	846.3840	423.6956	829.3574	415.1823	S	**966.4891**	483.7482	949.4625	475.2349	8
9	933.4160	467.2116	916.3894	458.6984	961.4109	481.2091	944.3843	472.6958	D	**879.4571**	440.2322	862.4305	431.7189	7
10	1046.5000	523.7537	1029.4735	515.2404	1074.4950	537.7511	1057.4684	529.2378	L	**764.4301**	382.7187	747.4036	374.2054	6
11	1209.5634	605.2853	1192.5368	596.7721	1237.5583	619.2828	1220.5317	610.7695	Y	**651.3461**	326.1767	634.3195	317.6634	5
12	1280.6005	640.8039	**1263.5739**	632.2906	**1308.5954**	654.8013	1291.5689	646.2881	A	**488.2827**	244.6450	471.2562	236.1317	4
13	1409.6431	705.3252	1392.6165	696.8119	1437.6380	719.3226	1420.6114	710.8094	E	417.2456	209.1264	400.2191	200.6132	3
14	1522.7271	761.8672	1505.7006	753.3539	1550.7221	775.8647	1533.6955	767.3514	L	**288.2030**	144.6051	271.1765	136.0919	2
15									R	**175.1190**	88.0631	158.0924	79.5498	1

Matched peptides shown in Bold Red

```
     1 MSLRLDTTPS CNSARPLHAL QVLLLLSLLL TALASSTKGQ TKRNLAKGKE
    51 ESLDSDLYAE LRCMCIKTTS GIHPKNIQSL EVIGKGTHCN QVEVIATLKD
   101 GRKICLDPDA PRIKKIVQKK LAGDESAD
```

Supplementary Figure S3



Supplementary Table S1. Plasma Proteins Whose Intensity Differed Significantly Between Patients with Pancreatic Cancer and Healthy Controls

Gene locus	Protein identification	Mascot score	AUC	No. of matched peptide peaks*
SCYB7_HUMAN	Platelet basic protein precursor (PBP) (Small inducible cytokine B7) (CXCL7) (Leukocyte-derived growth factor) (LDGF) (Macrophage-derived growth factor) (MDGF) [Contains: Connective tissue-activating peptide III (CTAP-III)]	99.60	0.84	2
FIBA_HUMAN	Fibrinogen alpha chain precursor [Contains: Fibrinopeptide A] - Homo sapiens (Human)	99.46	0.83	12
THRB_HUMAN	Prothrombin precursor (EC 3.4.21.5) (Coagulation factor II) - Homo sapiens (Human)	97.81	0.83	2
K2C1B_HUMAN	Keratin, type II cytoskeletal 1b (Keratin-77) - Homo sapiens (Human)	73.07	0.80	1
RETBP_HUMAN	Plasma retinol-binding protein precursor (PRBP) (RBP) [Contains: Plasma retinol-binding protein(1-182); Plasma retinol-binding protein(1-181); Plasma retinol-binding protein(1-179); Plasma retinol-binding protein(1-176)] - Homo sapiens	66.52	0.83	2
CAH2_HUMAN	Carbonic anhydrase 2 (EC 4.2.1.1) (Carbonic anhydrase II) (Carbonate dehydratase II) (Carbonic anhydrase C) - Homo	56.67	0.80	1
ABTB1_HUMAN	Ankyrin repeat and BTB/POZ domain-containing protein 1 (Elongation factor 1A-binding protein) - Homo sapiens (Human)	37.08	0.85	2
IF3A_HUMAN	Eukaryotic translation initiation factor 3 subunit 10 (eIF-3 theta) (eIF3 p167) (eIF3 p180) (eIF3 p185) (eIF3a) - Homo	34.93	0.88	1
NR1I3_HUMAN	Orphan nuclear receptor NR1I3 (Constitutive androstane receptor) (CAR) (Constitutive activator of retinoid response) (Constitutive active response) (Orphan nuclear receptor MB67) - Homo sapiens (Human)	34.70	0.80	1
HYAL3_HUMAN	Hyaluronidase-3 precursor (EC 3.2.1.35) (Hyal-3) (Hyaluronoglucosaminidase-3) (LUCA-3) - Homo sapiens (Human)	32.24	0.88	1

NOTE. Only proteins with the Mascot scores of >30 are listed.

Supplementary Table S2. Sensitivity and Specificity of CXCL7, CA19-9, and Their Combination in Pancreatic Cancer Detection

	Sensitivity					
	All stages	Stage I	Stage II	Stage III	Stage IV	Specificity
Training cohort						
No. of cancer patients	24	1	6	4	13	Healthy cases n=19[†]
CXCL7[*], n (%)	11 (46)	1 (100)	2 (33)	3 (75)	5 (38)	18 (95)
CA19-9, n (%)	19 (79)	1 (100)	4 (67)	4 (100)	10 (77)	17 (89)
Combination, n (%)	20 (83)	1 (100)	5 (83)	4 (100)	10 (77)	16 (84)
Validation cohort						
No. of cancer patients	140	5	25	40	70	Healthy cases n=87
CXCL7[*], n (%)	51 (36)	2 (40)	8 (32)	9 (23)	32 (46)	83 (95)
CA19-9, n (%)	110 (79)	3 (60)	19 (76)	33 (83)	55 (79)	83 (95)
Combination, n (%)	117 (84)	3 (60)	20 (80)	35 (88)	59 (84)	79 (91)

NOTE. *Measured using reverse-phase protein microarray. The cutoff point was set at the lower 5 percentile for healthy individuals in each cohort.
†Two patients whose samples were not available for reverse-phase protein microarrays were excluded.
Abbreviations: CXCL7, CXC chemokine ligand 7; SD, standard deviation.

Cancer Science
The official journal of the Japanese Cancer Association JCA

Plasma biomarker discovery and validation for colorectal cancer by quantitative shotgun mass spectrometry and protein microarray

Yusuke Murakoshi,[1,2] Kazufumi Honda,[1] Shizuka Sasazuki,[3] Masaya Ono,[1] Ayako Negishi,[1] Junichi Matsubara,[1] Tomohiro Sakuma,[4] Hideya Kuwabara,[4] Shoji Nakamori,[5] Naohiro Sata,[6] Hideo Nagai,[6] Tatsuya Ioka,[7] Takuji Okusaka,[8] Tomoo Kosuge,[8] Masashi Shimahara,[9] Yohichi Yasunami,[10] Yoshinori Ino,[11] Akihiko Tsuchida,[2] Tatsuya Aoki,[2] Shoichiro Tsugane[3] and Tesshi Yamada[1,12]

[1]Chemotherapy Division, National Cancer Center Research Institute; [2]Third Department of Surgery, Tokyo Medical University; [3]Epidemiology and Prevention Division, Research Center for Cancer Prevention and Screening, National Cancer Center; [4]BioBusiness Group, Mitsui Knowledge Industry, Tokyo; [5]Department of Surgery, Osaka National Hospital, National Hospital Organization, Osaka; [6]Department of Surgery, Jichi Medical University, Tochigi; [7]Department of Hepatobiliary and Pancreatic Oncology, Osaka Medical Center for Cancer and Cardiovascular Diseases, Osaka; [8]Hepatobiliary and Pancreatic Oncology and Hepatobiliary and Pancreatic Surgery Divisions, National Cancer Center Hospital, Tokyo; [9]Department of Oral Surgery, Osaka Medical College, Osaka; [10]Department of Regenerative Medicine and Transplantation, Fukuoka University Faculty of Medicine, Fukuoka; [11]Pathology Division, National Cancer Center Research Institute, Tokyo, Japan

(Received August 6, 2010/Revised November 29, 2010/Accepted November 30, 2010/Accepted manuscript online December 3, 2010/Article first published online December 28, 2010)

The development of a new plasma biomarker for early detection would be necessary to improve the overall outcome of colorectal cancer. Here we report the identification and validation of the ninth component of complement (C9) as a novel plasma biomarker for colorectal cancer by cutting-edge proteomic technologies. Plasma proteins were enzymatically digested into a large array of peptides, and the relative quantity of a total of 94 803 peptide peaks was compared between 31 colorectal cancer patients and 59 age/sex-matched healthy controls using 2D image-converted analysis of liquid chromatography and mass spectrometry. The selected biomarker candidates were validated in 345 subjects (115 colorectal cancer patients and 230 age/sex-matched healthy controls) using high-density reverse-phase protein microarrays. Plasma levels of Apo AI and C9 in colorectal cancer patients significantly differed from healthy controls with P values of 7.94×10^{-4} and 1.43×10^{-12} (Student's t-test), respectively. In particular, C9 was elevated in patients with colorectal cancer, including those with stage-I and -II diseases ($P = 3.01 \times 10^{-3}$ and $P = 1.13 \times 10^{-5}$, respectively). Although the significance of the present study must be validated in an independent clinical study, the increment of plasma C9 may be applicable to the early detection of colorectal cancer. (*Cancer Sci* 2011; 102: 630–638)

Colorectal cancer is the third most common cancer worldwide, with an estimated one and half million newly diagnosed cases every year.[1] In Japan, colorectal cancer is currently the third cause of cancer death in men and the first in women, but its incidence is predicted to increase and become the leading cause by 2015, most likely due to changing dietary habits as well as environmental conditions.[2,3] Successful prevention of death from colorectal cancer depends on its early detection. The surgical management of early stage colorectal cancer without metastasis is relatively uncomplicated. There is a significant level of evidence that the application of the fecal occult blood test to mass screening reduces the risk of colorectal cancer death, but its sensitivity and specificity, especially for early stage colorectal cancer, seem to be insufficient. Barium enema, flexible sigmoidoscopy/colonoscopy and ^{18}F-fluorodeoxyglucose positron emission tomography have higher specificity, but may not be cost- and labor-effective for mass screening of the asymptomatic general population.

The circulating blood proteome holds great promise as a reservoir of disease information, and a large variety of plasma/serum proteins have been used as disease biomarkers. Carcinoembryonic antigen (CEA) is a serum biomarker most widely used for colorectal cancer. However, the serum level of CEA often does not elevate in patients with colorectal cancer in the early stages and cannot be applicable to early detection of the disease.[4] We therefore have been searching for new serum/plasma biomarkers that can be used for mass survey of colorectal cancer.

Recently, various mass spectrometry (MS)-based proteomic technologies have been applied to clinical samples in the hope of identifying new disease biomarkers.[5–9] Among those technologies, shotgun proteomics has been considered the most promising because of its high sensitivity and protein identification capability: protein samples are enzymatically digested into a large array of peptides with uniform physical characteristics, and every peptide fragment is analyzed by liquid chromatography (LC) and MS. However, the number of samples that can be compared by shotgun proteomics is usually limited because isotope labeling is necessary to give a quantitative dimension to shotgun proteomics. The protein contents of each human sample vary significantly among individuals, and biomarker candidates can be distinguished from simple personal heterogeneity only by comparing a sufficient number of cases and controls. To overcome this limitation, we developed a software named 2D image converted analysis of liquid chromatography and MS (2DICAL),[9,10] which enabled accurate quantitative comparison across a theoretically unlimited number of LC-MS data without isotope labeling.[10] Using this powerful software we successfully identified biomarkers that can predict the hematological toxicities and survival of pancreatic cancer patients receiving gemcitabine.[11,12]

In the present study we searched for a biomarker that can be used for the early detection of colorectal cancer using 2DICAL. We carefully selected cases and controls to be compared by matching their age and gender distributions, as well as residential areas. We identified the significance increment of complement component C9 in the sera of patients with colorectal cancer, and its significance was validated in a large cohort using another innovative proteome technology, high-density reverse-phase protein microarray (RPPM).

[12]To whom correspondence should be addressed. E-mail: tyamada@ncc.go.jp

doi: 10.1111/j.1349-7006.2010.01818.x

Materials and Methods

Patients and plasma samples. Plasma samples (n = 345) were collected from 115 patients diagnosed with colorectal cancer and 230 healthy volunteers matched with cancer patients by sex, age and residential area (two controls for each cancer patient; Table 1) between October 1998 and March 2002 with informed consent, as described previously.[13] Thirty-one cancer patients and 59 controls were randomly selected and subjected to 2DICAL analysis.

Another set of plasma samples were collected prospectively from 378 individuals including healthy volunteers and newcomers mainly to the Department of Gastroenterology at the National Cancer Center Hospital (Tokyo), Osaka National Hospital (Osaka), Jichi Medical School Hospital (Shimotsuke), Osaka Medical College (Osaka), Tokyo Medical University Hospital (Tokyo), Osaka Medical Center for Cancer and Cardiovascular Diseases (Osaka), and Fukuoka University Hospital (Fukuoka). This multi-institutional study was conducted as part of the "Third-Term Comprehensive Control Research for Cancer" conducted by the Ministry of Health, Labour and Welfare of Japan. Written informed consent was obtained from every individual, and the study protocol was approved by the ethics committee of each participating institution.

Sample preparation and LC-MS. Twelve abundant plasma proteins including albumin, IgG, α1-antitrypsin, IgA, IgM, transferrin, haptoglobin, α1-acid glycoprotein, α2-macroglobulin, Apo AI, apolipoprotein AII and fibrinogen were deduced from plasma samples using IgY-12 spin columns (Beckman Coulter, Fullerton, CA, USA) prior to MS analysis.[11] The deduced samples were precipitated with acetonitrile, dried and digested to peptide with modified trypsin. LC-MS and data acquisition were performed as described previously.[14] Briefly, MS spectra were acquired using nano-electrospray ionization (nano-ESI)-quadruple time-of-flight (QqTOF) MS (QTOF Ultima; Waters, Milford, MA, USA) every second for 60 min in the 250–1600 m/z range. The MS peaks of each sample with the same m/z were extracted every 1 m/z and aligned. Peak lists were created using the MassNavigator software package (version 1.2; Mitsui Knowledge Industry, Tokyo, Japan). Targeted tandem MS (MS/MS) data were analyzed with the Mascot software package (version 2.2.1; Matrix Sciences, London, UK) against the NCBInr database (NCBInr_20070419. fast).

Antibodies. Anti-Apo AI rabbit polyclonal antibody was purchased from Calbiochem (Darmstadt, Germany), anti-C9 mouse monoclonal antibody from AntibodyShop (Gentofte, Denmark) and anti-α2-macroglobulin mouse monoclonal antibody from R&D Systems (Minneapolis, MN, USA).

Immunoblot analysis. Protein was separated by SDS-PAGE and transferred to polyvinyldene difluoride membrane, as described previously.[15] The membrane was incubated with anti-Apo AI, anti-C9 or anti-α2-macroglobulin (loading control) and then with relevant horseradish peroxidase (HRP)-conjugated secondary antibody. Blots were detected by ECL western blotting detection reagent according to the manufacturer's instruction (Amersham Biosciences, Buckinghamshire, UK).

RPPM. Plasma samples were serially diluted using PBS containing 0.01% triton X with or without 1% bovine serum albumin (BSA). The diluted plasma samples were printed onto slide glasses coated with ProLinker (ProteoChip; Proteogen, Seoul, Korea)[16,17] using a protein microarrayer equipped with 32 stainless steel pins of 100 μm diameter (Kakengeneqs, Matsudo, Japan) at 4°C. Printed microarray slides were incubated overnight at 37°C and stored at 4°C under desiccation. After returning to room temperature, the array slides were blocked with PBS containing 0.5% casein for 30 min and hybridized overnight with the first antibodies at 4°C. After washing, the array slides were hybridized with relevant biotinylated second antibodies (Vector Laboratories Inc. Burlingame, CA, USA) for 1 h and subsequently with avidin–HRP conjugated reagent (Amersham Biosciences) for 30 min. The fluorescent Cy5 signals were amplified using the tyramide signal amplification sys-

Table 1. Clinicopathological characteristics of individuals examined in the present study

	All cases (n = 345)			Cases analyzed by 2DICAL (n = 90)		
	Cancer (n = 115)	Healthy (n = 230)	P value	Cancer (n = 31)	Healthy (n = 59)	P value
Age (mean ± SD) (years)	59.3 ± 8.9	59.4 ± 8.9	0.93‡	56.8 ± 9.9	56.4 ± 9.8	0.79‡
Gender						
Male	71	142	0.92§	20	36	0.75§
Female	44	88		11	23	
Primary site						
Cecum	6			0		
Ascending colon	27			9		
Transverse colon	8			4		
Descending colon	5			3		
Sigmoid colon	18			5		
Rectum	51			10		
Clinical stage†						
0	17			5		
I	35			10		
II	28			7		
III	25			6		
IV	10			3		
Histology						
Well-differentiated adenocarcinoma	74			21		
Moderately differentiated adenocarcinoma	37			8		
Poorly differentiated adenocarcinoma	2			1		
Others	2			1		

†According to TNM Classification of Malignant Tumors (International Union Against Cancer), 6th Edition. ‡Student's t-test. §Chi-square test.

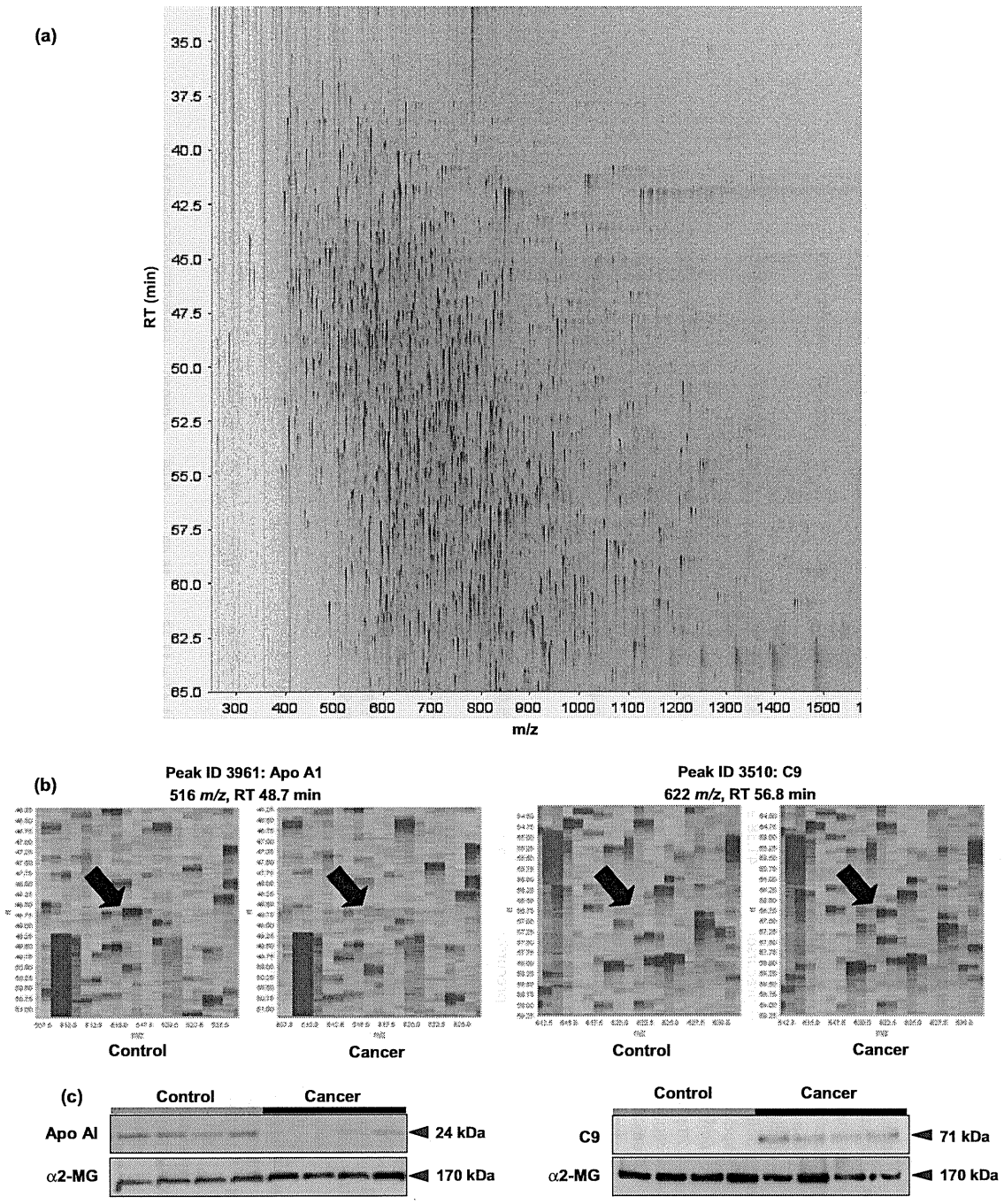


Fig. 1. Plasma biomarker discovery by 2D image converted analysis of liquid chromatography and mass spectrometry (2DICAL). (a) Two-dimensional display of >90 000 peaks of a representative sample with retention time (RT; in minutes) along the vertical (*X*) axis and with mass to charge (*m/z*) ration value along the horizontal (*Y*) axis. (b) Peptide peaks derived from apolipoprotein AI (Apo AI; left; ID 3961) and complement component C9 (C9; right; ID 3510) in a representative colorectal cancer patient (Cancer) and a representative healthy individual (Control). (c) Detection of Apo AI, C9 and α2-macroglobulin (α2-MG, loading control) in plasma samples of four representative colorectal cancer patients and four healthy controls by immunoblotting.

tem (Perkin Elmer, Boston, MA, USA) according to the manufacturer's instructions. Counterstaining was performed with Alexa 488-labeled anti-human IgG antibody (Invitrogen, Carlsbad, CA, USA). Dual-color fluorescent signals (green and red) were detected with a confocal laser microarray scanner (InnoScan 700 AL; Innopsys, Carbonne, France). The median signal intensity of quadruplicates was calculated using the Mapix software (Innopsys).

Measurement of CEA. The plasma level of CEA was determined using the CEA Enzyme Immunoassay kit (Hope Laboratories, Belmont, CA, USA) according to the manufacturer's instruction.

Statistical analysis. Statistically significant differences were detected using Paired *t*-test and Student's *t*-test. Interquartile range (IQR), receiver operator characteristics (ROC) and area under the curve (AUC) analyses were performed using the

doi: 10.1111/j.1349-7006.2010.01818.x

Table 2. Protein identification of peptide peaks with significant difference between colorectal cancer patients and healthy controls

ID	M/Z	RT	prot_acc	prot_desc	Cancer†	Control†	pep_score	AUC	*P*-value‡	pep_seq
664	651.33	46.18	APOA1_HUMAN	Apolipoprotein A-I precursor (Apo-AI) (ApoA-I) (contains: apolipoprotein A-I [1-242])	34.01 ± 18.34	60.26 ± 41.89	58.87	0.77	0.00038	THLAPYSDELR
3961	516.78	48.71	APOA1_HUMAN	Apolipoprotein A-I precursor (Apo-AI) (ApoA-I) (contains: apolipoprotein A-I [1-242])	11.40 ± 3.85	17.32 ± 7.43	49.08	0.81	0.00006	LSPLGEEMR
2384	632.03	63.40	A2GL_HUMAN	Leucine-rich alpha-2-glycoprotein precursor (LRG)	42.66 ± 23.89	24.61 ± 10.75	46.04	0.79	0.00094	ENQLEVLEVSWLHGLK
2819	506.81	45.11	APOA1_HUMAN	Apolipoprotein A-I precursor (Apo-AI) (ApoA-I) (contains: apolipoprotein A-I [1-242])	16.56 ± 7.20	26.31 ± 14.57	45.5	0.77	0.00063	AKPALEDLR
516	642.31	52.07	APOA1_HUMAN	Apolipoprotein A-I precursor (Apo-AI) (ApoA-I) (contains: apolipoprotein A-I [1-242])	41.38 ± 21.38	79.15 ± 64.65	33.51	0.76	0.00040	WQEEMELYR
5489	728.37	57.22	CO9_HUMAN	Complement component C9 precursor (contains: complement component C9a; complement component C9b)	22.00 ± 8.48	15.71 ± 4.73	32.3	0.75	0.00000	AIEDYINEFSVR
3510	621.88	56.75	CO9_HUMAN	Complement component C9 precursor (contains: complement component C9a; complement component C9b)	34.36 ± 18.88	18.48 ± 8.90	26.47	0.77	0.00003	LSPIYNLVPVK
3401	618.36	59.06	APOA1_HUMAN	Apolipoprotein A-I precursor (Apo-AI) (ApoA-I) (contains: apolipoprotein A-I [1-242])	15.13 ± 4.59	25.00 ± 13.78	25.32	0.77	0.00022	DLATVYVDVLK
393	626.82	50.77	APOA1_HUMAN	Apolipoprotein A-I precursor (Apo-AI) (ApoA-I) (contains: apolipoprotein A-I [1-242])	45.28 ± 25.93	88.60 ± 65.21	25.1	0.79	0.00005	VQPYLDDFQK
3615	743.04	48.25	STX16_HUMAN	Syntaxin-16 (Syn16)	16.95 ± 14.15	9.81 ± 3.18	24.74	0.75	0.02199	RPPPKWVDGVDEIQYDVGR
348	615.86	61.43	APOA1_HUMAN	Apolipoprotein A-I precursor (Apo-AI) (ApoA-I) (contains: apolipoprotein A-I [1-242])	55.73 ± 34.80	109.93 ± 75.99	23.96	0.77	0.00008	QGLLPVLESFK
5664	622.33	57.48	PDGFD_HUMAN	Platelet-derived growth factor D precursor (PDGF D)	25.96 ± 9.12	21.18 ± 7.81	21.52	0.66	0.00001	LIFVYTLICANFCSCR
11415	1057.01	48.26	EP400_HUMAN	E1A-binding protein p400 (EC 3.6.1.-) (p400 kDa SWI2/SNF2-related protein)	11.57 ± 4.37	8.79 ± 2.29	18.02	0.75	0.01037	GRPPIATFSANPEAKAAAAPFQTSQASASAPR
2259	941.45	60.01	KCD10_HUMAN	BTB/POZ domain-containing protein KCTD10	71.39 ± 18.04	61.47 ± 17.33	16.49	0.66	0.00010	EEMSGESVVSSAVPAAATR
6778	804.81	46.45	DIP2A_HUMAN	Disco-interacting protein 2 homolog A	10.73 ± 5.51	7.57 ± 1.59	16.21	0.77	0.00017	KAVLSMNGLSYGVIR
7297	804.98	46.47	K0460_HUMAN	Uncharacterized protein KIAA0460	9.99 ± 5.47	6.87 ± 1.49	16.02	0.78	0.00005	AAGGGGGSSKASSSSASSAGALESSLDR
7749	662.33	46.17	ADPPT_HUMAN	L-Aminoadipate-semialdehyde dehydrogenase-phosphopantetheinyl transferase (EC 2.7.8.-)	10.70 ± 2.16	14.59 ± 5.05	15.92	0.77	0.00206	FTNKEWETIR

†Data are presented as mean ± SD. ‡Paired *t*-test. AUC, area under curve; RT, retention time.

components available in R-project (http://www.r-project.org/).[11] A composite index of two markers was generated using the results of multivariate logistic regression analysis, which also enabled calculation of the ROC curve.

Results

Identification of biomarker candidates. We compared plasma proteome data between 31 colorectal cancer patients and 59 healthy volunteers (Table 1) using 2DICAl. A total of 94 803 independent peptide peaks were detected across the 90 cases (Fig. 1a) and numbered from ID1 to ID94803. Ninety MS peaks showed a statistically significant difference between healthy controls and colorectal cancer patients (maximum mass peak intensity > 10 [arbitrary unit] and $P \leq 0.0001$ [Paired t-test] or AUC ≥ 0.75). Twenty-nine peaks were further selected by visual inspection (Fig. 1b) and subjected to MS/MS analysis. MS/MS spectra obtained from 17 peptide peaks matched 10 proteins deposited in the NCBInr database (Table 2). We selected two proteins, Apo AI and C9, for further analyses because the same protein annotation was obtained from at least two independent peaks.

Identification and differential expression of Apo AI and C9 were confirmed by immunoblotting plasma samples of representative cases (Fig. 1c). Apo AI was downregulated in cancer patients, while C9 was upregulated.

Establishment of RPPM. For the rapid selection and validation of plasma biomarker candidates, we established a high-density protein microarray platform. Plasma samples were serially diluted and printed in quadruplicate onto a hydrophobic glass surface in a format of 6144 spots within an area of 17.65 × 34.57 mm. The location of each spot was visualized by staining human IgG (green), and the relative amounts of Apo A1 and C9 proteins were quantified by hybridization with antibodies (red). Fluorescent signal intensity showed linearity in the plasma dilution range of ×32 to ×4096 in a quality control experiment (Fig. 2a) and was highly reproducible among four independent experiments (Fig. 2b). Over 78% of spots showed coefficients of variation (CV) values of < 0.1 (Fig. 2c).

Validation of biomarker candidates. In order to examine the significance of Apo AI and C9 in a larger cohort, we used RPPM, onto which plasma samples were spotted in a high-density manner. The plasma samples were serially diluted and randomly printed four times onto a microarray. Figure 3a depicts the entire image of RPPM stained with anti-C9 antibody. Representative blots of colorectal cancer and control samples are shown in Figure 3b. The RPPM revealed that Apo AI was downregulated in colorectal cancer patients compared with healthy controls, and the difference between colorectal cancer and healthy controls was statistically significant ($P = 0.000794$; Fig. 3c). C9 was significantly upregulated in colorectal cancer ($P = 1.43 \times 10^{-12}$; Fig. 3d). The results of RPPM were well correlated with those of immunoblot assay (Fig. S1), thus confirming the preciseness of RPPM.

The AUC values of Apo AI and C9 were 0.621 and 0.730, respectively (Fig. 3e,f). Although the level of plasma Apo AI was significantly different between colorectal cancer patients and healthy controls, the utility of Apo AI as a biomarker of colorectal cancer seems to be limited due to its relatively low AUC value. Statistically significant differences for upregulation of C9 were observed in patients with stage I, II, III and IV[18] colorectal cancer ($P = 3.02 \times 10^{-3}$, 1.13×10^{-5}, 5.22×10^{-8} and 3.45×10^{-4}, respectively; Fig. 3g). The AUC values of C9 for the early (stages 0–II) and advanced (stages III and IV) colorectal cancer patients were 0.667 and 0.862, respectively.

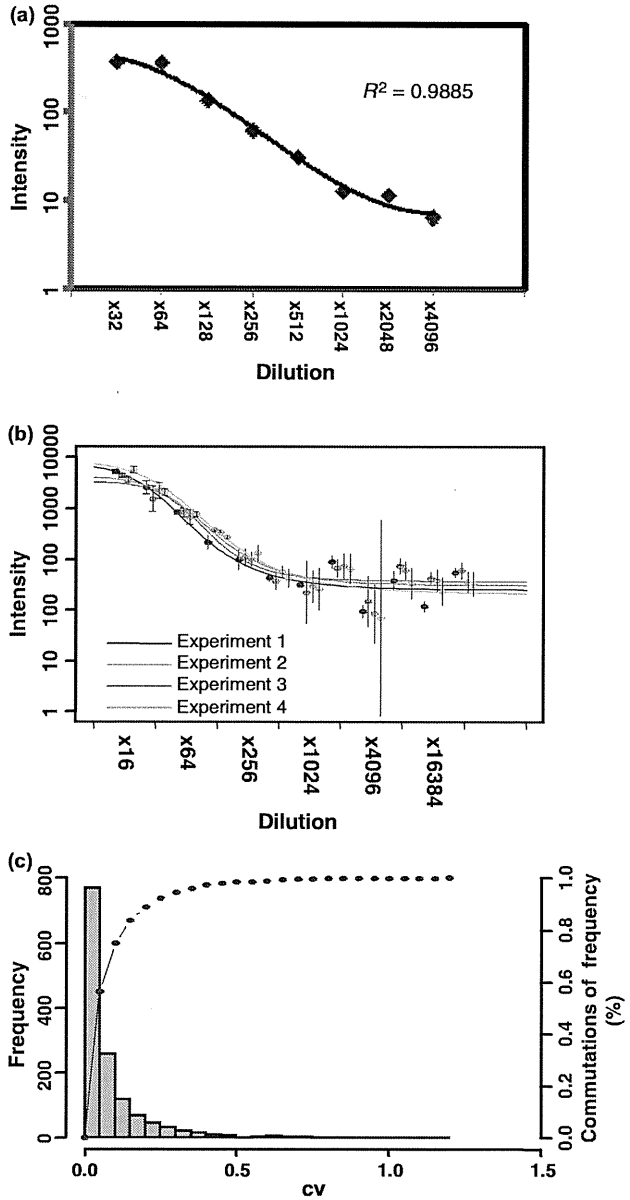


Fig. 2. Reproducibility of reverse-phase plasma microarray (RPPM). (a) Linearity of RPPM. A plasma sample was serially diluted from 32 to 4096 folds and spotted onto a ProteoChip glass. The glass was stained with anti-apolipoprotein AI antibody as described in the Materials and Methods. The mean fluorescence intensity of quadruplicates (vertical axis in arbitrary units) is plotted against plasma dilution (horizontal axis). (b) Reproducibility of RPPM. A plasma sample was serially diluted from 16 to 32 768 folds and processed as described above. The same experiment was conducted four times (Experiments 1–4), and their dilution curves overlapped. The dots represent the median intensity of quadruplicates. Bars represent interquartile ranges (IQR). (c) Distribution of the coefficients of variation (CV) values among quadruplicates in the 6144 spots (1536 quadruplicates) of RPPM stained with anti-Apo AI antibody. The dotted line represents cumulative frequency (%; right vertical axis label).

We measured the plasma level of CEA in 42 samples (20 healthy controls and 22 colorectal cancer patients: stage 0 [$n = 2$], stage I [$n = 5$], stage II [$n = 5$], stage III [$n = 5$] and stage IV [$n = 5$]), for which the residual sample volume was sufficient for the measurement with enzyme-linked immunosorbent assay (ELISA). The AUC values of CEA for the early

doi: 10.1111/j.1349-7006.2010.01818.x