Short Communication

Proteomic analysis of nipple aspirate fluid to detect biologic markers of breast cancer

ER Sauter*,[1], W Zhu[1], X-J Fan[2], RP Wassell[2], I Chervoneva[3] and GC Du Bois[2]

[1]Department of Surgery, Thomas Jefferson University, 1025 Walnut St., Ste 605, Philadelphia, PA 19107, USA; [2]Department of Microbiology and Immunology, Kimmel Cancer Center of Thomas Jefferson University, Philadelphia, PA 19107, USA; [3]Department of Biostatistics, Kimmel Cancer Center of Thomas Jefferson University, Philadelphia, PA 19107, USA

The early detection of breast cancer is the best means to minimise disease-related mortality. Current screening techniques have limited sensitivity and specificity. Breast nipple aspirate fluid can be obtained noninvasively and contains proteins secreted from ductal and lobular epithelia. Nipple aspirate fluid proteins are breast specific and generally more concentrated than corresponding blood levels. Proteomic analysis of 1 μl of diluted nipple aspirate fluid over a 5–40 kDa range from 20 subjects with breast cancer and 13 with nondiseased breasts identified five differentially expressed proteins. The most sensitive and specific proteins were 6500 and 15 940 Da, found in 75–84% of samples from women with cancer but in only 0–9% of samples from normal women. These findings suggest that (1) differential expression of nipple aspirate fluid proteins exists between women with normal and diseased breasts, and (2) analysis of these proteins may predict the presence of breast cancer.

Standard screening for breast cancer involves physical examination and mammography. These tools alert the physician to the presence of a mass which is palpated and/or an abnormality which is visualised. Treatment of breast cancer requires a morphologic diagnosis characterised by visual changes in the nuclei in a cytologic or histologic preparation of breast cells. In order to obtain the cells for review by the pathologist, a diagnostic needle, core, or surgical biopsy must be performed. These procedures are painful, the needle and core biopsies are subject to sampling error, and only approximately 15–20% of the procedures detect malignancy (Watson *et al*, 1999). Non- and minimally invasive procedures are presently under review to detect breast cancer without submitting the subject to an invasive diagnostic procedure. These procedures include nipple aspiration, ductal lavage, and ductoscopy. Of these, only nipple aspiration is totally noninvasive and requires a device with minimal cost. In addition, only nipple aspiration provides concentrated secreted proteins undiluted with irrigation fluid which is required to perform ductal lavage and ductoscopy.

While cytologic and histologic evaluation are the gold standards to detect breast cancer, recent advances in comprehensive molecular technologies allow the simultaneous analysis of multiple protein expression targets. Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and mass spectrometry have been used to perform proteomic analysis of human breast epithelium for comparison of cancer with normal breast (Czerwenka *et al*, 2001; Paweletz *et al*, 2001), changes in protein synthesis after growth factor induction (Vercoutter-Edouart *et al*, 2001), to detect new breast cancer markers and proteins involved in cell pathways such

as signalling and the humoural response to antigens (Hondermarck *et al*, 2001; Le Naour *et al*, 2001), and to develop a breast cancer protein expression map database (Harris *et al*, 2002). Technologies in development include protein chips using antibodies or aptamers (short strings of DNA or RNA that bind with high affinity to specific target proteins) as capture molecules (Abbot, 2002).

A wide array of proteins are secreted into and highly concentrated in NAF and have been associated with breast cancer. We hypothesised that the proteome wide analysis of NAF using surface-enhanced laser desorption/ionization time of flight-mass spectrometry (SELDI-TOF) would identify one or more proteins differentially expressed in women with breast cancer compared to subjects without disease. The SELDI technique can be performed with 1 μl of NAF, can detect components in the high femtomole range, and the chip surface which allows the rapid evaluation of 8–24 samples has high throughput potential. Candidate breast cancer biomarkers can be identified using SELDI with a rapid immunoassay.

MATERIALS AND METHODS

After informed consent of an Institutional Review Board approved protocol, NAF was collected using a modified breast pump (Sauter *et al*, 1996) and coded so that all analyses were blinded. Specimens were collected from women 35–81 years old scheduled to undergo surgery for a suspected malignancy in the aspirated breast, as well as from women without evidence of disease (Table 1). NAF was diluted with Tris buffer to a concentration of 3.6 mg total protein ml^{-1} Tris buffer. Proteome analysis with the SELDI PBSII system (Ciphergen Biosystems, Fremont, CA, USA) was carried out on three chips, normal phase (NP), hydrophobic (H4) and anion exchange (SAX), using 1 μl of NAF. The chromatographic surfaces of these ProteinChips allow the capture of generic proteins (NP),

Table I Patient demographics

Menopausal status	Cancer	No cancer
Pre-	9	10
Post-	11	3
Race		
White	17	9
Nonwhite	3	4
Age		
Median	52	44
Range	33–81	29–55
Ever pregnant		
Yes	17	8
No	3	5
Median age (years) at 1st pregnancy	27	21
Ever used birth control pills		
Yes	11	8
No	9	5
Ever used hormone replacement medications		
Yes	6	1
No	14	12

proteins having exposed hydrophobic surfaces (H4), and proteins binding by anion exchange (SAX). NP and SAX chips were washed first with binding buffer containing detergent, then with detergent-free buffer and water to remove nonspecific proteins. H4 chips were loaded with 10% acetonitrile, then washed with this solvent and water.

RESULTS

A representative ProteinChip SELDI MS analysis of NAF demonstrates discrete peaks (Figure 1) at 6500 and 8000 Da from three of four tumour bearing women which are not present or are present at a reduced level in NAF from women with normal breast tissue. We considered a protein expressed if the value assigned by SELDI software had a signal-to-noise ratio $\geqslant 3:1$. Protein peaks from different samples were judged similar if the values in Da were within 0.05% of each other. The figure also demonstrates a distinct peak at 15 940 Da in four out of four NAF samples from women with breast cancer that is not observed in NAF from normal subjects. Two larger proteins (a broad-based peak centered at 28 100 Da and a peak at 31 770, not shown) were also identified in a high percentage of women with breast cancer but were absent or present at a reduced level in NAF from normal women.
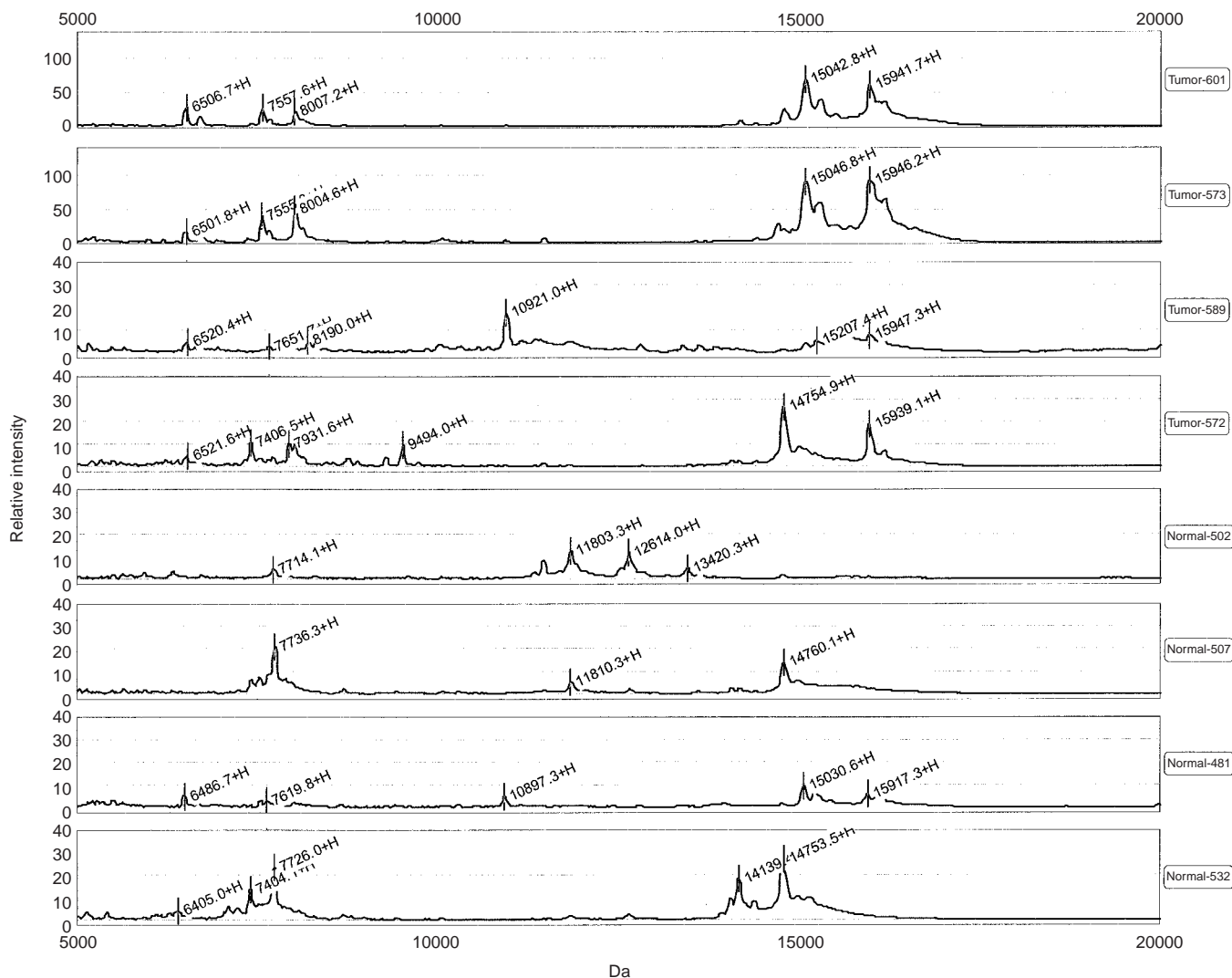
Figure I SELDI-TOF profile (5–20 kDa) from eight NAF samples. The top four lanes are from subjects with breast cancer.

Molecular and Cellular Pathology

Percentages of positively expressed proteins in NAF from women with and without cancer and odds ratios (ORs) with 95% confidence intervals (CIs) are listed in Table 2. Because of the small sample size, all estimates and P-values were computed using exact methods (LogXact 4.0, Cytel Software Corporation, Cambridge, MA, USA). Protein peaks at 15 940 Da in all chips, at 8000 Da in the H4 and NP chips, at 6500 Da in the H4 and SAX chips, and at 28 100 Da in the NP chip were expressed significantly more often in NAF samples from subjects with breast cancer than from normal subjects (Table 2). For each chip, maximum sensitivity and specificity were achieved with the single protein peak: 15 940 Da in the H4 and NP chips, and 6500 Da in the SAX chip (Table 2, bolded rows). These proteins were separately evaluated in multivariable logistic regression models controlling for age, race, past pregnancy, and past use of oral contraceptives. The adjusted odds ratios were: 29.2 (95% CI: 4.44 to ∞, $P < 0.001$) for the 15 940 Da protein using the H4 chip, 27.3 (95% CI: 3.93 to ∞, $P < 0.001$) for the 6500 Da using the SAX chip, and 18.7 (95% CI: 2.52 to 6.36, $P = 0.002$) for the 15 940 Da using the NP chip.

DISCUSSION

While both RNA and protein profiling can be applied to tissue samples, analysis of body fluids such as NAF is restricted to proteomics due to their low cellularity (Kennedy, 2001). Proteomic analysis of breast tissue has been reported using 2-D PAGE (Hondermarck et al, 2001). Although sensitive and powerful, 2-D PAGE is a labour intensive and low throughput method, has the drawback of selecting against proteins which are extremely acidic or basic, and there is a detection bias toward highly abundant proteins. The recently developed SELDI technique (Davies et al, 1999) allows for the rapid profiling of extracts from cells, tissues, and physiological fluids and can screen large numbers of samples in a clinical setting by differential protein capture according to chemical (ionic, hydrophobic, hydrophilic or metal ion affinity) surfaces on a protein chip.

Of the differentially expressed proteins identified using SELDI MS, the 6500 and 15 940 Da proteins are particularly striking since they were detected in a high percentage of subjects with breast cancer but in no one (6500 using the SAX chip and the 15 940 using the NP chip) or very few women without breast cancer. The identity of the 6500 Da protein is not known but may represent epithelin, while the 8000 Da peak may by mammaglobin. Mammaglobin has been reported to be a breast cancer marker (Watson et al, 1999). The identity of the 15 940 protein has not been determined. The 31 770 peak is most likely a dimer of the 15 940 protein.

The identity of the 28 100 Da peak may represent one or more members of the kallikrein family. A recent report using SELDI MS (Watkins et al, 2001) analysed the sera of 46 women with breast cancer and 23 controls. A 28.3 kDa protein was found in 100% of women with invasive and 80% with non-invasive breast cancer, respectively, but in only 4% of women without disease. This protein may correspond to the 28.1 kDa peak we observed in NAF, although confirmation is required. The putative breast cancer marker proteins that we have detected in this study may represent different isoforms of the candidate proteins or other known proteins. We are currently characterising the MS peaks identified in this study to determine their protein signature. We have also instituted a prospective analysis of NAF using SELDI-TOF to confirm the differential expression of these proteins in breast cancer vs normal breast, and plan to use a bioinformatics approach to identify the proteomic patterns in NAF that distinguish cancer from normal (Petricoin et al, 2002).

In conclusion, proteomic analysis of NAF from subjects with and without breast cancer identified five differentially expressed proteins. The most sensitive and specific proteins were 6500 and 15 940 Da, found in 75 – 84% of samples from women with cancer but in only 0 – 9% of samples from normal women. While current tools to detect breast cancer are widely available and clinically effective, they have limited sensitivity and specificity. Analysis of NAF proteins secreted from the breast epithelium may increase our ability to detect breast cancer at its earliest stages when used in combination with mammography and physical examination.

Table 2 Protein expression profile in NAF (by size) obtained from normal women and from women with breast cancer, using three different SELDI ProteinChips (NP,SAX, and H4)

Chip	Protein peak	Positive in tumour (sensitivity)		Positive in normal (1-specificity)		OR (95% CI)	P-value
NP							
	6500 Da	12/19	63%	3/11	27%	4.33 (0.72, ∞)	0.128
	8000 Da	15/19	79%	1/11	9%	31.73 (3.11, ∞)	<0.001
	15 940 Da	**16/19**	**84%**	**1/11**	**9%**	**43.10 (4.01, ∞)**	**<0.001**
	28 100 Da	8/19	42%	0/11	0%	9.61 (1.37, ∞)	0.026
	31 770 Da	6/19	32%	0/11	0%	6.16 (0.77, ∞)	0.091
SAX							
	6500 Da	**15/20**	**75%**	**0/12**	**0%**	**54.67 (7.26, ∞)**	**<0.001**
	8000 Da	4/20	20%	0/13	0%	2.91 (0.24, 160)	0.661
	15 940 Da	11/20	55%	0/13	0%	19.04 (2.61, ∞)	0.002
	28 100 Da	8/20	40%	2/13	15%	3.53 (0.53, 41.3)	0.264
	31 770 Da	0/20	0%	0/13	0%	n/e	
H4							
	6500 Da	13/20	65%	2/13	15%	9.44 (1.46, 111)	**0.013**
	8000 Da	12/20	60%	2/13	15%	7.71 (1.20, 90.0)	0.027
	15 940 Da	**16/20**	**80%**	**0/13**	**0%**	**72.02 (9.32, ∞)**	**<0.001**
	28 100 Da	5/19	26%	0/12	0%	5.63 (0.69, ∞)	0.116
	31 770 Da	1/19	5%	0/12	0%	0.68 (0.02, ∞)	1.000

OR: odds ratio. CI: confidence interval. n/e: not estimable.

REFERENCES

Abbot A (2002) Betting on tomorrow's chips. *Nature* **415:** 112–114

Czerwenka KF, Manavi M, Hosmann J, Jelincic D, Pischinger KI, Battistutti WB, Behnam M, Kubista E (2001) Comparative analysis of two-dimensional protein patterns in malignant and normal human breast tissue. *Cancer Detect Prev* **25:** 268–279

Davies H, Lomas L, Austen B (1999) Profiling of amyloid beta peptide variants using SELDI Protein Chip arrays. *Biotechniques* **27:** 1258–1261

Harris RA, Yang A, Stein RC, Lucy K, Brusten L, Herath A, Parekh R, Waterfield MD, O'Hare MJ, Neville MA, Page MJ, Zvelebil MJ (2002) Cluster analysis of an extensive human breast cancer cell line protein expression map database. *Proteomics* **2:** 212–223

Hondermarck H, Vercoutter-Edouart AS, Revillion F, Lemoine J, el-Yazidi-Belkoura I, Nurcombe V, Peyrat JP (2001) Proteomics of breast cancer for marker discovery and signal pathway profiling. *Proteomics* **1:** 1216–1232

Kennedy S (2001) Proteomic profiling from human samples: the body fluid alternative. *Toxicol Lett* **120:** 379–384

Le Naour F, Misek DE, Krause MC, Deneux L, Giordano TJ, Scholl S, Hanash SM (2001) Proteomics-based identification of RS/DJ-1 as a novel circulating tumor antigen in breast cancer. *Clin Cancer Res* **7:** 3328–3335

Paweletz CP, Trock B, Pennanen M, Tsangaris T, Magnant C, Liotta LA, Petricoin III EF (2001) Proteomic patterns of nipple aspirate fluids obtained by SELDI-TOF: potential for new biomarkers to aid in the diagnosis of breast cancer. *Dis Markers* **17:** 301–307

Petricoin III EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, Steinberg SM, Mills GB, Simone C, Fishman DA, Kohn EC, Liotta LA (2002) Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* **359:** 572–577

Sauter ER, Daly M, Linahan K, Ehya H, Engstrom PF, Bonney G, Ross EA, Yu H, Diamandis E (1996) Prostate-specific antigen levels in nipple aspirate fluid correlate with breast cancer risk. *Cancer Epidemiol Biomarkers Prev* **5:** 967–970

Vercoutter-Edouart AS, Czeszak X, Crepin M, Lemoine J, Boilly B, Le Bourhis X, Peyrat JP, Hondermarck H (2001) Proteomic detection of changes in protein synthesis induced by fibroblast growth factor-2 in MCF-7 human breast cancer cells. *Exp Cell Res* **262:** 59–68

Watkins B, Szaro R, Ball S, Knubovets T, Briggman J, Hlavaty JJ, Kusinitz F, Stieg A, Wu Y-Y (2001) Detection of early-stage cancer by serum protein analysis. *American Laboratory* **33:** 32–36

Watson MA, Dintzis S, Darrow CM, Voss LE, DiPersio J, Jensen R, Fleming TP (1999) Mammaglobin expression in primary, metastatic, and occult breast cancer. *Cancer Res* **59:** 3028–3031

Molecular and Cellular Pathology