CANCER GENOMICS & PROTEOMICS 6: 177-182 (2009)

Diagnosis of Breast Cancer by Tear Proteomic Pattern

ANTJE LEBRECHT[1], DANIEL BOEHM[1], MARKUS SCHMIDT[1], HEINZ KOELBL[1], RAPHAELA L. SCHWIRZ[1,2] and FRANZ H. GRUS[2]

[1]Department of Obstetrics and Gynecology, and [2]Experimental Ophthalmology, Department of Ophthalmology, Johannes Gutenberg University Mainz, D-55101 Mainz, Germany

Abstract. *Background: Early detection of breast cancer reduces breast cancer-related mortality. Breast cancer biomarkers offer a promising means of detecting this disease at the earliest and most treatable stages. Patients and Methods: The aim of this study was to generate a protein biomarker profile in tear fluid for breast cancer patients. This established biomarker profile was then used to discriminate between cancer patients and healthy controls. Potential biomarkers were screened in tear fluid from 50 women with breast cancer and 50 healthy women, matched for age. Tear fluid was drawn prior to surgery. Surface-enhanced laser desorption-ionisation time-of-flight mass spectrometry was used for protein profiling with two different active surfaces on the protein chips: a cationic exchanger (CM-10) and a reverse-phase surface (H50). The data were analyzed by multivariate statistical techniques and artificial neural networks. Results: A total of 404 peaks were found with different molecular weights at different laser intensities and a statistically significant ($p<0.05$) panel with 20 biomarkers was generated. Use of the biomarker panel resulted in 71.19% of the samples being correctly classified as cancer samples (42 out of 59) and 70.69% as control samples (41 out of 58), thus overall 70.94% were correctly classified. The diagnostic pattern was able to differentiate cancer patients from healthy women with a specificity and sensitivity of approximately 70% using tear fluid. Conclusion: In this study a biomarker panel in tear fluid was successfully generated to allow breast cancer patients to be discriminated from healthy women. The study suggests that the proteomic pattern of tear fluid may be useful in the diagnosis of breast cancer and for high-throughput biomarker discovery.*

Correspondence to: Antje Lebrecht, MD, Johannes Gutenberg University Mainz, Department of Obstetrics and Gynecology, Langenbeckstrasse 1, D-55101 Mainz, Germany. Tel: +49 6131177311, Fax: +49 613117472609, e-mail: lebrecht@uni-mainz.de

Key Words: SELDI-TOF-MS, breast cancer, biomarkers, proteomics, tear fluid.

Breast cancer is the most frequent neoplasm and the leading cause of cancer mortality in women worldwide (1). Screening mammography is the most sensitive and reliable method for early detection (2). Unfortunately, small lesions are frequently missed and may not be visible even by mammography, particularly in young women or women with dense breast tissue (3). Additionally, for many women, breast compression during mammography is an uncomfortable and painful experience (2, 4, 5). This can lead to the decision of their not returning for a future screening round (2, 4, 6).

As yet, serum tumor markers of breast cancer such as CEA and CA 15.3 has not been successful in molecular diagnosis of breast cancer in early stages (7). Mass spectrometry technologies offer the opportunity to discover novel biomarkers by protein profiling (8). Surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF MS; Ciphergen, Fremont, CA, USA) separates proteins on different solid-phase chromatographic surfaces and the proteins are subsequently ionized and detected by TOF-MS. Our previous work indicated that SELDI-TOF-MS may be a useful tool in the development of tear biomarkers in breast cancer (9). Beside our findings, other authors have also shown that ProteinChip technology is able to analyze protein profiles from a varity of complex biological materials, *e.g.* serum, tears, saliva, urine, and cell lysates (10).

Altered protein profiles were found in tear fluid in some systemic diseases, *e.g.* cancer (11-15), diabetes mellitus (16-17), Sjögren syndrome (18), atypical rosacea (19). The tear film is a very complex mixture of proteins, lipids, glycoproteins, neuropeptides and cytokines and is far from being just a simple reflection of blood plasma (20-23). In particular, the lack of highly abundant proteins such as albumin and immunoglobulin in tears allows purification and preparation of the samples to be much reduced as compared to that for other biological materials, *e.g.* serum (9).

The aim of the study was to generate a protein biomarker profile from tear fluid for breast cancer patients. We subsequently used this established biomarker profile to discriminate between cancer patients and healthy controls.

Patients and Methods

Patient characteristics. This prospective study included 100 individuals: 50 patients with breast cancer and 50 healthy controls matched for age. Informed consent was obtained from all participating in the study and the protocols were approved by the institutional Ethics Committee and conformed to the provisions of the Declaration of Helsinki.

Breast cancer stages pT1, pT2, and pT3 were seen in 26, 19, and 1 cases, respectively. Axillary lymph node dissection was performed in all 50 cancer patients. Lymph node metastases were found in 17 cases. Histologically, 8, 20, and 18 tumors were graded as well, moderately and poorly differentiated, respectively. The median age of the patients at the time of diagnosis was 63 years (range 37-86 years).

Tear fluid was drawn 24-48 hours prior to surgery by Schirmer strips. The Schirmer strips were stored at −80˚C until examination. The Schirmer strips were eluted overnight in 500 μl 0.1% dodecylmaltoside and then for two hours in 300 μl 0.1% trifluoroacetic acid (TFA) resulting in two elutions, a TFA elution and a dodecyl elution. For automatic handling of all binding and washing steps with the TFA elution, a robotic laboratory automation station (Biomek 2000; Beckman Coulter, Fullerton, CA, USA) was used (9). Sodium acetate (pH 4) was used as a washing buffer. The washing steps with the dodecyl eluate were performed without using the Biomek.

The two eluates were assayed on two different chromatographic surfaces: a week cation exchange surface (CM10) and a reversed-phase surface (H50). All ProteinChip® Arrays (Ciphergen Biosystems, Inc., Fremont, CA, USA) were pretreated according to the standard protocols of the manufacturer. Binding buffers were 5% acetonitrile/0.1% TFA (H50), and 20 mM sodium acetate buffer (pH 5; CM10). Twenty microliters of the TFA elution and 20 μl of binding buffer were applied to each spot using the Biomek. The arrays were incubated on a DPC shaking platform for 1 hour, and the solution in each well was removed by the Biomek. All wells were washed with sodium acetate (pH 4) followed by a wash step with distilled water for 5 minutes to remove buffer salts. After the wells were dry, the Biomek was used to apply 2 μl of the saturated sinapinic acid solution (an energy-absorbing molecule) in 50% acetonitrile and 0.5% TFA to each spot. After the spots were air dried, each spot was analyzed in a ProteinChip Reader. Each sample (2 μl per spot) was bound to each array surface in duplicate on separate arrays and BioProcessors.

The dodecyl eluate (2 μl per spot) was applied directly to the ProteinChips. The washing steps with sodium acetate (pH 4) and distilled water were carried out as mentioned above.

Data acquisition and preprocessing. ProteinChip Arrays were analyzed on a PBS-IIc ProteinChip Reader equipped with a ProteinChip Array AutoLoader using the ProteinChip Software version 3.2 (Ciphergen Biosystems, Inc., Fremont, CA, USA) The AutoLoader is able to analyze up to 24 ProteinChip Arrays (192 spots) at one time. Each array was read at two laser intensities: low intensity optimized for low molecular mass proteins and high intensity for high molecular mass proteins. The high-intensity protocol averaged 195 laser shots from each spot with a laser intensity of 190, a deflector setting of 3,000 Da, a detector sensitivity of 9, and a molecular mass detection range of 2,000 to 200,000 Da. For low-intensity measurements, the laser intensity was set at 180 and the deflector set at 1500 Da. The raw data were transferred to the CiphergenExpress Data Manager Software version 2.1 (CE; Ciphergen Biosystems) for analysis.

Data analysis. The CE data manager software was used to normalize the spectra, to automatically detect peaks, and to create the peak cluster lists. The peak intensities were normalized according to the total ion current. The cluster lists were exported as ASCII files to a statistical analysis program (Statistica, ver. 6.2; StatSoft, Tulsa, OK, USA). The cluster lists contained normalized peak intensity values for each sample within a group. Based on these normalized peak intensities, probabilities based on *t*-tests and multivariate discriminant analysis were calculated.

The separation between clinical groups for diagnostic purposes can generally be enhanced by using multimarker panels of protein biomarkers rather than single biomarkers. The statistical software package was used to perform a multivariate discriminant analysis based on combinations of multiple biomarker peaks. The discriminant analysis selected the 20 most important protein biomarkers for best discrimination between the groups.

To assess the diagnostic power of this biomarker panel, a neural network was trained using these markers as input neurons. The traditional artificial neural network (ANN) model is the most widely used today and is the multiple-layer feed-forward network (MLFN) with the back-propagation training algorithm (24). The network was trained by the biomarker peaks and the output neurons defined whether the protein pattern classified the patient as having breast cancer or not. To assess the performance of the network, the data set was randomly divided in two parts: The first half of the patients and controls was used as training set, the second as a test set. After completion of training, the success of the algorithm was tested using the test set, which comprised patients and controls to which the neural network had not previously been exposed.

The software generated a receiver operating characteristic (ROC) curve by plotting sensitivity against 1-specificity.

Results

Tear samples were analyzed on two different ProteinChip Array chemistries as described. Figures 1 and 2 demonstrate SELDI-TOF mass spectra of tear fluid that illustrates differences in profiles corresponding to two patient groups by different laser intensities. We found 404 peaks with different molecular weights at different laser intensities. From these we were able to establish a panel with 20 biomarkers (Figure 3) with significant differences between the tear protein profiles of control and cancer samples ($p<0.05$). Their intensities are shown in Table I.

As Figure 4 exemplarily shows for one biomarker, the spread around the average was small. The result of the ANOVA is shown in Table II. The *p*-values for 15 biomarkers were $p<0.05$ and for five biomarkers were $p<0.059$. For the discriminant analysis we present seven of

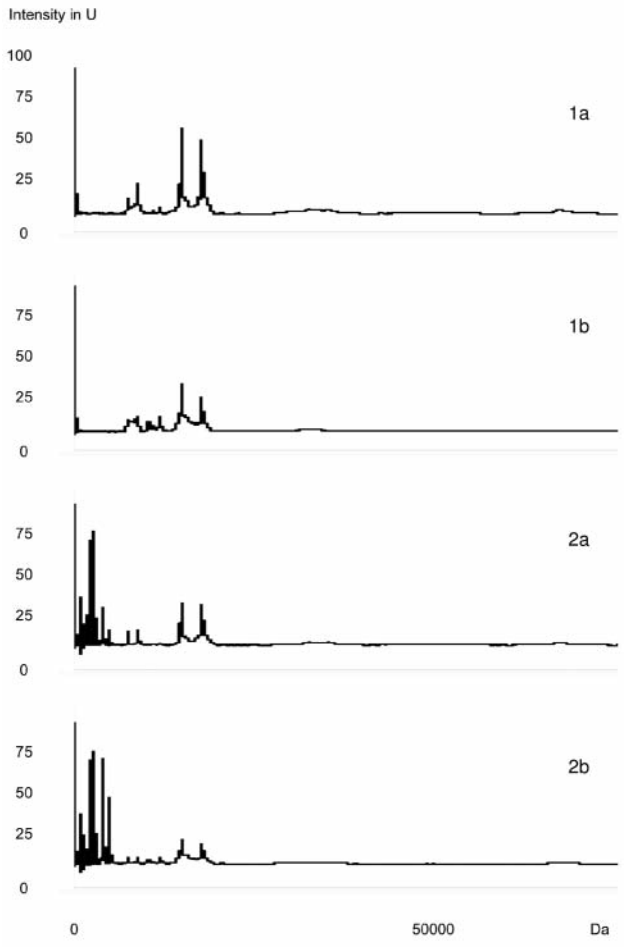

Figure 1. *SELDI-TOF mass spectra of tear protein profiles of two healthy individuals (a and b) analyzed with a low (1) and a high (2) laser energy level.*

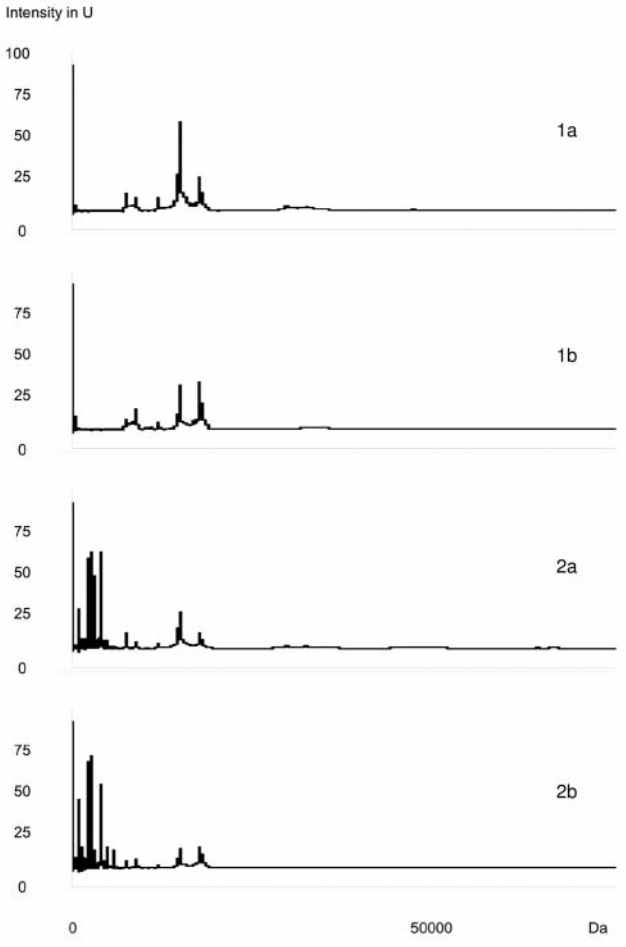

Figure 2. *SELDI-TOF mass spectra of tear protein profiles of two breast cancer patients (a and b) analyzed with a low (1) and a high (2) laser energy level.*

the 20 biomarkers exemplarily in Table III. The correlation coefficient was between 0.29 and 0.43. So each marker contributes moderately to the separation of the two groups. The Malanobis distance was 1.038. This means that the biomarker panel was statistically significantly able to discriminate between healthy and cancer patients ($p<0.005$). The classification matrix (Table IV) demonstrates that 71.19% of the samples were correctly classified as cancer samples (42 of 59) and 70.69% of the samples correctly as control samples (41 of 58); overall 70.94% of the samples were correctly classified.

Using this biomarker panel for input, an artificial neural network was trained with a training data set. The performance of the trained net was assessed using a test data set. Figure 5 show the ROC curve with an AUC of 0.75 and a specificity and sensitivity of approximately 70% .

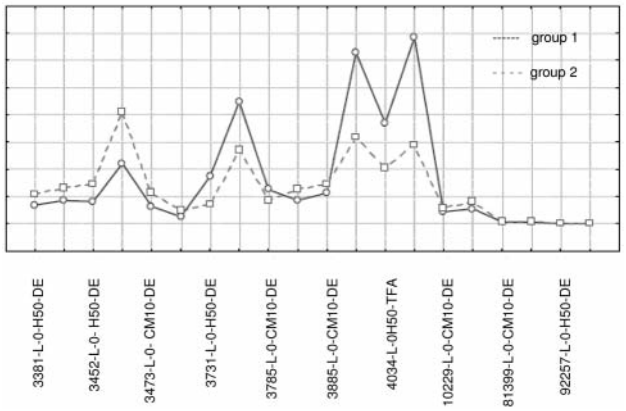

Figure 3. *Protein profiles of tear samples from cancer patients (group 1) and controls (group 2).*

Table I. *Twenty significant biomarkers: comparison of intensities of the detected biomarkers in controls and cancer patients.*

Biomarker (molecular weight - protein chip surface - eluate)	Biomarker intensity	
	Cancer patients	Controls
3381-L-O-H50-dodecyl-eluate	1.385	2.150
3409-L-O-CM10-dodecyl-eluate	1.697	2.621
3452-L-O-H50-dodecyl-eluate	1.633	2.944
3458-L-O-CM10-dodecyl-eluate	4.436	8.215
3473-L-O-CM10-dodecyl-eluate	1.277	2.317
3629-L-O-CM10-dodecyl-eluate	0.522	1.011
3731-L-O-H50-dodecyl-eluate	3.475	1.442
3736-L-O-CM10-dodecyl-eluate	8.972	5.452
3785-L-O-CM10-dodecyl-eluate	2.513	1.718
3847-L-O-CM10-TFA	1.722	2.577
3885-L-O-CM10-dodecyl-eluate	2.266	2.915
4031-L-O-H50-dodecyl-eluate	12.561	6.352
4034-L-O-H50-TFA	7.373	4.102
4038-L-O-CM10-dodecyl-eluate	13.704	5.806
10229-L-O-CM10-dodecyl-eluate	0.873	1.133
10883-L-O-CM10-dodecyl-eluate	1.111	1.627
81399-L-O-CM10-dodecyl-eluate	0.139	0.178
83094-L-O-CM10-dodecyl-eluate	0.129	0.168
92257-L-O-H50-dodecyl-eluate	0.025	0.026
146095-N-O-CM10-dodecyl-eluate	0.027	0.037

Table II. *Result of the ANOVA comparing controls with patients and the p-values of the biomarkers.*

Biomarker	*p*-Value
4038-L-0-CM10-dodecyl-eluate	0.0026
3736-L-0-CM10-dodecyl-eluate	0.0082
3785-L-0-CM10-dodecyl-eluate	0.0114
3731-L-0-H50-dodecyl-eluate	0.0124
3629-L-0-CM10-dodecyl-eluate	0.0178
92257-L-0-H50-dodecyl-eluate	0.0203
3409-L-0-CM10-dodecyl-eluate	0.0305
10883-L-0-CM10-dodecyl-eluate	0.0331
3473-L-0-CM10-dodecyl-eluate	0.0400
3458-L-0-CM10-dodecyl-eluate	0.0410
4031-L-0-H50-dodecyl-eluate	0.0418
3452-L-0-H50-dodecyl-eluate	0.0435
81399-L-0-CM10-dodecyl-eluate	0.0441
10229-L-0-CM10-dodecyl-eluate	0.0476
83094-L-0-CM10-dodecyl-eluate	0.0492
4034-L-0-H50-TFA	0.0511
3847-L-0-CM10-TFA	0.0534
146095-N-0-CM10-dodecyl-eluate	0.0560
3381-L-0-H50-dodecyl-eluate	0.0585
3885-L-0-CM10-dodecyl-eluate	0.0600

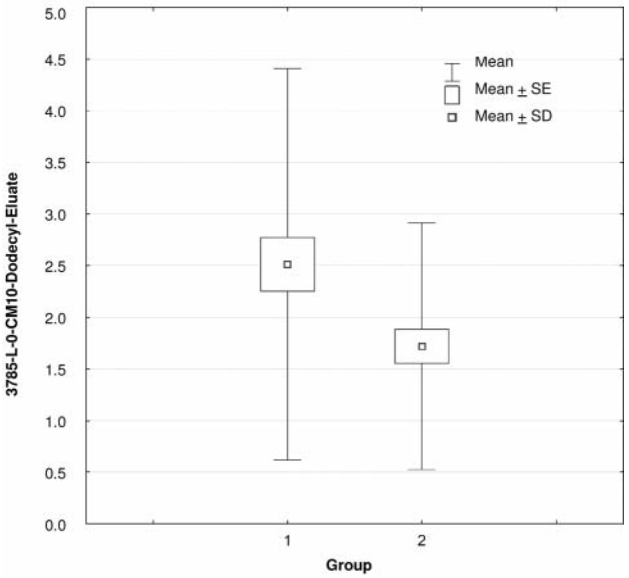

Figure 4. *Box plots for the mean value of the biomarker 3785-L-O-CM10-dodecyl-eluate: 2.51 for cancer patients (1) and 0.17 for controls (2).*

Table III. *Example of discriminant analysis for seven biomarkers.*

Biomarker	Standardised canonical correlation coefficient
Variable	Root 1
4038-L-O-CM10-dodecyl-eluate	0.29
3629-L-O-CM10-dodecyl-eluate	–0.36
92257-L-O-H50-dodecyl-eluate	–0.34
3785-L-O-CM10-dodecyl-eluate	0.43
3847-L-O-CM10-TFA	–0.35
3885-L-O-CM10-dodecyl-eluate	–0.31
3452-L-O-H50-dodecyl-eluate	–0.30

Table IV. *Classification matrix.*

	Classification matrix		
	%	1 (cancer patients)	2 (controls)
1 cancer	71.19	42	17
2 control	70.69	17	41
Overall	70.94	59	58

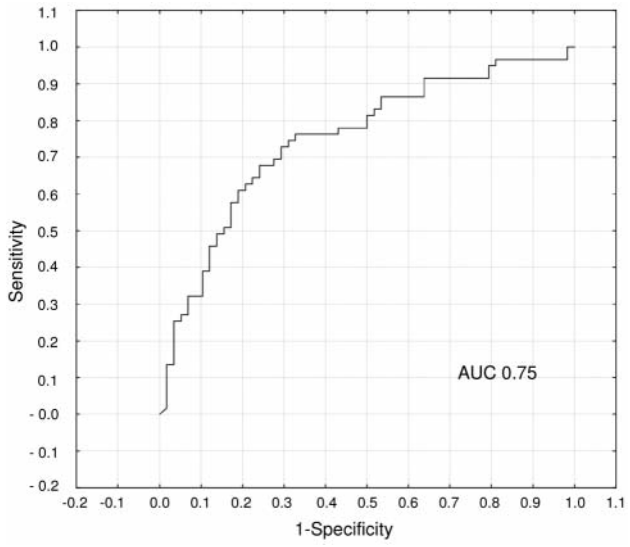

Figure 5. *ROC curves of tear samples.*

Discussion

Early detection of breast cancer reduces breast cancer-related mortality (25-26). Breast cancer biomarkers offer a promising means of detecting this disease at the earliest and most treatable stages (27). Because of the heterogenous nature of the most tumors, no single tumor marker provides the sensitivity or specificity required for widespread screening (28). Recent advances in image analysis, automated mass spectrometry and bioinformatics have provided the tools necessary for high-throughput biomarker discovery (29). Several thousand proteins from multiple patients can simultaneously be compared and analyzed to arrive at distinct phenotypic proteomic profiles that are associated with the malignant state (27).

Compared to serum, the analysis of tear fluid offers several opportunities. First it is easy to obtain with the help of Schirmer strips or capillaries and it is easy to store pending analysis. Secondly, tear fluid does not need any purification or depletion of highly abundant proteins. This together with the SELDI-TOF-MS technology offers the opportunity to find and also analyse proteins with a molecular weight lower than 8 kDa, as we did in our biomarker panel.

Furthermore, SELDI-TOF-MS technology allows rapid sample analysis because very small sample volumes can be directly applied to the ProteinChip Array surfaces and the process can easily be automated for high-throughput analysis (30).

To discriminate breast cancer patients from healthy women, some authors screened nipple aspirate fluid (28), serum (27, 31) and saliva (32-33) to establish a characteristic biomarker profile. Pawlik *et al*. (28) and Li *et al*. (31) found biomarkers with similar molecular weight in nipple aspirate fluid and serum, respectively, as we did. Serum protein profiles in spectra to classify breast cancer vs. normal benign breast disease showed 90% sensitivity and 94% specificity (8).

In the first step of our study, we successfully generated a biomarker panel in tear fluid from breast cancer patients. We found an acceptable discrimination between breast cancer patients and healthy controls. With the help of neural networks this panel of 20 biomarkers shows a sensitivity and specificity of approximately 70% . At the present time we do not know the identity of these ion signals. After validating the biomarker profiles in a larger group of patients, we will identify the proteins.

Acknowledgements

This investigation was supported in part by the Maifor programme of the Johannes Gutenberg University, Mainz, Germany.

References

1 Parkin DM, Bray F, Ferlay J and Pisani P: Global cancer statistics. CA Cancer. J Clin *55*: 74-108, 2005.
2 Keemers-Gels ME, Groenendijk RP, van den Heuvel JH, Boetes C, Peer PG and Wobbes TH: Pain experienced by women attending breast cancer screening. Breast Cancer Res Treat *60*: 235-240, 2000.
3 Antman K and Shea S: Screening mammography under age 50. JAMA *281*: 1470-1472, 1999.
4 Jackson VP, Lex AM and Smith DJ: Patient discomfort during screen-film mammography. Radiology *168(2)*: 421-423, 1988.
5 Cockburn J, Cawson J, Hill D and De Luise T: An analysis of attenders at an Australian pilot breast cancer screening programm. Australas Radiol *36*: 115-119, 1992.
6 Aro AR, Absetz-Ylostalo P, Eerola T, Pamilo M and Lonnqvist J: Pain and discomfort during mammography. Eur J Cancer *32A(10)*: 1674-1679, 1996.
7 Cheung KL, Graves CR and Robertson JF: Tumour marker measurements in the diagnosis and monitoring of breast cancer. Cancer Treat Rev *26(2)*: 91-102, 2000
8 Clarke CH, Buckley JA and Fung ET: SELDI-TOF-MS proteomics of breast cancer. Clin Chem Lab Med *43(12)*: 1314-1320, 2005.
9 Lebrecht A, Boehm D, Schmidt M, Koelbl H and Grus FH: Surface-enhanced laser desorption/ionisation time-of-flight mass spectrometry to detect breast cancer markers in tears and serum. Cancer Genomics Proteomics *6*: in press, 2009.
10 Grus FH, Joachim SC and Pfeiffer N: Analysis of complex autoantibody repertoires by surface-enhanced laser desorption/ionization-time of flight mass spectrometry. Proteomics *3*: 957-961, 2003.
11 Petricoin EF 3rd, Ornstein DK, Paweletz CP, Ardekani A, Hackett PS, Hitt BA, Velassco A, Trucco C, Wiegand L, Wood K, Simone CB, Levine PJ, Linehan WM, Emmert-Buck MR, Steinberg SM, Kohn EC and Liotta LA: Serum proteomic patterns for detection of prostate cancer. J Natl Cancer Inst *94*: 1576-1578, 2002.

12 Semmes OJ, Feng Z, Adam BL, Banez LL, Bigbee WL, Campos D, Cazares LH, Chan DW, Grizzle WE, Izbicka E, Kagan J, Malik G, McLerran D, Moul JW, Partin A, Prasanna P, Rosenzweig J, Sokoll LJ, Srivastava S, Srivastava S, Thompson I, Welsh MJ, White N, Winget M, Yasui Y, Zhang Z and Zhu L: Evaluation of serum protein profiling by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry for the detection of prostate cancer: I. Assessment of platform reproducibility. Clin Chem *51*: 102-112, 2005.
13 Petricoin EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, Steinberg SM, Mills GB, Simone C, Fishman DA, Kohn EC and Liotta LA: Use of proteomic patterns in serum to identify ovarian cancer. Lancet *359*: 572-577, 2002.
14 Chen YD, Zheng S, Yu JK and Hu X: Artificial neural networks analysis of surface-enhanced laser desorption/ionization mass spectra of serum protein pattern distinguishes colorectal cancer from healthy population. Clin Cancer Res *10*: 8380-8385, 2004.
15 Yu JK, Cheng YD and Zheng S: An integrated approach to the detection of colorectal cancer utilizing proteomics and bioinformatics. World J Gastroenterol *10*: 3127-3131, 2004.
16 Herber S, Grus FH, Sabuncuo P and Augustin AJ: Changes in the tear protein patterns of diabetic patients using two-dimensional elecrophoresis. Adv Exp Med Biol *506*: 623-626, 2002.
17 Stolwijk TR, Kuizenga A van Haeringen NJ, Kijlstra A, Oosterhuis JA and van Best JA: Analysis of tear fluid proteins in insulin-dependent diabetes mellitus. Acta Ophthalmol *72*: 357-362, 1994.
18 Tomosugi N, Kitagawa K, Takahashi N, Sugai S and Ishkawa I: Diagnostic potential of tear proteomic patterns in Sjögren's syndrome. J Proteome Res *4*: 820-825, 2005.
19 An HJ, Ninonuevo M, Agulan J, Liu H, Lebrilla CB, Alvarenga LS and Mannis MJ: Glycomics analyses of tear fluid for the diagnostic detection of ocular rosacea. J Proteome Res *4*: 1981-1987, 2005.
20 Janssen PT and van Bijsterveld OP: Origin and biosynthesis of human tear fluid proteins. Invest Ophthalmol Vis SCI *24*: 623-630, 1983.
21 Sariri R and Ghafoori H: Tear proteins in health, disease, and contact lens wear. Biochemistry *73*: 381-392, 2008.
22 Van Haeringen NJ: Clinical biochemistry of tears. Surv Ophthalmol *26*: 84-96, 1981.
23 Tiffany JM: The normal tear film. Dev Ophthalmol *41*: 1-20, 2008.
24 Rumelhart D, Hinton G and Williams J: Learning representations by back-propagating procedures. Nature *323*: 533-536, 1986.
25 Francisci S, Capocaccia R, Grande E, Santaquilani M, Simonetti A, Allemani C, Gatta G, Sant M, Zigon G, Bray F and Janssen-Heijnen M; EUROCARE Working Group: The cure of cancer: a European perspective. Eur J Cancer *45(6)*: 1067-1079, Epub 2009.
26 Smith RA, Cokkinides V and Eyre HJ: American Cancer Society. American Cancer Society guidelines for the early detection of cancer. CA Cancer J Clin *54*: 41-52, 2004.
27 Davis MA and Hanash S: High-throughput genomic technology in research and clinical management of breast cancer. Plasma-based proteomics in early detection and therapy. Breast Cancer Res *8*: 217, 2006.
28 Pawlik TM, Fritsche H, Coombes KR, Xiao L, Krishnamurthy S, Hunt KK, Pusztai L, Cheng J, Clarke CH, Arun B, Hung M and Kuerer HM: Significant differences in nipple aspirate fluid protein expression between healthy women and those with breast cancer demonstrated by time-of-flight mass spectrometry. Breast Cancer Res Treat *89(2)*: 149-157, 2005.
29 Kuerer HM, Goldknopf IL, Fritsche H, Krishnamurthy S, Sheta EA and Hunt KK: Identification of distinct protein expression patterns in bilateral matched pair breast ductal fluid specimens from women with unilateral invasive breast carcinoma. Cancer *95*: 2276-2282, 2002.
30 Wulfkuhle JD, Paweletz CP, Steeg PS, Petricoin EF and Liotta L: Proteomic approaches to the diagnosis, treatment, and monitoring of cancer. Adv Exp Med Biol *532*: 59-68, 2003.
31 Li J, Zhang Z, Rosenzweig J, Wang YY and Chan DW: Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. Clin Chem *48*: 1296-1304, 2002.
32 Streckfus CF, Bigler LR and Zwick M: The use of surface-enhanced laser desorption/ionization time-of-flight mass spectrometry to detect putative breast cancer markers in saliva: a feasibility study. J Oral Oathol Med *35*: 292-300, 2006.
33 Streckfus CF and Bigler LR: Saliva as a diagnostic fluid. Oral Disease *8*: 69-76, 2002.

Received February 15, 2009
Revised April 27, 2009
Accepted May 14, 2009