BRIEF COMMUNICATIONS

Serum Proteomic Patterns for Detection of Prostate Cancer

Emanuel F. Petricoin III, David K. Ornstein, Cloud P. Paweletz, Ali Ardekani, Paul S. Hackett, Ben A. Hitt, Alfredo Velassco, Christian Trucco, Laura Wiegand, Kamillah Wood, Charles B. Simone, Peter J. Levine, W. Marston Linehan, Michael R. Emmert-Buck, Seth M. Steinberg, Elise C. Kohn, Lance A. Liotta

Pathologic states within the prostate may be reflected by changes in serum proteomic patterns. To test this hypothesis, we analyzed serum proteomic mass spectra with a bioinformatics tool to reveal the most fit pattern that discriminated the training set of sera of men with a histopathologic diagnosis of prostate cancer (serum prostate-specific antigen [PSA] ⩾4 ng/mL) from those men without prostate cancer (serum PSA level <1 ng/mL). Mass spectra of blinded sera (N = 266) from a test set derived from men with prostate cancer or men without prostate cancer were matched against the discriminating pattern revealed by the training set. A predicted diagnosis of benign disease or cancer was rendered based on similarity to the discriminating pattern discovered from the training set. The proteomic pattern correctly predicted 36 (95%, 95% confidence interval [CI] = 82% to 99%) of 38 patients with prostate cancer, while 177 (78%, 95% CI = 72% to 83%) of 228 patients were correctly classified as having benign conditions. For men with marginally elevated PSA levels (4–10 ng/mL; n = 137), the specificity was 71%. If validated in future series, serum proteomic pattern diagnostics may be of value in deciding whether to perform a biopsy on a man with an elevated PSA level. [J Natl Cancer Inst 2002;94:1576–8]

Each organ and tissue perfused by blood can contribute to, modify, or remove circulating proteins or peptides. Consequently, the serum proteome may reflect the abnormality or pathologic state of organs and tissues. Surface-enhanced laser desorption ionization time-of-flight (SELDI-TOF) mass spectrometry has been used successfully to analyze microdissected prostate cells (described in detail at http://clinicalproteomics.steem.com) *(1)* and to identify individual serum biomarkers *(2)* for prostate cancer. We have recently combined SELDI-TOF serum profiling with artificial intelligence analysis to move beyond single-marker discovery to proteomic pattern analysis *(3)*. We investigated the relationship between benign or malignant prostate disorders diagnosed histopathologically and proteomic patterns in serum. We trained our bioinformatics algorithm *(1)* using a training set of sera from known case subjects with benign prostate conditions or with prostate cancer. We then tested the pattern discovered in the training set against a blinded sample set (N = 266) that included men with serum PSA level greater than or equal to 4 ng/mL who had had a biopsy (serum PSA level ⩾4 ng/mL) to obtain a histologic diagnosis. The goal of this study was to evaluate the ability of proteomic pattern diagnostics to detect and discriminate prostate cancer from benign prostate conditions in men with normal or elevated PSA levels.

A prostate cancer screening trial (1996 through 2001) was approved by the Institutional Review Board (IRB) at the Catholic University of Chile (Santiago, Chile). Asymptomatic men 50 years of age or older with no previous history of prostate cancer were eligible. On study entry in 1996, patients provided serum samples and then had a digital rectal exam (DRE). If a serum PSA level was greater than or equal to 4 ng/mL or if the results of the DRE were suspicious, the patient underwent a single sextant biopsy set (six simultaneous biopsies). Subjects who had prostate cancer were identified as having stage I, II, or III cancer and Gleason scores of 4–9. An additional 20 serum specimens were collected at the National Cancer Institute and recoded to preserve subject identity under the approval of the IRB. Six of 20 serum samples were obtained from normal healthy male volunteers (serum PSA level <1 ng/mL). Fourteen were matched serum samples from seven men undergoing radical prostatectomy for organ-confined cancer. Serum samples were obtained preoperatively and 6 weeks postoperatively. Twenty-five additional samples were obtained from the Simone Protective Cancer Institute (Lawrenceville, NJ) under IRB approval from asymptomatic men (⩾50 years old, serum PSA level <1 ng/mL) prior to DRE.

Sera were thawed, applied to a C16 hydrophobic interaction protein chip (Ciphergen Biosystems, Freemont, CA), and analyzed as previously described *(1–3)*. Our analytical tool *(3)* combines elements from genetic algorithms first described by Holland *(4)* and cluster analysis methods from Kohonen *(5)*. The input data for analysis were ASCII files of proteomic spectra generated by SELDI-TOF. Each spectrum is composed of peak amplitude measurements at approximately 15 200 points defined by a corresponding *m/z* value. The output of the bioinformatics tool is a pattern defined by the combination of relative amplitudes that acts as a discriminator

Affiliations of authors: E. F. Petricoin III, P. S. Hackett, L. Wiegand, K. Wood, Food and Drug Administration (FDA)–National Cancer Institute (NCI) Clinical Proteomics Program, Department of Therapeutic Proteins/Center for Biologics Evaluation and Research (CBER), FDA, Bethesda, MD; D. K. Ornstein, Division of Urology, University of North Carolina, Chapel Hill; C. P. Paweletz, FDA–NCI Clinical Proteomics Program, Department of Therapeutic Proteins/CBER and Laboratory of Pathology, Center for Cancer Research (CCR), NCI, National Institutes of Health (NIH), Bethesda, and Department of Chemistry, Georgetown University, Washington, DC; A. Ardekani, FDA–NCI Clinical Proteomics Program, Department of Therapeutic Proteins/CBER Laboratory of Pathology, CCR, NCI, NIH; B. A. Hitt, P. J. Levine, Correlogic Systems, Inc., Bethesda; A. Velassco, C. Trucco, Department of Urology, Catholic University of Chile, Santiago; C. B. Simone, Simone Protective Cancer Institute, Lawrenceville, NJ; W. M. Linehan, Urologic Oncology Branch, NCI, NIH; M. R. Emmert-Buck, E. C. Kohn, L. A. Liotta, FDA–NCI Clinical Proteomics Program, Laboratory of Pathology, CCR, NCI, NIH; S. M. Steinberg, Biostatistics and Data Management Section, CCR, NCI, NIH.

Correspondence to: Emanuel F. Petricoin III, Ph.D., Bldg. 29A, Rm. 2D12, 8800 Rockville Pike, Bethesda, MD 20892 (e-mail: petricoin@cber.fda.gov).

See "Note" following "References."

for the training dataset cohorts. Analysis is divided into two phases: 1) training with known serum samples, and 2) testing validation with blinded sera. (A detailed description of the methodology and presentation of all the raw spectra can be found at http://clinicalproteomics.steem.com.)

Spectra generated from blinded sera were compared with two populations in the training set: patients with biopsy-proven cancer and those with benign prostate conditions. The algorithm used an iterative searching process to identify a subset of key amplitude values along the *x*-axis of the spectrum. This subset of key *m/z* values completely segregated the sera from men with benign conditions from that of cancer patients. Unknown test serum spectra not included in the training set were analyzed in phase 2. Only the key subset of the *m/z* amplitude values identified in phase 1 was used for classification. Each unknown was classified into two possible categories: 1) cancer or 2) benign condition.

The chi-square test was used to test the statistical significance of the classification of cancer versus benign, according to PSA levels. To compare correctness of the classifications when stratified by PSA level, an exact test for homogeneity of odds ratios was used *(6).* All *P* values are two-tailed.

The training set is outlined in Table 1. The optimal discriminatory pattern discovered during training consisted of the combined relative amplitudes at seven *m/z* values (2092, 2367, 2582, 3080, 4819, 5439, and 18 220). A representative spectra comparison of sera from men with benign versus cancerous conditions at *m/z* = 3080 is shown in Fig. 1. The blinded test set is shown in Table 2 (N = 266). The algorithm correctly classified prostate cancer in 36 (95%, 95% CI = 82% to 99%) of 38 case subjects in the blinded test set, including 18 (90%, 95% CI = 68% to 99%) of 20 sera with PSA levels of 4–10 ng/mL. The algorithm also correctly classified prostate cancer in seven (100%, 95% CI = 59% to 100%) of seven stage I and 29 (94%, 95% CI = 79% to 99%) of 31 stage II and III case subjects. Seventy (93%, 95% CI = 85% to 98%) of 75 asymptomatic men whose PSA levels were less than 4 ng/mL, including 31 with PSA levels less than 1 ng/mL, were correctly classified as having benign conditions. Ninety-seven (71%, 95% CI = 62% to 78%) of 137 men whose PSA levels were in the indeterminate range (4–10 ng/mL) were correctly classified as having benign conditions.

Ten (63%, 95% CI = 35% to 85%) of 16 men with high PSA levels (>10 ng/mL) and a negative biopsy were classified correctly as having benign conditions. Statistical analysis by using homogeneity of odds ratios verified that the proteomic classification patterns were independent of intrinsic PSA levels (*P* = 1). Complex serum proteomic patterns may reflect the underlying pathologic state of a solid organ such as the prostate. The pattern changes dis-

Table 1. Sera used for training set*

Disease status	N
No evidence of disease and PSA level ⩽1 ng/mL	25
Biopsy-proven prostate cancer and PSA level ⩾4 ng/mL	31
Total	56

*All patients were asymptomatic at the time sera were taken (1996) and 50 years old or older. N = number of independent subject serum specimens.

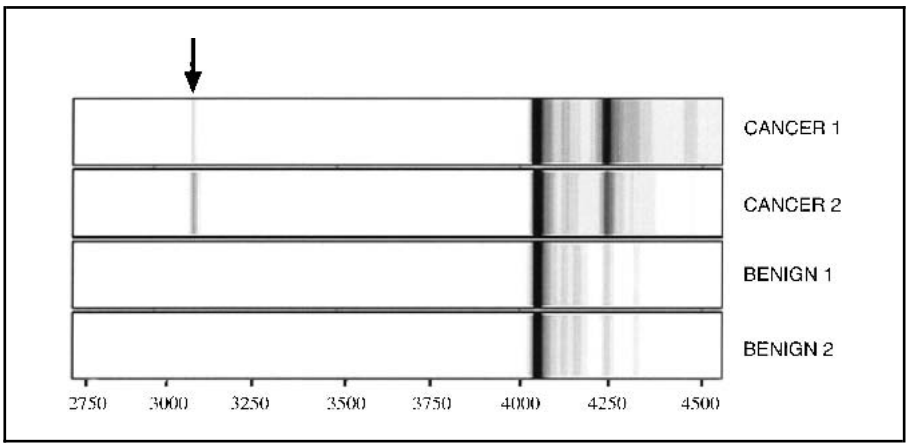

Fig. 1. Proteomic spectra. Representative surface-enhanced laser desorption ionization time-of-flight (SELDI-TOF) mass spectra (gel view) from sera of two prostate cancer patients (**top**) are compared with spectra from a set of sera from two men with benign conditions (**bottom**). A serum protein spectrum in our study is composed of *m/z* amplitudes over a range of 0–20 000. A small window (*m/z* values from 2750 to 4500) surrounding one of the seven discriminatory *m/z* values (3080) within the pattern is shown. As indicated by the **arrow,** the peak ion signature value appears distinctly different between sera from men with benign conditions and cancer.

Table 2. Comparison of the actual histopathologic diagnosis following a single sextant biopsy set with the predicted diagnosis from proteomic pattern analysis of patients' serum samples obtained prior to biopsy*

		Predicted diagnosis by proteomic pattern analysis	
Actual histopathologic diagnosis	N	Cancer N (%)	Benign N (%)
Prostate cancer†	38	36 (95)	2 (5)
Stage I	7	7	0
Stage II and III	31	29	2
Benign disease PSA level, ng/mL‡			
<4	75	5 (7)	70 (93)
4–10	137	40 (29)	97 (71)
>10	16	6 (37)	10 (63)

*Data are from a blinded test series. PSA = prostate-specific antigen.

†Includes seven patients with organ-confined cancer who were correctly classified by proteomic pattern analysis as having cancer before surgery and then correctly classified as being free of prostate cancer 6 weeks after surgical treatment. PSA levels for patients correctly classified as having cancer: 16 of 16 patients, >10 ng/mL; 18 of 20 patients, 4–10 ng/mL; 2 of 2 patients, <4 ng/mL.

‡Thirty-one patients had serum PSA levels of less than 1 ng/mL and did not have a biopsy. Benign pathology included benign prostatic hyperplasia (BPH) and prostatitis. Five-year follow-up data were available on 70 patients with BPH and/or prostatitis. Seven of 70 patients were found to have developed prostate cancer after 1998. All seven were correctly predicted as having cancer by proteomic pattern analysis of the initial screening serum samples obtained in 1996.

covered in this study set could classify prostate cancer with 95% sensitivity, including sera from men whose PSA levels were in the indeterminate range of 4–10 ng/mL ($P<.001$; Table 2). Specificity was 71% for men with PSA levels greater than or equal to 4 ng/mL. Such specificity does not support the application of this technology as a replacement for biopsy. Nevertheless, this test has high clinical relevance as a potential secondary screen for men who have marginally elevated PSA levels. Fifty-one of the 197 men (26%) with a biopsy diagnosis of benign prostatic hyperplasia (BPH) or prostatitis were classified, based on their serum proteomic patterns, as having cancer. These may not all be false positives. Some men may harbor occult cancer missed by the sextant biopsy series. Over 20% of subjects diagnosed as cancer-free on the first biopsy have been found to have cancer on the subsequent biopsy *(7,8).* To determine whether some of the men with benign conditions actually had prostate cancer, available follow-up data were obtained on 70 of the 197 men. Seven (10%) of the 70 subjects developed prostate cancer during the 5-year follow-up period. All seven men who subsequently developed cancer were correctly classified as having cancer, using the 1996 enrollment serum, including four who had PSA levels of less than 10 ng/mL on enrollment. The data presented in Table 2 reflect this classification of cancer for these seven subjects who subsequently developed cancer.

The data in Table 2 support the hypothesis that pathologic conditions confined to the prostate modify the serum proteomic pattern. Indeed, the signature patterns in all seven of the men that had elective prostatectomies for organ-confined cancer reverted from a cancer classification to a benign classification following removal of the prostate (Table 2). The proteins comprising the optimal discriminatory pattern could be derived from the organ itself or from the host and may be proteins (e.g., cytokines and chemokines), metabolites, or enzymatic cleavage products. Proteomic patterns can distinguish prostate cancer even when PSA levels are in the indeterminate range. Thus, serum protein pattern analysis has the potential to supplement or complement medical decisions that are now based on physical examination, imaging, and serum PSA levels *(7).* A major clinical application of proteomic analysis would be its use in the clinical management of men with marginally elevated PSA levels. Currently, PSA level is commonly used to determine the need for prostate biopsy in asymptomatic men. Unfortunately the specificity of elevated PSA levels in this cohort is 25%–35%; therefore, 70%–75% of men undergoing biopsy because of an abnormal PSA level do not have prostate cancer *(8).* Therefore, we envision that serum proteomic pattern analysis may be used in the future to aid clinicians so that fewer men are subjected to unnecessary biopsies, while the detection of curable cancer in the affected cohort is not compromised, because biopsy will remain the gold standard for confirmation of the presence of cancer.

REFERENCES

(1) Paweletz CP, Gillispie JW, Ornstein DK, Simone NL, Brown MR, Cole KA, et al. Rapid protein display profiling of cancer progression directly from human tissue using a protein biochip. Drug Dev Res 2000;49:34–42.

(2) Xiao Z, Adam BL, Cazares LH, Clements MA, Davis JW, Schellhammer PF, et al. Quantitation of serum prostate-specific membrane antigen by a novel protein biochip immunoassay discriminates benign from malignant prostate disease. Cancer Res 2001;61:6029–33.

(3) Petricoin EF, Ardekani A, Hitt P, Levine P, Fusaro VA, Steinberg S, et al. Use of proteomic patterns in serum to identify ovarian cancer. Lancet 2002;359:572–7.

(4) Holland JH, editor. Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. 3rd ed. Cambridge (MA): MIT Press; 1992.

(5) Kohonen T. Self-organized formation of topologically correct feature maps. Biol Cybern 1982;43:59–69.

(6) Breslow NE, Day NE. Statistical methods in cancer research. Volume I - The analysis of case-control studies. IARC Sci Publ 1980;(32):5–338.

(7) Arcangeli CG, Ornstein DK, Keetch DW, Andriole GL. Prostate-specific antigen as a screening test for prostate cancer. The United States experience. Urol Clin North Am 1997;24:299–306.

(8) Keetch DW, Catalona WJ, Smith DS. Serial prostatic biopsies in men with persistently elevated serum prostate specific antigen values. J Urol 1994;151:1571–4.

NOTE

Manuscript received March 11, 2002; revised July 9, 2002; accepted August 9, 2002.