available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Short Communication

A comparative proteinomic analysis of nipple aspiration fluid from healthy women and women with breast cancer

J.L. Noble[a,b,d], R.S. Dua[a,b,d], G.R. Coulton[c], C.M. Isacke[b], G.P.H. Gui[a,*]

[a]Department of Academic Surgery (Breast Unit), The Royal Marsden NHS Foundation Trust, Fulham Road, London SW3 6JJ, UK
[b]Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, 237 Fulham Road, London SW3 6JB, UK
[c]Division of Basic Medical Sciences and Division of Cardiac and Vascular Sciences, St. George's Hospital Medical School, University of London, Cranmer Terrace, London SW17 0RE, UK

ARTICLE INFO

Article history:
Received 5 May 2007
Accepted 6 August 2007
Available online 27 September 2007

Keywords:
Nipple aspiration fluid
Proteomics
Breast cancer
SELDI-TOF/MS

ABSTRACT

This pilot study examines the feasibility of nipple aspiration to distinguish women with breast cancer from healthy women using surface-enhanced laser desorption ionisation time-of-flight mass spectrometry (SELDI-TOF/MS). Nipple aspiration fluid (NAF) was collected from each breast in 21 women newly diagnosed with unilateral breast cancer and 44 healthy women. No differences were found when proteomic profiles of NAF from the cancer-bearing breast and the contralateral non-cancerous breast were compared. In contrast, 9 protein peaks were significantly different between the cancer-bearing breast compared with healthy women and 10 peaks were significantly different between the contralateral healthy breast and healthy women ($P < 0.05$). These data suggest that invasive breast cancer may result in a field change across both breasts and that proteomic profiling of NAF may have more value in breast cancer risk assessment than as a diagnostic or screening tool.

1. Introduction

In breast cancer, the general approach to evaluation has become formalised as triple assessment, involving clinical examination, imaging (usually mammography and/or ultrasound) and pathology (core biopsy and/or fine needle aspiration). In breast cancer screening, although mammography remains the gold standard for breast imaging, the sensitivity varies from 62.9% to 87%.[1] The lack of sensitivity is in part due to increased breast density in younger women and in women taking hormone replacement therapy obscuring potential breast cancers.[2–4] A further limitation of conventional screening is the faster rate of tumour growth resulting in higher rates of interval cancers in younger women.[5,6] Together this has prompted the search for alternative approaches for the diagnosis of breast cancer.

The intraductal approach comprising nipple aspiration, ductal lavage and duct endoscopy is a minimally invasive technique which allows direct access to the ductal system from which the majority of breast cancers arise.[7,8] This approach could provide an alternative or adjunct to current diagnostic techniques particularly in younger women where the sensitivity of mammography is reduced. Studies using cytological analysis of nipple aspiration fluid (NAF) have found that atypical hyperplasia of breast ductal epithelium is associated with an increased risk of subsequent development of breast cancer.[9,10] The disadvantage of NAF lies in the low cellular yield making cytological analysis more

* *Corresponding author:* Tel.: +44 020 7808 2783; fax: +44 020 7808 2673.
E-mail address: gerald.gui@rmh.nhs.uk (G.P.H. Gui).
[d] These authors contributed equally to the work.
0959-8049/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2007.08.009

difficult. The emerging field of proteomics allows the identification of many proteins in a cell or biological fluid in a single experiment unlike older biochemical techniques which identify single proteins. Recent advances in molecular technologies have enabled rapid, protein biomarker identification from small volumes of NAF using surface-enhanced laser desorption ionisation time-of-flight mass spectrometry (SELDI-TOF/MS). SELDI-TOF/MS combines on-array separation of complex protein mixtures via retention coupled directly with mass spectrometry and has the advantage that its high sample throughput allows the generation of sufficient data to perform strong statistical tests. The purpose of this study was to use SELDI-TOF/MS to compare the NAF proteome from women with unilateral breast cancer and healthy women to discover differences that might be exploited in a diagnostic or screening test.

2. Materials and methods

The study protocol was approved by the Royal Marsden Hospital Committee for Clinical Research and the Research Ethics Committee.

2.1. Collection and SELDI-TOF/MS analysis of NAF

Paired NAF samples were obtained from the cancer (affected) and contralateral (unaffected) breast from 21 women recently diagnosed with unilateral invasive breast cancer and undergoing primary surgery at the Royal Marsden Hospital. NAF samples were obtained from a further 44 healthy at risk women attending the Royal Marsden Hospital for breast screening due to a family history of breast cancer. Healthy at risk women were defined by NICE guidelines (http://www.nice.org.uk/page.aspx?o=CG014&c=cancer). For both groups, the criteria for exclusion were previous breast surgery, pregnancy, breast feeding within the last 12 months, breast implants, active infection of the breast or allergy to lidocaine. For the healthy at risk women, the criteria for exclusion were an abnormal mammogram or clinical examination at the time of NAF collection. The details of patient characteristics are provided in Table 1. The volume of nipple fluid collected ranged from 0 to >100 μl, with an average of 16 μl per breast. The number of productive ducts was 1–3 per breast and where NAF was collected from more than one duct, it was pooled. For each patient, the collected NAF was stored separately from the cancer-bearing and contralateral uninvolved breasts. Samples were stored for up to 12 months and underwent a maximum of three freeze–thaw cycles. The process of collection and storage of NAF, preparation of the IMAC30 and CM10 (Ciphergen, Freemont, USA) ProteinChip arrays and the Ciphergen SELDI-TOF/MS analysis have been described previously.[11] As recommended,[12] a strict protocol for the handling, storage and analysis of NAF was adhered to.

2.2. Data analysis

The spectra from all samples were collated and normalised. Consistent peak clusters were generated using the Biomarker Wizard ProteinChip software (Ciphergen). The first pass used a signal-to-noise ratio set to 5 with a threshold of 20%. The second pass used a signal-to-noise ratio of 2 and included peaks within 0.3% of the mass of the peaks found in the first pass. Data below 3000 Da were excluded to avoid matrix noise. A majority of peaks fell within this range. A cluster graph was generated where each peak intensity was plotted against *m/z* for the three sample groups (cancer, contralateral and healthy). Comparisons

Table 1 – Patient characteristics

Characteristic	Cancer patients (%), *n* = 21	Healthy patients (%), *n* = 44
Age		
Median (years)	47	46
Range (years)	31–57	25–62
Menopausal status		
Pre (%)	14 (67)	31 (70)
Post (%)	5 (24)	13 (30)
Not known (%)	2 (9)	0
Risk		
Standard (%)	11 (52)	5 (12)
Moderate (%)	6 (29)	20 (45)
High (%)	3 (14)	19 (43)
Not known (%)	1 (5)	0
Operation		
WLE (%)	12 (57)	
Mastectomy (%)	9 (43)	
Nodal status		
Positive (%)	8 (33)	
Negative (%)	13 (62)	
Pathology		
IDC (%)	13 (62)	
ILC (%)	3 (14)	
IDC/ILC (%)	1 (5)	
Mucinous (%)	2 (9)	
DCIS (%)	1 (5)	
Ungradeable (%)	1 (5)	
Tumour grade[a]		
1 (%)	2 (9)	
2 (%)	10 (38)	
3 (%)	7 (33)	
High grade DCIS (%)	1 (5)	
Ungradeable (%)	1 (5)	
Tumour size		
<2 cm	4	
2–5 cm	11	
>5 cm	5	
Not known	1	
Extensive intraductal component		
Yes	6	
No	13	
Not known	2	
ER status		
Positive (%)	16 (76)	
Negative (%)	5 (24)	

IDC, invasive ductal carcinoma; ILC, invasive lobular carcinoma; WLE, wide local excision; ER, oestrogen receptor.
a Highest grade.

between groups were made using the Mann–Whitney *U*-test. A *P*-value of <0.05 for a given peak indicated that this peak was differentially expressed.

3. Results

Paired NAF samples were collected from 21 women with breast cancer. Of the 44 healthy at risk women, bilateral NAF samples were obtained from 27 and unilateral samples from 17 women. The two groups were well matched for age and menopausal status, but a higher percentage of women in the healthy group had a family history of breast cancer (88% versus 43%) (Table 1).

A comparison of the proteomic profiles by SELDI-TOF/MS from the cancer-bearing and contralateral breasts from the women with breast cancer revealed no peaks with *P*-values less than 0.05 on the CM10 surface. On the IMAC30 surface, there was one apparent peak at a mass of 4306 Da ($P < 0.05$) but on examination of the spectra this was seen as an amalgamation of several peaks rather than a peak in its own right (data not shown).

A comparison of NAF from the cancer-bearing breast and from healthy patients revealed nine statistically significant peaks on the CM10 surface (Fig. 1a). Of these, the peaks at 3471 Da, 3511 Da, 4151 Da, 4586 Da, 4646 Da and 4698 Da were single peaks. On close inspection putative peaks at 15,895 Da, 19,263 Da and 23,174 Da comprised multiple peaks and were therefore not included in this analysis (Table 2). On the IMAC30 array, four statistically significant peaks were identified with *P*-values <0.05 (Fig. 1a). Of these, three well discriminated peaks were seen at 3501 Da, 3627 Da and 4147 Da, whereas the peak at 23,232 Da was again a composite and hence was excluded (Table 2). Interestingly, the peaks at 4151 Da on the CM10 array and 4147 Da on the IMAC30 array were within 0.3% of each other and are likely to be the same protein.

Next, a comparison was made between NAF from the contralateral breast of women with cancer and NAF from healthy patients. On the CM10 array there were seven statistically significant peaks (Fig. 1b). The peaks at 3471 Da, 3869 Da, 4151 Da, 4646 Da and 14,720 Da were single peaks. The peaks at 19,263 and 23,174 Da were multiple peaks and hence excluded from further analysis (Table 2). Three out of the five significant single peaks in the healthy versus contralateral group were also significant in the cancer versus healthy group; two peaks at 3869 Da and 14,720 Da were not found to be significant in the cancer versus healthy group. Given the homogeneity between the cancer-affected and the contralateral breast, a large amount of overlap would be expected and indeed was observed in this analysis. On the IMAC30 array there were seven statistically significant peaks (Fig. 1b). The peaks at 3501 Da, 3627 Da, 4147 Da, 4376 Da and 5890 Da were single peaks, while the peaks at 23,232 Da and 25,784 Da were groups of peaks and were excluded. Three of the significant single peaks in the healthy versus contralateral group were also present in the cancer versus healthy group; two peaks at 4646 and 5890 Da were not found to be significant in the cancer versus healthy group (Table 2).

4. Discussion

Current diagnostic tools of breast cancer have limitations even when used in combination. The advent of new technologies, which allow the rapid identification of secreted proteins in readily accessible body fluids, may provide an alternative to the current practice. The intraductal approach, and in particular the collection of NAF, is a simple, non-invasive method which allows access to fluid from the breast ductal system. The advantage of using SELDI-TOF/MS to analyse these samples is that this high throughput technology allows statistical verification of data and good reproducibility once, as is the case here, sample preparation and handling is standardised.

This study compared NAF from 21 women with unilateral breast cancer and 44 healthy women. Examination of the paired samples from the 21 women with breast cancer revealed no statistically significant differences between the cancer-bearing breast and the contralateral non-cancerous breast. Similar findings have been reported previously[13,14] when paired NAF samples from 12 and 23 women, respectively, with unilateral breast cancer were examined. These studies and the study reported here noted that in women with unilateral breast cancer, there was a high degree of homology between breasts in the same woman. Consequently, one explanation for the inability of these studies to identify differences between the cancer-bearing and non-cancerous breast is due to the heterogeneity between patients. An alternative explanation is that the effect of breast cancer on NAF composition is systemic. To address this, NAF obtained from women with breast cancer was compared to that obtained from healthy women.

The comparison between NAF from the cancer-bearing breast and the breasts of healthy women revealed a number of statistically significant peaks. Interestingly, in the study by Pawlik et al.,[14] NAF from five healthy volunteers was compared to the NAF collected from the 23 women with breast cancer using WCX2 and IMAC3 arrays. Seventeen peaks were statistically significant between the NAF from the cancer-bearing breasts and the breasts of healthy volunteers and three statistically significant peaks were found when the contralateral non-cancerous breast was compared to the breasts of healthy volunteers. It is also notable that in a study comparing pathological nipple discharge in 114 women (27 with breast cancer and 87 benign), three proteins were found to be associated with breast cancer.[15] More recently,[16] a study has progressed to both validation and protein identification of proteins in NAF. In this study NAF from 5 women with cancer and 5 healthy women was used to find three discriminatory peaks. The peaks were validated using ductal lavage specimens and then identified as human neutrophil peptide 1–3 by database searching. Identification was confirmed using monoclonal antibodies and a prospective study is now in progress to measure the predictive value of these proteins.[16] However, there is no overlap between the proteins found in these two studies or those found in the current study. The most likely explanation for this lack of overlap is that different protocols were used in each study. In particular are the differences in the arrays surfaces which select for different

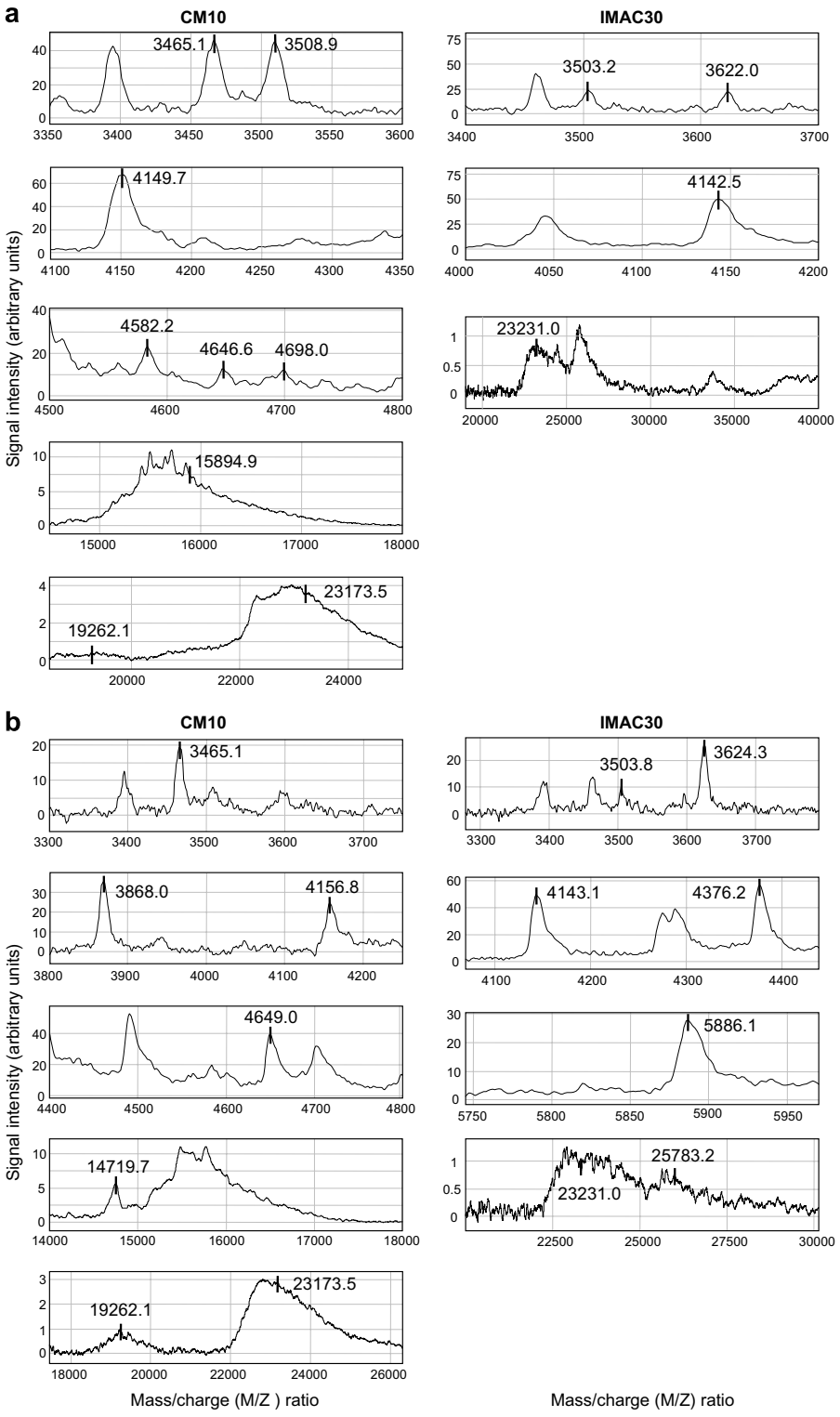

Fig. 1 – SELDI-TOF/MS spectra of NAF. (a) NAF samples from the cancer-bearing and healthy breasts were analysed on CM10 and IMAC30 arrays. Panels show selected regions of representative spectra from a single cancer NAF sample. Peaks that were statistically significant when all cancer and healthy NAF samples were compared are indicated by a line and the mass/charge (*m*/*z*) ratios for the illustrated sample are shown. *m*/*z* ratios from different spectra that are within 0.3% of each other are considered as potentially the same protein. (b) NAF samples from the healthy and contralateral breast were analysed on CM10 and IMAC30 arrays. Panels show selected regions of representative spectra from a single healthy NAF sample. Peaks that were statistically significant when all contralateral and healthy NAF samples were compared are indicated by a line and the mass/charge (*m*/*z*) ratios for the illustrated sample are shown. *m*/*z* ratios from different spectra that are within 0.3% of each other are considered as potentially the same protein.

Table 2 – Mass/charge ratio (m/z) and P-values of statistically significant peaks between the cancer-bearing breast, the contralateral non-cancerous breast and the healthy breast

Array	m/z	Cancer versus contralateral (P-value)	Cancer versus healthy (P-value)	Healthy versus contralateral (P-value)
CM10	3471	n/s	<0.001	0.046
IMAC30	3501	n/s	0.026	0.040
CM10	3511	n/s	0.014	n/s
IMAC30	3627	n/s	0.010	0.004
CM10	3869	n/s	n/s	0.043
IMAC30	4147	n/s	0.037	0.018
CM10	4151	n/s	0.006	0.018
IMAC30	4376	n/s	n/s	0.024
CM10	4586	n/s	0.004	n/s
CM10	4646	n/s	<0.001	0.029
CM10	4698	n/s	0.016	n/s
IMAC30	5890	n/s	n/s	0.011
CM10	14720	n/s	n/s	0.046

n/s, not significant.

classes of proteins, instrument settings and bioinformatic tools that were employed.[17] However, it is also important to consider other shortcomings in current proteomic-based technologies. As discussed by Diamandis,[17] the failure to identify tumour makers in the circulation has raised concern as to the sensitivity of SELDI-TOF/MS to detect low abundance proteins and that proteins which have been identified may represent acute phase reactants or epiphenomena of tumour presence rather than the products of the tumour cells themselves. It is as a result of such concerns that attention in the breast cancer community has focussed on the analysis of NAF as this involves direct sampling of the breast microenvironment so that any tumour-associated components would be present at a higher concentration compared to serum. Despite this advantage, we demonstrate here the limitation of this approach for the diagnosis or screening of women with breast cancer due to the homology of the NAF proteomic profiles between the cancer-bearing and contralateral breast. This lack of difference may result from the inherent heterogeneity between patients and problems in sensitivity associated with SELDI-TOF/MS. In addition, NAF is only collected from a subset of the ducts in the breast, and tumour-bearing ducts may be less likely to yield NAF due to physical obstruction of the lumen and duct collapse.[18] However, as this study did reveal proteins whose abundance was significantly different between the cancer-bearing breast and healthy women and also between the contralateral non-cancerous breast of women with breast cancer and healthy women, an alternative explanation is that NAF proteomic profiling reveals a systemic change in women with breast cancer.

In conclusion, the study presented here is the largest to date to compare NAF from women with breast cancer and healthy women. The findings and our identification of unique peaks add to the growing body of literature on this subject and offer some insight into NAF as a representation of the breast microenvironment. Together these data demonstrate the limitations to using the intraductal approach in the diagnosis or screening of women with breast cancer, particularly as the identification of the cancer-bearing breast may be impossible using conventional imaging techniques. However, as NAF proteomic profiling did distinguish between women with and without breast cancer, this study indicates the value of exploring the use of NAF proteomic profiling as a risk assessment tool and/or to monitor the course of disease and response to treatment. Consequently future work should focus on the predictive value of validated proteins, singly and groups, within new sample sets and then to employ complementary techniques for biomarker identification. This will allow the generation of autologous immuno-based assays that could be used for the rapid analysis of specific proteins in large sample groups and ultimately to follow women prospectively to assess the predictive value of these biomarkers.

Conflict of interest statement

The authors of this manuscript declare that there are no financial or personal relationships with other people or organisations that could inappropriately influence or bias their work.

Acknowledgements

This work was supported by Breakthrough Breast Cancer and the General Cancer Research Fund at the Royal Marsden NHS Foundation Trust. We wish to thank Ms. Marie-Catherine Le Bihan (Division of Basic Medical Sciences, SGUL) and Dr. Edward Tarelli and Mr. Glen Alder (St. George's Medical Biomics Centre) for their excellent technical support, and Sue Ashley and Yolanda Barbachano (Royal Marsden NHS Foundation Trust) for their advice and help with the statistical analysis.

REFERENCES

1. Carney PA, Miglioretti DL, Yankaskas BC, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Ann Intern Med* 2003;**138**:168–75.
2. Wolfe JN. Breast parenchymal patterns and their changes with age. *Radiology* 1976;**121**:545–52.
3. El-Bastawissi AY, White E, Mandelson MT, Taplin SH. Reproductive and hormonal factors associated with

mammographic breast density by age (United States). *Cancer Causes Control* 2000;**11**:955–63.
4. Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. Effect of age, breast density, and family history on the sensitivity of first screening mammography. *Jama* 1996;**276**:33–8.
5. Gilliland FD, Joste N, Stauber PM, et al. Biologic characteristics of interval and screen-detected breast cancers. *J Natl Cancer Inst* 2000;**92**:743–9.
6. Brekelmans CT, van Gorp JM, Peeters PH, Collette HJ. Histopathology and growth rate of interval breast carcinoma. Characterization of different subgroups. *Cancer* 1996;**78**:1220–8.
7. Dua RS, Isacke CM, Gui GP. The intraductal approach to breast cancer biomarker discovery. *J Clin Oncol* 2006;**24**:1209–16.
8. Fabian CJ, Kimler BF, Mayo MS, Khan SA. Breast-tissue sampling for risk assessment and prevention. *Endocr Relat Cancer* 2005;**12**:185–213.
9. King EB, Chew KL, Petrakis NL, Ernster VL. Nipple aspirate cytology for the study of breast cancer precursors. *J Natl Cancer Inst* 1983;**71**:1115–21.
10. Wrensch MR, Petrakis NL, King EB, et al. Breast cancer incidence in women with abnormal cytology in nipple aspirates of breast fluid. *Am J Epidemiol* 1992;**135**:130–41.
11. Noble J, Dua RS, Locke I, et al. Proteomic analysis of nipple aspirate fluid throughout the menstrual cycle in healthy pre-menopausal women. *Breast Cancer Res Treat* 2007;**104**:191–6.
12. Grizzle WE, Semmes OJ, Bigbee W, et al. The need for review and understanding of SELDI/MALDI massa spectroscopy data prior to analysis. *Cancer Inform* 2005;**1**:86–97.
13. Paweletz CP, Trock B, Pennanen M, et al. Proteomic patterns of nipple aspirate fluids obtained by SELDI-TOF: potential for new biomarkers to aid in the diagnosis of breast cancer. *Dis Markers* 2001;**17**:301–7.
14. Pawlik TM, Fritsche H, Coombes KR, et al. 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* 2005;**89**:149–57.
15. Sauter ER, Shan S, Hewett JE, Speckman P, Du Bois GC. Proteomic analysis of nipple aspirate fluid using SELDI-TOF-MS. *Int J Cancer* 2005;**114**:791–6.
16. Li J, Zhao J, Yu X, et al. Identification of biomarkers for breast cancer in nipple aspiration and ductal lavage fluid. *Clin Cancer Res* 2005;**11**:8312–20.
17. Diamandis EP. Mass spectrometry as a diagnostic and a cancer biomarker discovery tool: opportunities and potential limitations. *Mol Cell Proteom* 2004;**3**:367–78.
18. Khan SA, Wiley EL, Rodriguez N, et al. Ductal lavage findings in women with known breast cancer undergoing mastectomy. *J Natl Cancer Inst* 2004;**96**:1510–7.