available at www.sciencedirect.com www.elsevier.com/locate/molonc

Up-regulated Proteins in the Fluid Bathing the Tumour Cell Microenvironment as Potential Serological Markers for Early Detection of Cancer of the Breast

Pavel Gromov^{a,d,*}, Irina Gromova^{a,d}, Jakob Bunkenborg^b, Teresa Cabezon^{a,d}, José M.A. Moreira^{a,d}, Vera Timmermans-Wielenga^{c,d}, Peter Roepstorff^b, Fritz Rank^{a,c,d}, Julio E. Celis^{a,d,*}

^aInstitute of Cancer Biology, Danish Cancer Society, DK-2100 Copenhagen, Denmark

^bDepartment of Biochemistry and Molecular Biology, University of Southern Denmark, DK-5230 Odense, Denmark

^cDepartment of Pathology, The Centre of Diagnostic Investigations, Copenhagen University Hospital, DK-2100 Copenhagen, Denmark

^dThe Danish Centre for Translational Breast Cancer Research, DK-2100 Copenhagen, Denmark

ARTICLE INFO

Article history:

Received 1 November 2009

Received in revised form

13 November 2009

Accepted 13 November 2009

Available online 23 November 2009

Keywords:

Breast cancer

Tumour microenvironment

Interstitial fluid

Secretome

Proteomics

Cancer markers

Early detection

ABSTRACT

Breast cancer is by far the most common diagnosed form of cancer and the leading cause of cancer death in women today. Clinically useful biomarkers for early detection of breast cancer could lead to a significant reduction in mortality. Here we describe a detailed analysis using gel-based proteomics in combination with mass spectrometry and immunohistochemistry (IHC) of the tumour interstitial fluids (TIF) and normal interstitial fluids (NIF) collected from 69 prospective breast cancer patients. The goal of this study was to identify abundant cancer up-regulated proteins that are externalised by cells in the tumour microenvironment of most if not all these lesions. To this end, we applied a phased biomarker discovery research strategy to the analysis of these samples rather than comparing all samples among each other, with inherent inter and intra-sample variability problems. To this end, we chose to use samples derived from a single tumour/benign tissue pair (patient 46, triple negative tumour), for which we had well-matched samples in terms of epithelial cell numbers, to generate the initial dataset. In this first phase we found 110 proteins that were up-regulated by a factor of 2 or more in the TIF, some of which were confirmed by IHC. In the second phase, we carried out a systematic computer assisted analysis of the 2D gels of the remaining 68 TIF samples in order to identify TIF 46 up-regulated proteins that were deregulated in 90% or more of all the available TIFs, thus representing common breast cancer markers. This second phase singled out a set of 26 breast cancer markers, most of which were also identified by a complementary analysis using LC-MS/MS. The expression of calreticulin, cellular retinoic acid-binding protein II, chloride intracellular channel protein 1, EF-1-beta, galectin 1, peroxiredoxin-2, platelet-derived endothelial cell growth factor, protein disulfide isomerase and ubiquitin carboxyl-terminal hydrolase 5 were further validated using a tissue microarray containing 70 malignant breast carcinomas of various grades of atypia. A significant number of these proteins have already been detected in the blood/plasma/secretome by others. The next steps, which include biomarker prioritization based on the hierarchical evaluation of these

* Corresponding authors. Institute of Cancer Biology and The Danish Centre for Translational Breast Cancer Research, Danish Cancer Society, Strandboulevarden 49, DK-2100 Copenhagen, Denmark. Tel.: +45 35 257 354; fax: +45 35 257 363.

E-mail addresses: psg@cancer.dk (P. Gromov), jec@cancer.dk (J.E. Celis).

1574-7891/\$ – see front matter © 2009 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

doi:10.1016/j.molonc.2009.11.003

markers, antibody and antigen development, assay development, analytical validation, and preliminary testing in the blood of healthy and breast cancer patients, are discussed.

© 2009 Federation of European Biochemical Societies.

Published by Elsevier B.V. All rights reserved.

1. Introduction

Breast cancer is the leading cause of cancer deaths in women today and is the most common cancer (excluding non-melanoma skin cancers) among women in the Western world. According to the World Health Organization more than 1.2 million people will be diagnosed with the disease this year worldwide, and the global incidence rates are increasing (Ferlay et al., 2007). In Denmark, approximately 4200 women were diagnosed with the disease in 2005, which means that the number of new cases diagnosed every year has tripled in the past 40 years. Today, the cumulative risk of breast cancer in Danish women is 10.13% at age 75 years (i.e. one in 10) and 12.03% at age 80 (one in eight) (<http://www-dep.iarc.fr/NORDCAN>).

Randomized clinical trials have demonstrated the value of mammography in reducing breast cancer mortality in women who are screened annually or biannually. Consequently, mammography is currently the only accepted screening procedure to detect (measured as the sensitivity) and to exclude (measured as the specificity) the presence of breast cancer in women who are asymptomatic (Humphrey et al., 2002; Brenner et al., 2002; Elmore et al., 2005; Nohacker et al., 2009). Even though mammography has reduced breast cancer mortality significantly, it suffers from some limitations: sensitivity ranges from 90% to as low as 75%, it leads to over diagnosis and overtreatment, and it is inadequate for detecting the disease at a very early-stage, that is before the tumour starts to manifest its malignant potential (Timins, 2005; Pisano, 2005). In addition, it has low-positive prognostic usefulness in younger women (Gillet et al., 1997).

When breast cancer is detected at a localized stage and has less than 10 mm in size, the 5-year survival rate is 98%. If the lesion is larger it has often spread to nearby lymph nodes (regional disease) and the 5-year survival rate drops to 50–80%. If the cancer has spread (metastasized) to distant organs such as the lungs, bone marrow or liver the 5-year survival rate is less than 25% (http://www.cancer.org/docroot/STT/content/STT_1x_Cancer_Facts_Figures_2005.asp). Hence, it is crucial to develop more sensitive diagnostic tools that will not only complement mammography, but also enable the detection and diagnosis of breast cancer much earlier than is currently possible, allowing therapy that is less invasive, thus causing less morbidity in patients, whilst more effective. The ideal screening strategy would involve the development of a panel of highly specific and sensitive biomarkers that can be used to screen high-risk groups, detect recurrence, and monitor treatment using a simple blood-based test that can be performed by a general practitioner. Serum/plasma is relatively easy to collect and it's the main biological fluid used to determine disease status and progression (Omenn, 2004; Issaq et al., 2007 and references therein). Eventually, these

biomarkers may also be used in combination with mammography using other imaging modalities (Fass, 2008 and references therein).

To date, a few proteins have been suggested as possible markers for the early detection of breast cancer; these include the carbohydrate antigen CA15.3 (Hayes, 1996; Martín et al., 2006; Duffy, 2006; Uehara et al., 2008; Kim et al., 2009; Dede et al., 2009), CA 27.29 (Gion et al., 2001; Lumachi and Basso, 2004), carcinoembryonic antigen (CEA) (Hayes, 1996; Duffy, 2006; Uehara et al., 2008; Dede et al., 2009), clusterin (Yom et al., 2009) and alpha-1-antichymotrypsin (Doustjalali et al., 2004). Due to lack of specificity and/or sensitivity for early disease, however, none of these markers is of value for the detection of early breast cancer (Duffy, 2006). Consequently, novel, highly sensitive and specific biomarkers for the early detection of breast cancer are badly needed. Given the intra and inter heterogeneity and complexity of breast tumours it is likely that multiple rather than individual markers may be necessary in order to generate a specific and sensitive serum/plasma protein-based diagnostic test with clinical value.

The application of proteomic technologies has great potential in delineating breast cancer biomarker discovery strategies (Aebersold and Mann, 2003; Ong and Mann, 2005; Veenstra et al., 2005; Chen and Yates, 2007; Conrotto and Souchelnytskyi, 2008; Gromov et al., 2008; Rajcevic et al., 2009; Anderson et al., 2009). To date, a number of sources of biomarkers have been used and these include biological fluids (serum/plasma, tears, CSF, saliva) (Jeffery and Murphy, 2001; Hu et al., 2006; Schmidt and Aebersold, 2006; Maurya et al., 2007; Decramer et al., 2008; Apweiler et al., 2009; Agha-Hosseini et al., 2009; Groves et al., 2009), conditioned media from cultured and primary cells (Celis et al., 1987, 1989; Kulasingam and Diamandis, 2007, 2008; Jacobs et al., 2008; Liang et al., 2009), as well as near fluids (nipple aspirate, fluid bathing the tumour microenvironment, the TIF) (Celis et al., 2004, 2005a; Mannello et al., 2009).

Serum/plasma has been extensively used to derive markers using proteomic technologies such as gel-based proteomics (Tracy et al., 1982; Rui et al., 2003; Musselman and Speicher, 2005; Souchelnytskyi et al., 2006), LC-MS/MS (Shen et al., 2004; Mbeunkui et al., 2007), SELDI-TOF (Gagnon et al., 2008; Garrisi et al., 2008; Gast et al., 2009a, b), high-density antibody arrays (Haab, 2005; Loch et al., 2007), and selective isolation of N-glycosites from the plasma proteome in combination with mass spectrometry (Stahl-Zeng et al., 2007; Zeng et al., 2009; Schiess et al., 2009). Serum/plasma pose a major challenge as source of biomarkers as the dynamic range of protein levels in the blood is very large and a few proteins make for the bulk of the protein content (Polanski and Anderson, 2006; Stahl-Zeng et al., 2007). Biomarkers may be present in the ng/ml range and as a result they may be difficult to reveal (Polanski and Anderson, 2006). Moreover, most tissues contribute proteins to the blood, a fact that may hamper the discovery

of specific markers (Hanash et al., 2008). Another approach that has met with some degree of success has been the analysis of antigens that induce autoantibodies during tumour progression. It has been shown that immune humoral response directed against tumour-associated antigens can be generated early in tumour progression and, therefore, has a great potential for early cancer detection (Casiano et al., 2006; Chapman et al., 2007; Lu et al., 2008; Qiu et al., 2008; Hamrita et al., 2008). By using this approach, Desmetz and co-authors recently described a novel panel of five autoantibodies that significantly discriminate early-stage breast cancer and CIS from healthy individuals (Desmetz et al., 2009). Proteomic analysis of conditioned media also represents an attractive source of biomarkers given the simplicity of the procedure and the homogeneity of the biological material (Celis et al., 1987, 1989; Kulasingam and Diamandis, 2007, 2008; Paliouras and Diamandis, 2008; Xue et al., 2008). The clinical relevance of the discoveries is however problematic as tumour cells in vivo are surrounded by a complex set of cells present in the tumour microenvironment (Liotta and Kohn, 2001; Tlsty and Coussens, 2006; Witz, 2009) and these complex cellular interactions are expected to influence the protein expression phenotype of the tumour cells. Even in more physiologically relevant ex vivo models such as 3D cell cultures the expression patterns one obtains are cell line specific (Kenny et al., 2007) and therefore not necessarily relevant. Analysis of the supernatant of cells present in ascites and pleural effusions, however, may be relevant in spite of the fact that these cells appear only at late stages of disease progression (Faça et al., 2008). Consequently, before embarking on a complex and long-term biomarker discovery process it is crucial to first identify the most relevant sources to initiate the process.

With consideration to the above mentioned constraints, our group has focused on near fluids, primarily on tumour tissue interstitial fluid. We developed procedures to collect and characterize the protein composition of the TIF as this fluid is a potential source for the systematic search for potential serological biomarkers (Celis et al., 2004, 2005a). We surmised that the protein composition of this fluid most likely reflects the physiological and pathological state of the underlying tissue. TIF proteins, which are derived from all cells in the tumour microenvironment, including endothelial cells and their supporting pericytes, inflammatory cells (neutrophils, macrophages, eosinophils, and mast cells), immune cells (lymphocytes and dendritic cells), smooth muscle cells, myofibroblasts, preadipocytes, and adipocytes, are either secreted via the classical ER/Golgi pathways (Harter and Reinhard, 2000), shed by membrane vesicles like exosomes (Koga et al., 2005; Hegmans et al., 2004; Oliver and Vidal, 2007; Simpson et al., 2009), or unconventional protein secretion (Nickel, 2003). Solid tumours display increased interstitial fluid pressure due to high vessel permeability, reduced lymphatic drainage, low perfusion, and increased number of cells in the region of the blood vessels (Lunt et al., 2008 and references therein). As a result, some of the proteins present in the TIF may reach the blood circulation through the lymphatic system and/or blood vessels and hence represent potential serological markers for the early detection of cancer.

Here we describe a detailed comparative proteomic analysis of TIF and normal interstitial fluid (NIF) recovered

prospectively, during the last 6 years, from 69 breast cancer patients that underwent mastectomy at the Copenhagen University Hospital. The goal of this study was to identify abundant cancer up-regulated proteins that are externalised by cells to the tumour microenvironment – in most if not all of the lesions studied – as a first step in our long-term systematic search for serological markers for early detection. The next steps, which include biomarker prioritization, antibody and antigen development, assay development, analytical validation, and preliminary testing in the blood of healthy and breast cancer patients, are discussed.

2. Results

To date, we have carried out a detailed proteomic analysis of the TIF and NIF collected prospectively from specimens obtained from breast cancer patients (newly detected primary tumours) that underwent mastectomy at the Copenhagen University Hospital (Table 1) using gel-based proteomics (2D PAGE) in combination with mass spectrometry and immunohistochemistry (IHC). Of the 69 tumour samples available, 54 corresponded to invasive ductal carcinomas (IDCs), 13 to lobular carcinomas (LCs), one to a tubular carcinoma and one to an invasive apocrine carcinoma. Seven of the IDCs represented triple negative breast cancers (Sørli et al., 2001; Rakha and Ellis, 2009 and references therein) and 14 belonged to the Her-2 amplified subtype (Slamon et al., 1987, Sørli et al., 2001; Ross et al., 2009 and reference therein). Only in 28 of the patients did we have sufficient excess non-malignant tissue available, containing enough epithelial cells, for NIF preparation and these are listed in Table 1.

To recover the TIF and NIF we incubated fresh pieces of tissues for 1 h at 37 °C in phosphate buffered saline (PBS) as described in Section 4 (Celis et al., 2004, 2005a). The protein concentration of the NIFs ranged from 0.3 to 0.6 mg/ml of protein, while those of the TIFs ranged from 0.3 to 1.4 mg/ml.

2.1. Identification of cancer up-regulated proteins in the TIF 46/NIF 46 pair: generation of the initial dataset

Considering that one of the main goals of our study was to identify putative cancer markers that are up-regulated in most if not all of the TIF/tumours, and bearing in mind that a significant fraction of all the proteins detected in silver stained 2D gels are shared by all TIFs, irrespectively of their histopathological type (results not shown), we decided to apply a phased discovery research strategy to the analysis of these samples rather than comparing all samples among each other, with inherent inter and intra-sample variability problems. To this end, we chose to use samples derived from a single tumour/benign tissue pair collected from patient 46 – for which we had well-matched samples in terms of epithelial cell numbers as determined by CK19 antibody staining – to generate an initial dataset of proteins highly up-regulated in the TIF.

Accordingly, tissue samples from patient 46 (triple negative tumour, Table 1) were selected for the analysis as the benign tissue exhibited adenosis (enlarged lobules) and as a consequence contained a significant number of glands (Figure 1A; CK19 staining). Still, the number of epithelial cells in this

Table 1 – Clinicopathological characteristics of breast cancer patients and availability of interstitial fluids.

DCTB number ^a	Age ^b	Tumour type ^c	Tumour size ^d	Grade ^e	Receptor status ^f		Her2-neu IHC ^g	Her2-neu FISH ^g	LN ^h	Availability of		
					ER	PgR				TIF ⁱ	NIF ^j	FIF ^k
27	47	D	20	2	ER+	PGR+	0		N – 0/15	Yes	Yes	
28	53	D	24	3	ER+	PGR+	3+		N + 3/12	Yes		
29	45	D	13	2	ER+	PGR+	0		ND 0/0	Yes		
30	69	D	26	2	ER+	PGR+	3+		N + 1/6	Yes		
31	89	D	22	1	ER+	PGR+	1+		N + 6/15	Yes		
33	68	D	21	2	ER+	PGR+	3+		N + 1/11	Yes		
34	69	D	23	2	ER+	PGR+	1+		N + 3/12	Yes	Yes	
35	75	L	40	1	ER+	PGR+	1+		N + 1/16	Yes		
36	48	D	50	2	ER+	PGR+	3+		N + 8/18	Yes		
37	80	D	16	1	ER+	PGR–	0		N – 0/1	Yes		
38	62	D	25	2	ER+	PGR–	2+	3.1	N + 2/13	Yes		
39	73	D	25	1	ER+	PGR–	2+	1	N – 0/13	Yes		
40	52	D	30	2	ER+	PGR+	2+	1.9	N – 0/14	Yes		
41	68	D	35	3	ER–	PGR–	3+		N + 1/18	Yes		
42	54	D	21	3	ER–	PGR–	2+	1.6	N + 5/10	Yes	Yes	Yes
46	65	D	30	3	ER–	PGR–	2+	1.2	N + 3/19	Yes	Yes	
49	78	D	35	2	ER+	PGR–	2+	1.65	N + 6/7	Yes		
50	46	D	26	2	ER+	PGR+	3+		N + 2/10	Yes		Yes
51	27	D	40	3	ER+	PGR+	2+	1.49	N + 4/26	Yes	Yes	Yes
54	88	D	40	1	ER+	PGR–	1+		N + 1/7	Yes		Yes
55	46	D	50	2	ER+	PGR+	0		N + 4/22	Yes		
56	44	L	20	1	ER+	PGR–	1+		N + 1/14	Yes	Yes	
57	66	D	30	3	ER–	PGR–	3+		N + 27/31	Yes		
58	38	D	20	3	ER–	PGR–	3+		N + 9/20	Yes		
60	72	D	30	2	ER+	PGR–	2+	1.2	N + 3/14	Yes		
61	99	D	40	2	ER+	PGR+	2+	1.5	N + 7/11	Yes		
62	85	L	25	2	ER+	PGR+	2+	1.31	N + 1/13	Yes		
63	92	D	16	2	ER+	PGR+	2+	1.46	N – 0/13	Yes		
64	52	Tu/Crib	23	1	ER+	PGR+	1+		N + 23/25	Yes	Yes	
65	81	L	50	1	ER+	PGR+	1+		N + 4/15	Yes		
66	60	L	70	2	ER+	PGR+	1+		N – 0/10	Yes		
67	78	L	22	2	ER+	PGR+	0		N – 0/7	Yes	Yes	
68	51	L	35	2	ER+	PGR+	1+		N + 2/12	Yes		
69	86	D	25	2	ER+	PGR+	0		N + 3/15	Yes	Yes	
70	83	D	33	3	ER–	PGR–	3+		N + 3/11	Yes		
73	75	L	50	1	ER+	PGR–	1+		N + 14/15	Yes	Yes	
74	49	D	21	3	ER–	PGR–	1+		N + 3/22	Yes	Yes	
76	76	D	30	2	ER+	PGR–	2+	1.21	N – 0/20	Yes	Yes	
78	31	D	32	3	ER–	PGR–	0		N + 14/17	Yes	Yes	
79	56	Apocrine	35	1	ER–	PGR–	1+		N – 0/17	Yes		
80	41	D	40	3	ER+	PGR+	3+		N + 1/15	Yes		
81	39	L	50	2	ER+	PGR+	1+		N + 8/15	Yes		
83	57	D	45	2	ER+	PGR+	2+	1.19	N + 10/16	Yes	Yes	
86	77	D	110	2	ER+	PGR+	1+		N + 20/20	Yes		
87	66	D	25	3	ER–	PGR–	3+		N – 0/10	Yes		
88	83	D	35	2	ER+	PGR+	2+	1.48	N + 8/15	Yes		
89	43	D	16	3	ER+	PGR–	0		N + 1/13	Yes		
91	59	D	30	2	ER+	PGR–	3+		N + 13/16	Yes		
93	60	D	35	2	ER+	PGR+	1+		N + 11/14	Yes	Yes	
94	57	D	21	3	ER–	PGR–	3+		N + 3/13	Yes		
102	73	D	60	2	ER+	PGR–	0		N + 1/11	Yes	Yes	
104	67	D	55	1	ER+	PGR+	1+		N + 12/12	Yes	Yes	
105	83	D	60	2	ER+	PGR+	0		N + 13/18	Yes	Yes	
106	86	D	22	2	ER+	PGR+	0		N – 0/7	Yes		
109	79	D	20	1	ER+	PGR+	0		N + 4/10	Yes	Yes	
111	51	D	23	2	ER–	PGR–	3+		N + 3/9	Yes	Yes	
113	83	D	15	3	ER–	PGR–	3+		N + 20/20	Yes	Yes	
115	79	L	22	2	ER+	PGR+	1+		N + 4/10	Yes		
116	84	D	60	3	ER+	PGR–	2+		N + 19/22	Yes		
117	74	D	30	3	ER+	PGR+	1+		N + 8/15	Yes	Yes	
119	79	D	35	3	ER+	PGR+	1+		N + 10/10	Yes	Yes	
125	68	L	40	2	ER+	PgR+	1+		N – 0/14	Yes		
128	84	D	21	2	ER+	PgR+	0		N – 0/19	Yes		

Table 1 (continued)

DCTB number ^a	Age ^b	Tumour type ^c	Tumour size ^d	Grade ^e	Receptor status ^f		Her2-neu IHC ^g	Her2-neu FISH ^g	LN ^h	Availability of		
					ER	PgR				TIF ⁱ	NIF ^j	FIF ^k
129	76	D	100	2	ER+	PgR–	2+	1.24	N + 3/5	Yes		
155	67	D	50	2	ER+	PgR+	1+		N + 4/10	Yes	Yes	
156	45	L	50	2	ER+	PgR+	0		N + 3/9	Yes	Yes	
157	75	L	19	2	ER+	PgR+	1+		N + 1/10	Yes	Yes	
158	53	D	32	3	ER–	PgR–	0		N – 0/20	Yes	Yes	
161	52	D	40	3	ER–	PgR–	0		N + 4/10	Yes	Yes	

a The column represents internal DCTB patient anonymization number codes that we have used in all our breast cancer studies and publications.

b In years.

c D, ductal carcinoma. L, lobular carcinoma. Tu/Crib, tubular/cribiform.

d The tumour size is given in mm.

e The histological malignancy grade was determined according to [Elston and Ellis \(1991\)](#).

f ER and PgR status was determined by IHC analysis according to DBCG guidelines. Tumours were regarded as negative when both receptors were expressed in less than 10% of tumour cells.

g HER2 positive: IHC staining 3+, or 2+ if amplified, i.e. FISH HER2/CEN-17 ratio >2.

h LN status. N+ metastasis in lymph nodes. N– no metastasis detected in lymph nodes.

i TIF, tumour interstitial fluid.

j NIF, normal interstitial fluid.

k FIF, fat interstitial fluid.

“normal sample” was somewhat lower per area than that observed in the tumour piece ([Figure 1B](#); CK19 staining). In addition to the TIF/NIF set of samples we had also obtained a high-quality sample from regional metastases (axillary lymph node metastasis) so that for this particular patient we had a high-quality, well-matched set of samples derived from benign, malignant and metastatic tissue. [Figures 2A,B](#) show representative silver stained 2D gels of TIF 46 and NIF 46 loaded with similar amount of proteins. Of the 623 proteins detected in the 2D gels of TIF 46, 110 relatively abundant proteins were found to be up-regulated by a factor of two or more in this fluid as compared to its corresponding NIF and these are indicated in [Table 2](#) together with their identity as determined by mass spectrometry (see also [Supplementary Table 1](#)). The presence of the 110 proteins in 2D gels of total protein extracts from tumour 46 ([Supplementary Figure 1A](#)), and metastasis 46 ([Supplementary Figure 1B](#)) was confirmed by image analysis and in several cases by mass spectrometry. As previously reported, several of the proteins present in the tumour were not detected or were present at very low levels in the TIF supporting the contention that the TIF proteins correspond to bona fide externalised components (arrows in [Supplementary Figure 2](#)).

To confirm the results obtained by the comparative 2D gel analysis we performed IHC staining of non-malignant and tumour paraffin-embedded tissue samples from patient 46 using antibodies against a few proteins such as calreticulin ([Figures 3A,B](#)), calumenin ([Figures 3C,D](#)), translationally-controlled tumour protein (TCPT; [Figures 3E,F](#)), and S100A9 (MRP14) ([Figures 3G,H](#)). As depicted in [Figure 3](#), the IHC analysis corroborated the results obtained by the gel-based approach. Most of the up-regulated proteins verified by IHC were derived from epithelial cells, although in the case of S100A9 they stemmed mainly from cells present in the stroma ([Figures 3G,H](#)). It should be noted that the non-malignant tissue sample exhibited a few areas with atypia as determined by p53 staining (arrow in [Figure 3I](#); compare with a similar staining of T46 with the p53 antibody shown in [Figure 3J](#)); hence, it is possible that the levels of the up-regulated proteins in this NIF may in fact be even lower in other NIFs lacking such abnormal features. Further verification by IHC of the remaining up-regulated proteins was not possible due to the limited number of tumour 46/benign 46 paraffin-embedded sections available.

Comparison of the protein profiles of NIF 46 with those of the 27 other NIFs available confirmed the low expression

Figure 1 – Paraffin-embedded sections of (A) benign breast tissue 46 and (B) tumour 46 stained with a monoclonal antibody against CK19 (dilution 1:1000).

Figure 2 – Silver stained IEF 2D gels of (A) TIF 46 and NIF 46. Proteins up-regulated by a factor of 2 or more are indicated with arrows.

levels of the set of 110 up-regulated proteins identified in TIF 46 (results not shown).

To strengthen our 2D gel-based analysis and address some of the inherent methodological limitations of gel-based proteomics, we carried out a set of LC-MS/MS based proteomic analysis of TIF. Several strategies with lower resolving power than 2D gel electrophoresis were employed to identify lower abundance proteins in the human albumin rich samples. As the TIF 46 sample was exhausted during our 2D gel comparative analysis we used another high-quality, highly-validated sample, TIF 68. By using LC-MS/MS we could identify 84 out of the 110 (75%) TIF 46 up-regulated protein set (Table 2) corroborating the protein identity in the dataset generated by 2D gel proteomics, as well as strengthening the potential usefulness of some of these proteins as markers. The detailed analysis of the LC-MS/MS data also revealed that in most cases additional subunits of proteins identified by the gel-based approach were also detected in the LC-MS/MS dataset. As expected, the datasets generated by two proteomic technologies were not identical, but were strikingly similar (Supplementary Tables 1 and 2).

2.2. Identification of proteins deregulated in 90% or more of the TIFs analyzed

Having established a working set of 110 up-regulated proteins in the first phase of our discovery strategy, we proceeded to the second phase. In this phase, we carried out a systematic computer assisted analysis of the 2D gels of the remaining 68 TIF samples in order to identify among the 110 TIF 46 up-regulated proteins those that were detected in most of the available tumour tissue fluid samples. Given the large number of gels that needed to be matched we divided the images into 12 sets of five to six samples in each (Figure 4), and a master image was prepared that contained all the proteins spots present in the different images. All 110 TIF 46 up-regulated proteins were matched throughout the 12 image sets and

the 26 proteins that were detected in more than 90% of the samples were selected for further quantitative analysis using the 69 TIFs and 28 NIFs available (Table 3). The results of the non-parametric Mann-Whitney analysis are presented in Figure 5 (see also Supplementary Tables 3 and 4). All 26 proteins were found to be up-regulated in the TIFs as compared to the NIF counterparts (*P* values are given in Figure 5). It should be noted that the 26 TIF up-regulated proteins were found to be expressed in the most of breast cancer tumours used to collect the TIFs (data not shown). As an example, Supplementary Figure 3 shows the expression of the 26 up-regulated proteins in six representative IDCs.

As before, with the exception of the proliferating cell nuclear antigen and T-complex protein 1 subunits epsilon and theta, all other proteins were independently identified by LC-MS/MS analysis of TIF 68 (see Table 2 and Supplementary Table 2).

It should be stressed that in many cases the matching procedure was difficult to perform due to the low abundance of some proteins, the fact that some proteins migrated close to a cluster of proteins, heterogeneity of tissue samples, as well as to the inherent variability in the gel runs. Consequently, it is possible that some of the proteins that fell short of the 90% sample presence threshold we imposed in this analysis may in reality meet this criterion. Moreover, the 90% presence inclusion is a subjectively defined criterion imposed solely for the purpose of prioritization of efforts for the third phase of the process, that of experimental validation of hits. As such, it is conceivable that some of the proteins we identified in the TIF46, but that are only deregulated in a small percentage of samples, may have clinical usefulness as markers for particular subsets of tumours or as part of a multimarker protein signature. In this respect we should stress that given that the ultimate aim of this study was to identify potential serological markers, our analysis did not include down-regulated proteins nor did we try to perform multivariate data analysis. This study presents a first approach to the process of

Table 2 – Proteins up-regulated in TIF46 as compared to NIF46.

Protein number ^a	Protein name	SwissProt number	M _r /pI ^b	Ratio of protein spot intensities TIF46/NIF46 ^c	Identified by LC-MS/MS in TIF68 ^d
1	14-3-3 protein epsilon	P62258	29/4.6	2.9	Yes
2	14-3-3 protein zeta/delta	P63104	27.9/4.7	3.8	Yes
3	14-3-3 protein zeta/delta, variant	P63104	27.9/4.7	3.3	Yes
4	26S proteasome non-ATPase regulatory subunit 11	O00231	47/6.1	2.6	No
5	34/67 kDa laminin receptor	P08865	33/4.8	>10	Yes
6	Acidic leucine-rich nuclear phosphoprotein 32 family, member A	P39687	28.7/4	>10	Yes
7	Acidic leucine-rich nuclear phosphoprotein 32 family, member A, variant	P39687	28.7/4	>10	Yes
8	Actin, alpha skeletal muscle	P02568	42/5.2	2.2	Yes
9	Actin-beta, fragment	P60709	71.8/5.8	>10	Yes
10	Actin-related protein 3	P61158	48/5.6	3.5	Yes
11	Acyl-CoA-binding protein	P07108	10/6.1	4.3	Yes
12	Adenosylhomocysteinase	P23526	48.4/6.0	4.5	Yes
13	Alanyl-tRNA synthetase, cytoplasmic	P49588	107/5.3	2.4	No
14	Aldehyde dehydrogenase family 9 member A1	P49189	54.6/6.01	4.3	Yes
15	Aldo-keto reductase family 1 member A1	P14550	36.8/6.4	2.4	Yes
16	Aldose reductase	P15121	36/6.5	4	Yes
17	Alpha-actinin-4	O43707	105/5.2	4.8	Yes
18	Bifunctional purine biosynthesis protein PURH	P31939	65/6.2	2.2	Yes
19	Calreticulin	P27797	48/4.3	3.1	Yes
20	Calumenin	O43852	37.2/4.5	>10	Yes
21	Caspase 14	P31944	28/5.4	3	No
22	Cathepsin D	P07339	38/5.6	2	Yes
23	Cellular retinoic acid-binding protein 1	P29762	15.7/5.3	>10	No
24	Cellular retinoic acid-binding protein 2	P29373	15.7/5.4	3.3	Yes
25	Centractin alpha	P61163	43/6.2	3.5	No
26	Chloride intracellular channel protein 1	O00299	27.4/5.02	2.5	Yes
27	Coastosin-like protein	Q14019	16/5.5	>10	No
28	Cytoplasmic dynein 1 intermediate chain 2	Q13409	71/5.0	2.2	No
29	Cytosolic non-specific dipeptidase	Q96KP4	52/5.6	7.3	Yes
30	Dihydropyrimidinase-related protein 2	Q16555	63/5.9	2.3	Yes
31	Dimethylargininase 2	O95865	30/5.6	>10	Yes
32	E-FABP	Q01469	15.5/6.6	>10	Yes
33	Elongation factor 1-beta	P24534	24.8/4.5	4.4	Yes
34	Elongation factor 1-gamma	P26641	50.5/6.25	>10	Yes
35	Elongation factor 2	P13639	94/6.4	6.2	Yes
36	F-actin-capping protein subunit alpha-1	P52907	33/5.45	2	Yes
37	F-actin-capping protein subunit alpha-2	P47755	33/5.5	>10	Yes
38	F-actin-capping protein subunit beta	P47756	31/5.3	>10	No
39	Filamin 1	P21333	283/5.7	2.5	Yes
40	Filamin 1, variant	P21333	283/5.7	5.8	Yes
41	Galectin-1	P09382	15/5.4	3.6	Yes
42	Gamma-enolase	P09104	47/4.9	3.7	No
43	Glutathione-S-transferase omega-1	P78417	28/6.2	8.4	Yes

(continued on next page)

Table 2 (continued)

Protein number ^a	Protein name	SwissProt number	M _r /pI ^b	Ratio of protein spot intensities TIF46/NIF46 ^c	Identified by LC-MS/MS in TIF68 ^d
44	Glutathione-S-transferase P	P09211	23/5.4	2.4	Yes
45	Glyoxalase domain-containing protein 4	Q9HC38	35/5.4	>10	No
46	GTPase Ran	P62826	24.6/7.01	6.7	Yes
47	Heat shock protein HSP 90-beta	P08238	83.6/4.9	2.2	Yes
48	Histone-binding protein RBBP7	Q16576	48/4.8	5.6	No
49	Immunity-related GTPase family Q protein	Q8WZA9	63/4.8	3.3	No
50	Importin subunit beta-1	Q14974	98.8/4.7	2.9	Yes
51	Inorganic pyrophosphatase (PPase)	Q15181	33/5.54	9.8	Yes
52	Inorganic pyrophosphatase (PPase), variant	Q15181	33/5.54	>10	Yes
53	Isocitrate dehydrogenase [NADP] cytoplasmic	O75874	46.9/6.5	3.8	Yes
54	L-lactate dehydrogenase B chain	P07195	36.7/5.7	2.2	Yes
55	L-lactate dehydrogenase B chain, variant	P07195	36.7/5.7	>10	Yes
56	Lysophospholipase II	O95372	25/6.7	>10	Yes
57	Macrophage-capping protein	P40121	38.8/5.8	3.1	Yes
58	Maspin	P36952	42/5.7	3.4	No
59	Myotrophin	P58546	13/5.2	>10	Yes
60	Neutral alpha-glucosidase AB	Q14697	107/5.4	3.4	Yes
61	Nicotinamide N-methyltransferase	P40261	30/5.5	>10	Yes
62	Nuclear transport factor	P61970	14.8/5.1	>10	Yes
63	Peroxiredoxin 1	Q06830	22.3/8.27	3.3	Yes
64	Peroxiredoxin-2	P32119	22/5.6	2.6	Yes
65	Peroxiredoxin-6	P30041	25/6.0	4.9	Yes
66	Peroxiredoxin-6, variant	P30041	25/6.0	3.8	Yes
67	Phosphoglycerate mutase 1	P18669	28.7/7.6	5.7	Yes
68	Platelet-activating factor acetylhydrolase IB subunit beta	P68402	26/5.6	2.9	Yes
69	Platelet-derived endothelial cell growth factor	P19971	50/5.4	3.3	Yes
70	Proliferating cell nuclear antigen	P12004	29/4.5	>10	No
71	Proliferation-associated protein 2G4	Q9UQ80	44/6.1	2.4	No
72	Proteasome activator complex subunit 2	Q9UL46	27/5.4	2.3	Yes
73	Proteasome subunit alpha type-2	P25787	26/6.9	>10	Yes
74	Proteasome subunit beta type-3	P49720	23.3/6.1	>10	Yes
75	Proteasome zeta chain	P28066	26/4.7	3	Yes
76	Protein disulfide isomerase	P07237	55/4.7	2.6	Yes
77	Protein disulfide isomerase A3	P30101	57/6.23	2.7	Yes
78	Protein DJ-1	Q99497	20/6.3	2.2	Yes
79	Rab GDP dissociation inhibitor beta	P50395	51/6.1	4.9	Yes
80	Rab GDP dissociation inhibitor beta, variant	P50395	51/6.1	3.6	Yes
81	S100 A11	P31949	11.8/6.5	3.4	Yes
82	S100 A6	P06703	10/5.3	>10	No
83	S100 A9	P06702	13.5/5.7	>10	Yes
84	Selenide, water dikinase 1	P49903	43/5.6	>10	No
85	Serum amyloid P-component	P02743	25.5/6.1	2.1	Yes
86	SH3 domain-binding glutamic acid-rich-like protein 3	Q9H299	10/4.8	>10	Yes
87	SM22-alpha	Q01995	22/8.8	>10	Yes
88	Stress-induced-phosphoprotein 1	P31948	63.4/6.4	4.2	Yes
89	T-complex protein 1 subunit alpha	P17987	61/5.8	2.7	No
90	T-complex protein 1 subunit alpha, variant	P17987	61/5.8	>10	No
91	T-complex protein 1 subunit beta	P78371	57.7/6.02	5.3	No
92	T-complex protein 1 subunit epsilon	P48643	60/5.4	>10	No

Table 2 (continued)

Protein number ^a	Protein name	SwissProt number	M _r /pI ^b	Ratio of protein spot intensities TIF46/NIF46 ^c	Identified by LC-MS/MS in TIF68 ^d
93	T-complex protein 1 subunit gamma	P49368	61/6.1	>10	No
94	T-complex protein 1 subunit theta	P50990	60/5.4	5.4	No
95	Thioredoxin	P10599	12/4.8	3.3	Yes
96	Translationally-controlled tumour protein	P13693	19.7/4.8	>10	Yes
97	Triosephosphate isomerase	P60174	26.8/6.5	3.6	Yes
98	Triosephosphate isomerase, variant	P60174	26/6.5	>10	Yes
99	Triosephosphate isomerase, variant	P60174	26.8/6.5	>10	Yes
100	Tropomyosin 3 isoform 2	gi 24119203	29/4.7	6.8	No
101	Tropomyosin alpha-4 chain	P67936	28.6/4.6	6.3	Yes
102	Tropomyosin alpha-4 chain, variant	P67936	28.6/4.6	3.2	Yes
103	Tryptophanyl-tRNA synthetase, cytoplasmic	P23381	53.5/5.8	6.1	No
104	Tryptophanyl-tRNA synthetase, cytoplasmic, variant	P23381	53.5/5.8	2.6	No
105	Ubiquitin carboxyl-terminal hydrolase 5	P45974	97/4.9	2.8	Yes
106	Ubiquitin-like modifier-activating enzyme 1	P22314	119/5.5	7.7	Yes
107	Ubiquitin-like modifier-activating enzyme 1, variant	P22314	119/5.5	6.1	Yes
108	Vimentin (C-terminal fragment)	P08670	54/5.06	>10	Yes
109	WD repeat-containing protein 1	O75083	66/6.1	3	Yes
110	WD repeat-containing protein 1, variant	O75083	66/6.1	4.7	Yes

a Protein numbers are listed according to their positions in the 2D gels as shown in Figure 2.

b Theoretical values of molecular mass (M_r) and isoelectric point (pI).

c Quantification of the gel images was performed using the PDQuest 8.0.1 software as described in Section 4. Ratio of more than 10 means that a corresponding spot was not detected in NIF46.

d Proteins were identified by shotgun LC-MS/MS proteomics as described in Section 4 (see also Supplementary Table 2).

identification of clinically useful markers. Our approach to this process is based on an iterative and incremental development of data analysis. As such, the proteomics data we generated can always be reanalyzed using different criteria than the ones we applied in this study.

2.3. IHC verification of the common TIF up-regulated cancer protein markers using an independent set of malignant breast carcinomas

The third phase in our discovery process was that of experimental validation of hits. For this purpose we used a commercially available tissue microarray (TMA), the Pantomics BRC1503 breast cancer tissue array that contains 70 malignant carcinomas with various grades of atypia (7 (10%) grade I; 20 (28.5%) grade II; 41 (57.1%) grade II/III and 3 (0.43%) grade III). IHC staining with various antibodies verified the up-regulation of the target antigens in tumour tissue compared to normal/benign tissue as well as the widespread expression of several of the markers in this independent set of breast tumours: these included calreticulin (staining of 69 out of 70 tumours), cellular retinoic acid-binding protein II (staining of 66 out of 70 the tumours), chloride intracellular channel protein 1 (staining of 68 out of the 70 tumours), EF-1-beta (staining of 69 out of the 70 tumours), galectin 1 (staining of 60 out of the 70 tumours), peroxiredoxin-2 (staining of 67 out of the 70 tumours), platelet-derived endothelial cell growth factor (staining of 64 out of the 70 tumours), PDI (staining of 68 out of the

70 tumours) and ubiquitin carboxyl-terminal hydrolase 5 (staining of 59 out of the 70 tumours). Representative examples of normal versus tumour cores for all these proteins are presented in Figure 6. A representative area comprising 12 tumour cores from TMAs stained with antibodies against calreticulin (Figure 7A), platelet-derived endothelial cell growth factor (Figure 7B), peroxiredoxin-2 (Figure 7C), protein disulfide isomerase (Figure 7D) are shown in Figure 7 to illustrate the widespread expression of these antigens. Some cores had only a few tumour cells, and this together with differences in staining intensities observed with various antibodies, explain the dissimilarity in staining intensities observed in the various cores. It should be stressed that these studies necessitated in every single case careful evaluation of antibodies (in some cases the use of more than one commercial source of antibody was required) as well as individual titration of the antibodies for differential analysis of this particular TMA.

The Human Protein Atlas is a publicly available Web-based analysis tool for exploring human tissue protein profiles in normal and cancer cells based on images representing immunohistochemically stained tissue sections. Although this resource contains a much smaller number of mammary tumours than the TMAs we used and our own sample collection, it provided laboratory-independent verification of our data. Verification was achieved by visual inspection of IHC images of “normal” and breast carcinomas available at the site (Table 3) (<http://www.proteinatlas.org>; version 4) for those

Figure 3 – IHC staining of paraffin-embedded sections of benign 46 (A, C, E, G and I) and tumour 46 (B, D, F, H and J) stained with antibodies against TIF 46 up-regulated proteins. (A) and (B) calreticulin, dilution 1:200. (C) and (D) calumenin, dilution 1:1000. (E) and (F) TCPT, dilution 1:2000. (G) and (H) SA100 A9, dilution 1:1000. (I) and (J) p53, dilution 1:250. The arrow in I indicates cells with atypia (compare with J).

Figure 4 – PDQuest image analysis of TIF and NIF samples. 2D images were matched and analysed in sets of five to six images each. A master image was established based on the gel of TIF 46. All proteins detected in the set were added to the master image.

markers included in the current version of the Atlas. Visual analysis of the IHC images confirmed up-regulation of EF-2, inorganic pyrophosphatase, glutathione-S-transferase pi and TCP-1-epsilon. It also supported our TMA-based analysis for the cellular retinoic acid-binding protein II, peroxiredoxin-2, platelet-derived endothelial cell growth factor, galectin 1, calreticulin, chloride intracellular channel protein 1, and ubiquitin carboxyl-terminal hydrolase 5 (Table 3). In some cases, however, it was difficult to compare the “normal” and tumour images in the Human Protein Atlas as the IHC signal, particularly in the “normal” preparations, was too high to assess differential expression.

2.4. Most of the 26 TIF up-regulated proteins are present at low levels in the interstitial fluid from fresh mammary fat tissue

Considering that in older women fat tissue is the major component of the mamma and that we have often detected breast tumour cells interdigitating with and spreading through the peripheral fat tissue (Figure 8; tumour double stained with CK 19 and A-FABP antibodies which stain fat cells) suggesting a close association between these two cell types (Celis et al., 2005a), we proceeded to analyse the interstitial fluid of fresh fat tissue samples (FIF) collected from five patients (Table 1) as some of the proteins externalised by the adipose tissue may contribute to the overall levels of the TIF up-regulated proteins. This is quite important as the high ratio of adipose tissue/glands in older women will make any protein externalised in large amounts by adipocytes unsuitable as a biomarker unless it occurs as a specific response to tumour cells. Gel analysis of the various FIFs revealed that the expression profiles and

protein levels were very similar among the various FIFs analysed, and a representative example of the FIF from patient 51 is shown in Figure 9A. With the exception of galectin 1, isocitrate dehydrogenase, lactate dehydrogenase B and peroxiredoxin-2 (red arrows), the levels of the 26 TIF up-regulated proteins were found to be substantially lower in all four FIFs analysed (FIF 51, black arrows in Figure 9A) as compared to the TIFs (TIF 51, red arrows in Figure 9B). Two of the major proteins present in FIF, A-FABP and GPDH-C, which are well known markers for adipose cells (indicated with blue arrows in Figures 9A,B), were not detected in the majority of the TIFs, suggesting that the specimens used to collect these fluids did not present adipose tissue infiltration, a fact that was verified by IHC staining of paraffin-embedded tumours with the A-FABP antibody. Only two lesions, tumours 61 and 66, exhibited significant levels of these proteins (results not shown).

3. Discussion

There is an urgent need to identify biomarkers that can help in the early diagnosis of breast cancer as cancer is easier to treat and cure if it is detected early at a localized stage. Proteomics methods have great potential for the discovery of novel biomarkers and promise to have a major impact on clinical practice, but to date the outcome from the great majority of biomarker discovery studies has been disappointing (Duffy, 2006; Brooks, 2009). This is for the most part due to a lack of coherent long-term strategies that look beyond the biomarker discovery-phase (Rifai et al., 2006). Most studies do not take into account the later stages of the process that goes from biomarker discovery to clinical application. This is intrinsically

Table 3 – Proteins up-regulated in 90% or more of the TIFs analyzed.

Protein number ^a	Protein name ^b	SwissProt number	GO annotation (process/function) ^c	Identification by LC-MS/MS in TIF68 ^d	Presence in serum/plasma ^e	Up-regulation verified using the BRC1503 breast cancer TMA ^f	Verification of expression in normal breast TMA available in the Human Protein Atlas ^g	Verification of expression in breast carcinoma TMA available in the Human Protein Atlas ^g	Identified in the secretome of breast cancer cell lines ^h	Up-regulation in the secretome of normal and breast cancer cell lines ⁱ
2	14-3-3 protein zeta/delta	P63104	Anti-apoptosis	Yes	Yes (Anderson et al., 2004)				Yes	No
15	Aldo-keto reductase family 1 member A1 (AKR1A1)	P14550	Protein binding	Yes	No data				Yes	No
16	Aldose reductase (AR)	P15121	Response to stress	Yes	No data				Yes	No
19	Calreticulin	P27797	Regulation of apoptosis/positive regulation of cell cycle/androgen receptor binding	Yes	Yes (Anderson et al., 2004)	Yes	Yes	Yes	Yes	No
24	Cellular retinoic acid-binding protein 2 (CRABP II)	P29373	Signal transduction/transporter activity	Yes	No data	Yes	Yes	Yes	Yes	No
26	Chloride intracellular channel protein 1 (CLIC1)	O00299	Signal transduction/transporter activity	Yes	Yes (Shen et al., 2004)	Yes	Yes	Yes	Yes	No
33	Elongation factor 1-beta (EF-1-beta)	P24534	Translational elongation/protein binding	Yes	Yes (Schenk et al., 2008)	Yes			Yes	No
35	Elongation factor 2 (EF-2)	P13639	Translational elongation/protein binding	Yes	Yes (Schenk et al., 2008)		Yes	Yes	Yes	No
41	Galectin-1	P09382	Regulation of apoptosis/positive regulation of I-kappaB kinase/NF-kappaB cascade/protein binding	Yes	Yes (Liu et al., 2009)	Yes	Yes	Yes	Yes	No
44	Glutathione S-transferase P (GSTP1-1)	P09211	Anti-apoptosis/protein binding	Yes	Yes (Rose et al., 2004)		Yes	Yes	Yes	No

50	Importin subunit beta-1 (Importin-beta-1)	Q14974	Protein transport	Yes	Yes (Shen et al., 2004)			Yes	No
51	Inorganic pyrophosphatase (PPase)	Q15181	Metabolic process	Yes	No data		Yes	Yes	No
53	Isocitrate dehydrogenase [NADP] cytoplasmic (IDH)	O75874	Metabolic process	Yes	Yes (HUPO plasma proteome 3020 list)			Yes	No
54	L-lactate dehydrogenase B chain (LDH-B)	P07195	Metabolic process	Yes	Yes (Anderson et al., 2004)			Yes	No
64	Peroxisomal protein 2 (PRDX2)	P32119	Regulation of apoptosis	Yes	Yes (HUPO plasma proteome 3020 list)	Yes	Yes	Yes	No
67	Phosphoglycerate mutase 1 (PGAM-B)	P18669	Metabolic process	Yes	No data			Yes	No
69	Platelet-derived endothelial cell growth factor (PD-ECGF)	P19971	Cell-cell signaling/ angiogenesis/ cell differentiation	Yes	Yes (Huang et al., 2004; Stefani et al., 2009)	Yes	Yes	Yes	No
70	Proliferating cell nuclear antigen (PCNA)	P12004	Cell proliferation/ DNA repair	No	Yes (Takahashi et al., 1997)			Yes	No
76	Protein disulfide-isomerase (PDI)	P07237	Protein folding	Yes	Yes (Anderson et al., 2004)	Yes		Yes	No
78	Protein DJ-1 (DJ-1)	Q99497	Cell proliferation/ response to oxidative stress/ protein folding	Yes	Yes (Anderson et al., 2004)			Yes	No
92	T-complex protein 1 subunit epsilon (TCP-1-epsilon)	P48643	Protein folding/ protein metabolic process	No	Yes (Anderson et al., 2004)		Yes	Yes	No
94	T-complex protein 1 subunit theta (TCP-1-theta)	P50990	Protein folding/ protein metabolic process	No	Yes (Anderson et al., 2004)			Yes	No
95	Thioredoxin (Trx)	P10599	Cell proliferation/ cell-cell signaling/ transport	Yes	Yes (Anderson et al., 2004)		Yes	Yes	No
97	Triosephosphate isomerase (TIM)	P60174	Metabolic process	Yes	Yes (Okano et al., 2006)			Yes	No
101	Tropomyosin alpha-4 chain (TPM-alpha-4)	P67936	Cell motion	Yes	Yes (Anderson et al., 2004)		Yes	Yes	No

Table 3 (continued)

Protein number ^a	Protein name ^b	SwissProt number	GO annotation (process/function) ^c	Identification by LC-MS/MS in TIF68 ^d	Presence in serum/plasma ^e	Up-regulation verified using the BRC1503 breast cancer TMA ^f	Verification of expression in normal breast TMA available in the Human Protein Atlas ^g	Verification of expression in breast carcinoma TMA available in the Human Protein Atlas ^g	Identified in the secretome of breast cancer cell lines ^h	Up-regulation in the secretome of normal and breast cancer cell lines ⁱ
105	Ubiquitin carboxyl-terminal hydrolase 5 (UCHL5)	P45974	Hydrolase activity	Yes	Yes (Anderson et al., 2004)	Yes	Yes	Yes	Yes	No

a Protein numbers correspond to the numbers given in Table 2 as well as in Figure 2.

b Protein names are listed in alphabetic order. Short protein names are given in brackets.

c Protein functions are annotated according to the APROPOS genomic and proteomic database (<http://apropos.mcw.edu/>).

d Proteins were identified by LC/MS-MS proteomics as described in Section 4 (see also Table 2 and Supplementary Table 2).

e Screening was performed through several plasma protein databases as well as via PubMed search.

f IHC analysis was performed in this study using the Pantomic BRC1503 breast cancer TMA.

g IHC images available in the Human Protein Atlas database (<http://www.proteinatlas.org>).

h Proteins that have been identified as externalised by several breast cancer cell lines (Kulasingam and Diamandis, 2007). The term “secretome” includes all proteins that are externalised into the media (secretion, shedding, cell death).

i Proteins that have been identified as deregulated proteins in the secretome of normal and breast cancer cell lines. Four independent studies were analysed (Kulasingam and Diamandis, 2007; Mbeunkui et al., 2007; Jacobs et al., 2008; Liang et al., 2009).

Figure 5 – Scatter dot plots showing comparisons of the levels of the 26 cancer up-regulated proteins between TIF (69 sample set) and NIF samples (28 sample set). Data are presented as protein spot intensities measured for each protein in both TIF (blue dots) and NIF (red dots) sample sets. P values were determined using the non-parametric Mann–Whitney test. Horizontal bars indicate mean with SEM.

Figure 6 – IHC staining of paraffin-embedded sections of normal (A, C, E, G, I, K, M, O and Q) and tumour cores (B, D, F, H, J, L, N, P and R) in the Pantomics BRC1503 breast cancer tissue array stained with various antibodies against up-regulated proteins. (A) and (B) calreticulin, dilution 1:200. (C) and (D) CRABP II, dilution 1:1500. (E) and (F) CLIC 1, dilution 1:250. (G) and (H) EF-1-beta, dilution 1:1500. (I) and (J) galectin 1, dilution 1:250. (K) and (L) PRDX 2, dilution 1:3000. (M) and (N) PD-ECGF 1, dilution 1:300. (O) and (P) PDI, dilution 1:800. (Q) and (R) USTH 5, dilution 1:50.

a multiphase process, and during study design one needs to take into consideration not only the overall workflow, but also the most effective strategies within each phase of the process.

Having in mind the conceptual biomarker development phases proposed by the Early Detection Research Network (<http://edrn.nci.nih.gov>; candidate discovery, prioritization, verification, reagent development and assay optimization, analytical validation, and biomarker validation) we have taken the first steps towards a long-term and systematic identification of biomarkers that may be pursued as candidates for a blood-based test for the early detection of breast cancer. We describe here the results of our strategic approach towards

the identification of proteins up-regulated in the tumour microenvironment that are potential serological markers for early detection of cancer of the breast. The study design we adopted in this first phase of the biomarker development process departs from the standard conceptual framework for biomarker discovery. Rather than doing a full comparative proteomic analysis of all samples to find deregulated proteins we first identified all proteins present in a single tumour interstitial fluid sample, then pinpointed those that were not present in normal tissue and only then using this selected dataset of deregulated proteins did we expand our analysis to a large number of samples to identify those commonly deregulated. In this respect it is important to consider that the tumour

Figure 7 – Representative tumour cores of the Pantomics BRC1503 breast cancer tissue array stained with various antibodies. (A) Calreticulin. (B) Platelet-derived endothelial cell growth factor. (C) Peroxiredoxin-2. (D) Protein disulfide isomerase.

burden detection limit of any given blood tumour biomarker will most likely be directly proportional to the biomarker secretion rate and expression levels. By using TIF as the source and 2D gel-based proteomics as discovery tool, thus prioritizing the discovery of relatively abundant, differentially expressed, externalized proteins we tried to maximize the probability of finding a clinically useful serological marker.

Figure 8 – Indirect immunofluorescence analysis of fat tissue peripheral to a breast tumour stained with A-FABP (dilution 1:250; Alexa Fluor 488; green) and keratin 19 (dilution 1:1000; Alexa Fluor 594; red) specific antibodies.

Of the 26 cancer up-regulated proteins identified in our study, only calreticulin (Franzén et al., 1997; Bini et al., 1997), lactate dehydrogenase (Franzén et al., 1997), GST-pi (Franzén et al., 1997), thioredoxin (Bini et al., 1997), PDI (Bini et al., 1997; Chahed et al., 2008), TPI (Zhang et al., 2005), EF-1-beta (Deng et al., 2006), and galectin 1 (Chahed et al., 2008) have been previously described as deregulated in breast cancer, albeit in a very limited number of tumours tissue samples. All 26 markers have been identified in conditioned media of three breast cell lines: MCF-10A (a basal B subtype with intact p53), BT474 (a luminal subtype) and MDA-MB-468 (a basal A-like subtype) (Kulasingham and Diamandis, 2007), confirming these proteins as *bona fide* externalised proteins and emphasizing the validity of the approach we took. Significantly, none of the 26 markers was found to be up-regulated in Kulasingham and Diamandis studies (2007), LC-MS/MS studies of differentially secreted proteins by isogenic cell lines in a breast cancer progression model (Mbeunkui et al., 2007; Jacobs et al., 2008; Liang et al., 2009) or by a pair of normal (HTB-125) and malignant breast carcinoma (HTB-126) cell lines analyzed by the SILAC technology (Liang et al., 2009). This apparent discrepancy between our tissue/TIF studies and the cell line *in vitro* results is most likely due to the striking changes that occur as a result of the cultivation process (Celis et al., 1999). Recently, Mann and colleagues (Pan et al., 2009) showed that the protein expression profile of a well established cell line differed fundamentally from the primary cell culture counterpart in a number of important aspects that included rearrangement of metabolic pathways, drastic up-regulation of cell cycle-associated functions, and important shut down of drug metabolizing enzymes. These results must be taken

Figure 9 – Silver stained 2D gels of (A) FIF 51 and (B) TIF 51. TIF up-regulated proteins (black arrows) as well as fat tissue specific proteins (blue arrows) are indicated in the figure.

into account when starting a long-term biomarker discovery process.

Considering that the gel-based proteomic strategy detects mainly proteins that are present at moderate to high levels comprehensive identification of proteins requires that one complements the approach with other technologies, such as LC-MS/MS (Chen and Yates, 2007 and references therein), antibody arrays (Haab, 2005), and autoantibodies (Kilic et al., 2008) in order to reveal lower abundance protein markers. In this study, we analyzed the TIF from one patient using LC-MS/MS and we found that many of the proteins detected by the gel approach, including the 26 candidate biomarkers we identified, were also identified by the shotgun procedure, supporting the gel-based data as well as the robustness of the procedure we have used to collect the externalised proteins. In addition, the shotgun proteomic analysis revealed several proteins that were not detected by the 2D gel approach and vice versa (Supplementary Table 2) emphasizing the complementarity of the technologies. Previously, we analyzed the TIF and NIF from breast cancer patients using antibody arrays containing antibodies against 150 cytokines and growth factors (RayBiotech C-1000 antibody array) and the results showed that it is feasible to use the TIF for this analysis as the fluid is recovered under non-denaturing conditions and could be directly applied to the arrays.

So far, our studies have focused mainly on TIF up-regulated proteins that are common to most of the breast histopathological types, but it is also possible to use the same approach to search for proteins that are unique/prominent to a given breast cancer subtype. For example, our proteomic studies of invasive apocrine carcinomas (IACs) (Celis et al., 2007b, 2008, 2009) have identified a protein, 15-prostaglandin dehydrogenase (15-PGDH) that is mainly expressed by these rare lesions. We have also shown that this protein is expressed by low grade bladder tumours and that it is externalised to the tumour microenvironment by low grade bladder tumours (Celis

et al., 1996). Thus, 15-PGDH may represent a potential serological candidate both for breast IACs and bladder cancer lesions. Given the rarity of the IACs and the need to have prospective fresh samples to prepare TIF, we have been unable to determine if this protein is externalised to the microenvironment by these lesions. Clearly, as we attain a better classification of breast tumour types using proteomic technologies (Celis et al., 2007a; Gonçalves et al., 2008) we should eventually be able to identify biomarkers that may form the basis for the identification of individual tumour subtypes using a serological test.

Among the validated up-regulated proteins we have identified, a significant number (calreticulin, chloride intracellular channel protein 1, galectin 1, peroxiredoxin-2, platelet-derived endothelial cell growth factor, protein disulfide isomerase, and ubiquitin carboxyl-terminal hydrolase 5) have already been detected in the blood/plasma by others (Table 3), and therefore the next step, which will require prioritization of markers, will have to take into consideration expression by other tissues and tumours, detection in the blood using independent assays, as well as availability of specific antibodies. The biomarker discovery-phase is expected to generate a set of three to five analytically validated biomarkers that will be tested individually or in combination for specificity and sensitivity using blood samples from well-defined cohorts of breast cancer patients, women with benign breast lesions and healthy female blood donors. For each of the selected biomarkers, we will construct conventional ELISA immunoassays to validate specificity, sensitivity, recovery, dilution and reproducibility (Puntmann, 2009). In order for a biomarker to be taken to the next validation step, it must show higher than 70% sensitivity and 98% specificity. Ideally, the markers should be present already in premalignant lesions and studies are currently underway to determine if this is the case for the markers we have already identified.

The next step in our biomarker development pipeline is the validation of selected markers and will include preoperative

samples from 450 women with primary breast cancer, 500 samples from women with benign breast lesions and 1000 samples obtained from healthy female blood donors. In addition, we will analyze a broad spectrum of preoperative blood samples that have already been collected from other cancer patients. Validations will be conducted using an automated antibody array based platform. Biomarkers that make it to the last validation step should in this material show the same degree of sensitivity and specificity as defined for the initial validation (>70% sensitivity and 98% specificity). The samples have been prospectively collected by N. Brnner's group during the last 3 years (Ethical Committee approval and informed consent have been obtained) as part of the long-term discovery-driven biomarker discovery process undertaken by the Danish Centre for Translational Breast Cancer Research (DCTB) that our group is a part of, and the corresponding database of clinical information is located and managed by the Danish Breast Cancer Cooperative Group (DBCG).

4. Experimental procedures

4.1. Sample collection and handling

Tissue biopsies from clinical high-risk patients (high-risk definition according to the Danish Breast Cooperative Group; <http://www.dbcg.dk> last accessed 22.10.2009) that underwent mastectomy were collected from the Pathology Department at the Copenhagen University Hospital. Samples for gel analysis were placed in liquid nitrogen and were rapidly transported to the Institute of Cancer Biology where they were stored at -80°C . Samples for interstitial fluid recovery were kept in PBS at 4°C and were routinely prepared within a maximum of 30–45 min from the time of surgical excision. The project was approved by the Scientific and Ethical Committee of the Copenhagen and Frederiksberg Municipalities (KF 01-069/03).

4.2. Preparation of TIF, NIF and FIF

About 0.1–0.3 g of clean fresh tissue (tumour, normal and fat) was cut to small pieces (3 mm^3), dipped into 2 ml of PBS at 4°C , and placed in a 10-ml conical plastic tube containing 1.0 ml of PBS. Samples were incubated for 1 h at 37°C in a humidified CO_2 incubator. Thereafter the samples were centrifuged at 1000 rpm for 2 min, and the supernatant was aspirated with the aid of an elongated Pasteur pipette. Samples were further centrifuged at 5000 rpm for 20 min in a refrigerated centrifuge (4°C). Protein concentration was determined by the Bradford assay (Bradford, 1976).

4.3. Two-dimensional gel electrophoresis (2D PAGE) and image analysis

4.3.1. 2D PAGE (isoelectric focusing, IEF) was performed as previously described (Celis et al., 2005b).

A. Tissues. Twenty to 30, 6- μm cryostat sections of frozen tissues were resuspended in 0.1 ml lysis solution (O'Farrell, 1975). The resulting lysate was frozen and kept at -20°C until used, usually within 24–48 h (Celis et al., 2005b). A total of 20–

40 μl were applied to the gels and each sample was run at least in duplicate. On average three experimental replicates were performed per sample. The first and last sections of each sample were used for immunofluorescence analysis using cytokeratin 19 (CK19) antibodies as this epithelial marker is ubiquitously expressed by mammary epithelial cells (Moll, 1998). The availability of these pictures greatly facilitated the interpretation of the gel data as it gave a rough estimate of the ratio of glands/tumour cells to stromal tissue.

B. Interstitial fluids (TIF, NIF, and FIF). A fraction of the interstitial fluids (0.2 ml) was freeze-dried resuspended in 0.2 ml of lysis solution (O'Farrell, 1975) and kept at -20°C until use. A total of 20–40 μl were applied to the IEF gels and each sample was run at least in duplicate.

After running the second dimension, gels were placed in 7.5% acetic acid, 50% ethanol, and 0.05% formalin for 1 h, washed $3 \times 30\text{ min}$ in 7.5% acetic acid, 10% ethanol and stained with silver nitrate according to a procedure compatible with mass spectrometry (Gromova and Celis, 2006).

4.3.2. Image analysis

2D gels were scanned and analyzed using the PDQUEST software from BioRad (version 8.0.1). Given the large number of gels that needed to be matched in case of TIFs (69 samples) and NIFs (28 samples) the images were divided into 12 sets of five to six samples in each and the sets were analysed one after the other. A master image was prepared using the gel of TIF 46. All protein spots selected for the analysis were added to the master image. Highly saturated proteins such as albumin, actin and others were excluded from the analysis. In some cases additional spots were found to be saturated and these could not be automatically quantified by Gaussian modelling. In those particular cases, a spot boundary was defined manually that allowed levels of saturated spots to be roughly compared.

4.4. Protein identification by mass spectrometry

Protein spots were excised from silver stained dry gels and the gel pieces were re-hydrated in water. Gel pieces were detached from the cellophane film and cut into 1 mm^2 pieces followed by "in-gel" digestion as previously described (Shevchenko et al., 1996).

4.4.1. Sample preparation

Samples were prepared for analysis by applying 2 μl of digested and extracted peptides on the surface of a 400/384 AnchorChip target plate (Bruker Daltonik, Germany) in duplicate, followed by co-crystallization with 0.35 μl of α -cyano-4-hydroxy-cinnamic acid (HCCA) (3 mg/ml in 66% acetonitrile, 0.5% TFA). Dry samples were washed twice by 2% TFA to remove contaminations.

4.4.2. MALDI measurements

MALDI-TOF-TOF data were acquired using an UltraflexTM III 200 time-of-flight mass spectrometer (Bruker Daltonik, Germany) equipped with a Smart beamTM laser and a LIFT-TOF/TOF unit. Data acquisition and data processing were performed by the Flex Control 3.0 and Flex Analysis 3.0 software (Bruker Daltonik, Germany). All of the spectra were obtained using reflector positive mode with an acceleration voltage of 25 kV,

reflector voltage of 26.38 kV and detection suppressed up to 450 Da. A total of 2000 shots in steps of 200 shots were added to one spectrum in the mass range of m/z 600–4000.

4.4.3. Spectral analysis

Post-acquisition two step calibration was automatically performed in FlexAnalysis using standard peptide calibration mixture (Bruker Daltonics, Germany) for external calibration followed by an additional post-acquisition internal calibration step to obtain better mass accuracies. Ubiquitous presented auto-digested tryptic mass values visible in all the spectra were used for internal calibration. The background masses (matrix, metal adducts, tryptic peptides from keratins) were automatically subtracted from finally generated pick list and were excluded from the further analysis. For protein identification, peptide masses were transferred to the BioTools 3.2 interface (Bruker Daltonik, Germany) to search in the National Center for Biotechnology non-redundant NCBIInr (20090905) database using in house MASCOT search engine (version 2.2, released 28.08.2009, Matrix Science Ltd.). No restriction on the protein molecular mass and taxonomy was applied as a first step. A number of fixed (acrylamide modified cysteine, i.e. propionamide/carbamidomethylation) and variable modifications (methionine oxidation and protein N-terminus acetylation) were included in the search parameters. The peptide tolerance did not exceed 30 ppm and a maximum of one trypsin missed cleavage was allowed. Protein identifications were considered to be confident when the protein score of the hit exceeded the threshold significance score of 65 ($p < 0.05$) and no less than six peptides were recognized. Often the peptides identified matched equally well to multiple database entries using the non-redundant NCBIInr database that is why the second/final search was performed using the same parameters, but the UniProtKB/SwissProt 55.3 (19372 human entries) database, and if the number and the sequence of the recognized peptides were identical to the first search, the SwissProt accession number was assigned for the identified protein. Whenever it was possible or necessary, in case of low significance score for PMF, the MS/MS analysis by LIFT technology was additionally performed using the same search parameters where peptide/fragment tolerance did not exceed 30 ppm and 0.5 Da correspondently.

4.5. LS-MS analysis of TIF samples

4.5.1. Sample preparation

A total of 0.8 ml of TIF sample from patient 68, containing the highest protein content was lyophilized and then dissolved in 0.4 ml of Milli-Q H₂O. The remaining pellet, contained undissolved proteins, was dissolved in lysis buffer for 1st dimensional gel electrophoresis (see below). Aliquots – 40 μ l per tube – were stored at -80°C before use. Four strategies were selected for protein separation/enrichment before LS-MS/MS analysis:

4.5.2. Separation of solubilised TIF proteins by 1D-gel followed by tryptic digestion

Twenty μ l of solubilised TIF proteins were separated on a 4–20% Tris-Glycine gel and stained with silver nitrate as described above. The gel was divided into 4 mm bands and tryptic digestion of proteins was performed as described above. After tryptic digestion the peptide extraction was performed using 5%

formic acid. The collected supernatants were lyophilized and stored at -20°C until LC-MS analysis.

4.5.3. Separation of solubilised TIF proteins by isoelectric focusing after albumin and IgG depletion

Forty μ l of solubilised TIF proteins were depleted of albumin and major immunoglobulins by affinity chromatography using the albumin depletion kit from Vivascience (Germany) following the manufacturers protocol. The resulting protein fractions were combined and dialysed against H₂O at 4°C . Dialysed proteins were freeze-dried, dissolved in lysis buffer for 1st dimensional electrophoresis and separated by isoelectric focusing (see 1st 2D PAGE dimension). The gel-stripe was cut into 11 bands and tryptic digestion of proteins was performed as described above.

4.5.4. Separation of TIF proteins recovered from pellet by isoelectric focusing

Proteins recovered from TIF pellet (see above) were separated directly by isoelectric focusing (see above 1st 2D PAGE dimension) followed by the same procedure as described in above.

4.5.5. “In-solution” tryptic digestion of ACN-precipitated proteins and SCX-fractionation

Forty μ l of solubilised TIF proteins were precipitated by addition of 160 μ l acetonitrile. After incubation during 2 h at -18°C the precipitate was separated from the supernatant by centrifugation. The proteins were lyophilized and resuspended in 21 μ l 400 mM NH₄HCO₃ in 8 M urea. Five μ l 45 mM DTT was added and the mixture was incubated for 15 min at 50°C , cooled to room temperature and 5 μ l of 100 mM iodoacetamide was added. After 15 min incubation in the dark 140 μ l of H₂O and 2 μ g sequence grade trypsin (Promega, Madison, WI) were added and incubated overnight at 37°C . The peptides were cleaned on C18 material and separated by SCX-fractionation.

Peptides were separated by SCX using an Ettan LC system (Amersham Pharmacia Biotech, Uppsala, Sweden). The peptide samples were diluted 50 times in solvent A (5 mM ammonium formate, 30% ACN adjusted to pH 3 with formic acid (FA) and loaded onto a 150 mm long 2.1 mm ID Polysulfoethyl A (PolyLC, Columbia, MD) column. Peptides were separated in a 20 min gradient at a constant flow rate of 0.2 mL/min from 0 to 40% solvent B (1.0 M ammonium formate, 30% ACN adjusted to pH 3 with FA), and 10–20 fractions were collected. The volume of each fraction was reduced by vacuum centrifugation to remove the ACN, loaded onto a reversed-phase stage-tip made of C18 disk material from Empore (3M, Minneapolis, MN), washed with 1% TFA, and eluted with 0.5% acetic acid and 80% ACN. The samples were reduced to a small volume by vacuum centrifugation, diluted to 10 μ l in 0.5% acetic acid and analysed by LC-MS.

4.5.6. Liquid chromatography-tandem mass spectrometry

Peptide separation, recovered from all strategies described above, was achieved using LC Packings nanoflow LC system (LC Packings, Amsterdam, The Netherlands) equipped with a Famos autosampler. Custom-made pre-columns (2 cm, 75 μ m ID fused silica) and analytical columns (10 cm, 50 μ m ID fused silica) were used with silica frits retaining the Reprosil-Pur C18 3 μ m beads (Dr. Maisch GmbH, Ammerbuch-

Entringen, Germany). Peptides were loaded onto the pre-column with a flow of 5 μ L/min. After loading, nanoflow reversed-phase LC was performed using a flow split before the pre-column to achieve a flow of approximately 200 nL/min going through the pre-column and the analytical column directly into the ESI source of a Q-TOF Ultima tandem mass spectrometer (Waters Corp., Milford, MA) using a linear gradient of 5–40% ACN in 90 min with 0.5% aqueous acetic acid (Merck) as solvent A and 0.5% aqueous acetic acid and 80% ACN as solvent B. The mass spectrometer acquired 1 survey spectrum of the *m/z* range 300–1700 followed by MS/MS of the three most intense multiply charged ions. In the window experiment the 10 SCX-fractions were acquired with survey spectra in *m/z* range 300–600 and 600–1700 in two runs.

4.5.7. Spectral analysis and database searching

The tandem mass spectra acquired during LC-MS/MS were smoothed, centroided and converted to peak-list text files using MassLynx 3.5 (Waters/Micromass). The spectral data was searched against the human subset of the non-redundant databases SWISS-PROT and TREMBL release 48 using Mascot version 2.1 (Matrix Sciences Ltd., Cheshire, UK). The following search parameters were used: variable modifications were carbamidomethylation or propionamide of cysteine residues, N-terminal acetylation of proteins, possible oxidation of methionine residues, carbamoylation of N-terminal amines, deamidation of NQ, possible conversion of terminal glutamine or glutamate residues into pyroglutamate, at most two missed tryptic cleavage sites, semi-tryptic cleavage restraints, a 0.15 Da error tolerance in both MS and MS/MS. A Mascot score of 25 or a sequence tag equal to or longer than four amino acids were used as acceptance criteria for peptide identification.

In summary, 832 unique proteins were identified based on 11,927 peptide sequences. The complete information of MS-MS data is presented in [Supplementary Table 2](#).

4.6. Antibodies

Antibodies against the following proteins were used in this study: A-FABP (rabbit polyclonal anti peptide: Eurogentec, Seraing, Belgium, used at a dilution of 1:200), calreticulin (rabbit polyclonal antibody 39818; Abcam, Cambridge, UK, used at a dilution 1:200), calumenin (rabbit polyclonal obtained from Reider Wallin, NC, used at a dilution of 1:1000), chloride intracellular channel protein 1 (rabbit polyclonal, AP7589a, ABGENT, San Diego, CA, used at a dilution of 1:250), cytokeratin 19 (mouse monoclonal clone A53-B/A2.26; Thermo Scientific, Fremont, CA, used at a dilution of 1:1000), elongation factor 1b (rabbit polyclonal ab 37957; Abcam, Cambridge, UK, used at a dilution of 1:1500), galectin 1 (goat polyclonal antibody (AF 1152); R & D Systems, MN, used at a dilution of 1:250), p53 (mouse monoclonal clone DO-7, DAKO, Copenhagen, Denmark, used at a dilution of 1:250), peroxiredoxin-2, (mouse monoclonal clone 4E10-2D2; ABGENT, San Diego, CA, used at a dilution of 1:3000), platelet-derived endothelial cell growth factor (mouse monoclonal clone P-GF.44C; IMGEX, San Diego, CA, used at a dilution of 1:300), protein disulfide isomerase (mouse monoclonal clone 1D3 (SPA-891); Assay Designs Stressgen, Ann Arbor, used at a dilution of 1:800), S100A9, MRP14, calgranulin (mouse monoclonal clone 47-8D3; Abcam, Cambridge, UK,

used at a dilution of 1:1000), translationally-controlled tumour protein (rabbit polyclonal antibody, Abcam, Cambridge, UK, used at a dilution of 1:2000), tropomyosin alpha-4 chain (mouse monoclonal clone AT4319a; ABGENT, San Diego, CA, used at a dilution 1:200) and ubiquitin carboxyl-terminal hydrolase 5 (rabbit polyclonal antibody SH0310161; ABGENT, San Diego, CA, used at a dilution of 1:50).

4.7. Immunohistochemistry

Following surgery, fresh tissue blocks were immediately placed in formalin fixative and paraffin-embedded for archival use. Six- μ m sections were cut from the tissue blocks and mounted on Super Frost Plus slides (Menzel-Gläser, Braunschweig, Germany), baked at 60 °C for 60 min, deparaffinised, and rehydrated through graded alcohol rinses ([Honma et al., 2005; Celis et al., 2008](#)). Heat induced antigen retrieval was performed by immersing slides in 10 mM citrate buffer (pH 6.0) and microwaving in a 750 W microwave oven for 10 min. The slides were then cooled at room temperature for 20 min and rinsed abundantly in tap water. Non-specific staining of slides was blocked (10% normal goat serum in PBS buffer) for 15 min, and endogenous peroxidase activity quenched using 0.3% H₂O₂ in methanol for 30 min. Antigen was detected with a relevant primary antibody, followed by a suitable secondary antibody conjugated to a peroxidase complex (HRP conjugated goat anti-rabbit or anti-mouse antibody (DakoCytomation, Glostrup, Denmark). Finally, colour development was done with 3,3'-diaminobenzidine (Pierce, IL) as a chromogen to detect bound antibody complex. Slides were counterstained with haematoxylin. Standardization of the dilution, incubation, and development times appropriate for each antibody allowed an accurate comparison of expression levels in all cases.

4.8. TMAs

Breast cancer tissue microarray slides were obtained from Pantomics (BRC1503 – Pantomics Inc., CA). Consecutive sections of the BRC1503 breast tissue array were used. These arrays contain 150 cores, including 75 cases in duplicates of normal (*n* = 1), benign lesions (*n* = 4) and malignant (*n* = 70, various grades and stages) tissues of the breast. All the tissues were from surgical resection. They were fixed in 10% neutral buffered formalin for 24 h and processed using identical SOPs. Sections were picked onto Super Frost Plus slides. Each slide has >95% tissue core retention. The slides were stained as above using an appropriate primary antibody. For detection of immune complexes we used a horseradish peroxidase-labelled polymer (Envision+ detection kit, DAKO, Denmark) as a secondary antibody. All slides were independently reviewed by two of the authors (JEC and JAM) and in the few discrepant cases a consensus was reached after joint review.

4.9. Indirect immunofluorescence analysis

Immunofluorescence analysis was performed essentially as described ([Moreira et al., 2005; Celis et al., 2007a](#)). Briefly, five- μ m sections were cut from paraffin blocks of breast tissue samples mounted on Super Frost Plus slides (Menzel-Gläser, Braunschweig, Germany), baked at 60 °C for 60 min,

deparaffinised, and rehydrated through graded alcohol rinses. Heat induced antigen retrieval was carried out as described above. Following antigen retrieval, sections were treated with Image-iT FX™ signal enhancer (Molecular Probes, OR) to block non-specific staining and sequentially incubated with A-FABP and CK19 specific antibodies (dilutions 1:250 and 1:2000, respectively). Detection of immune complexes was done with species-specific secondary antibodies conjugated to Alexa Fluor® 488 and Alexa Fluor® 594 (Molecular Probes). Nuclear material was counterstained with TO-PRO-3™. The sections were washed three times with cold phosphate buffered saline (PBS) between incubations. Normal rabbit or mouse sera instead of primary antibody were used as a negative control. Sections were imaged using a Zeiss LSM510Meta confocal laser scanning microscope (Carl Zeiss MicroImaging GmbH, Germany).

Acknowledgements

We would like to thank Dorte Holm, Lene Jørgensen, Sofia Svensson, and Signe Trentemøller for expert technical assistance. This work was supported by the Danish Cancer Society through the budget of the Institute of Cancer Biology, and by grants from the John and Birthe Meyer Foundation, Novo Nordisk, the Kai Lange and Gundhild Kai Lange Fond, the Saint Albans Church, the Lisa and Gudmund Jørgensens Fond and the “Race against Breast Cancer”. The earmarked support of the Marketing Department at the Danish Cancer Society through their fundraising activities on behalf of DCTB is greatly appreciated.

Appendix. Supplementary material

Supplementary material associated with this article can be found in the online version at [doi:10.1016/j.molonc.2009.11.003](https://doi.org/10.1016/j.molonc.2009.11.003).

REFERENCES

- Aebersold, R., Mann, M., 2003. Mass spectrometry-based proteomics. *Nature* 422, 198–207.
- Agha-Hosseini, F., Mirzaii-Dizgah, I., Rahimi, A., 2009. Correlation of serum and salivary CA15-3 levels in patients with breast cancer. *Med. Oral. Patol. Oral. Cir. Bucal.* 14, e521–e524.
- Anderson, N.L., Polanski, M., Pieper, R., Gatlin, T., Tirumalai, R.S., Conrads, T.P., Veenstra, T.D., Adkins, J.N., Pounds, J.G., Fagan, R., Lobley, A., 2004. The human plasma proteome: a nonredundant list developed by combination of four separate sources. *Mol. Cell. Proteomics* 3, 311–326.
- Anderson, N.L., Anderson, N.G., Pearson, T.W., Borchers, C.H., Paulovich, A.G., Patterson, S.D., Gillette, M., Aebersold, R., Carr, S.A., 2009. A human proteome detection and quantitation project. *Mol. Cell. Proteomics* 8, 883–886.
- Apweiler, R., Aslanidis, C., Deufel, T., Gerstner, A., Hansen, J., Hochstrasser, D., Kellner, R., Kubicek, M., Lottspeich, F., Maser, E., Mewes, H.W., Meyer, H.E., Müllner, S., Mutter, W., Neumaier, M., Nollau, P., Nothwang, H.G., Ponten, F., Radbruch, A., Reinert, K., Rothe, G., Stockinger, H., Tarnok, A., Taussig, M.J., Thiel, A., Thiery, J., Ueffing, M., Valet, G., Vandekerckhove, J., Verhuven, W., Wagener, C., Wagner, O., Schmitz, G., 2009. Approaching clinical proteomics: current state and future fields of application in fluid proteomics. *Clin. Chem. Lab. Med.* 47, 724–744.
- Bini, L., Magi, B., Marzocchi, B., Arcuri, F., Tripodi, S., Cintonaro, M., Sanchez, J.C., Frutiger, S., Hughes, G., Pallini, V., Hochstrasser, D.F., Tosi, P., 1997. Protein expression profiles in human breast ductal carcinoma and histologically normal tissue. *Electrophoresis* 18, 2832–2841.
- Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254.
- Brenner, J., Sawant, S.G., Hande, M.P., Miller, R.C., Elliston, C.D., Fu, Z., Randers-Pehrson, G., Marino, S.A., 2002. Routine screening mammography: how important is the radiation-risk side of the benefit-risk equation? *Int. J. Radiat. Biol.* 78, 1065–1067.
- Brooks, M., 2009. Breast cancer screening and biomarkers. *Methods Mol. Biol.* 472, 307–321.
- Casiano, C.A., Mediavilla-Varela, M., Tan, E.M., 2006. Tumor-associated antigen arrays for the serological diagnosis of cancer. *Mol. Cell. Proteomics* 5, 1745–1759.
- Celis, J.E., Ratz, G.P., Celis, A., 1987. Secreted proteins from normal and SV40 transformed human MRC-5 fibroblasts: toward establishing a database of human secreted proteins. *Leukemia* 1, 707–717.
- Celis, J.E., Ratz, G.P., Madsen, P., Gesser, B., Lauridsen, J.B., Hansen, K.P., Kwee, S., Rasmussen, H.H., Nielsen, H.V., Crüger, D., et al., 1989. Computerized, comprehensive databases of cellular and secreted proteins from normal human embryonic lung MRC-5 fibroblasts: identification of transformation and/or proliferation sensitive proteins. *Electrophoresis* 10, 76–115.
- Celis, J.E., Ostergaard, M., Basse, B., Celis, A., Lauridsen, J.B., Ratz, G.P., Andersen, I., Hein, B., Wolf, H., Orntoft, T.F., Rasmussen, H.H., 1996. Loss of adipocyte-type fatty acid binding protein and other protein biomarkers is associated with progression of human bladder transitional cell carcinomas. *Cancer Res.* 56, 4782–4790.
- Celis, A., Rasmussen, H.H., Celis, P., Basse, B., Lauridsen, J.B., Ratz, G., Hein, B., Ostergaard, M., Wolf, H., Orntoft, T., Celis, J.E., 1999. Short-term culturing of low-grade superficial bladder transitional cell carcinomas leads to changes in the expression levels of several proteins involved in key cellular activities. *Electrophoresis* 20, 355–361.
- Celis, J.E., Gromov, P., Cabezon, T., Moreira, J.M., Ambartsumian, N., Sandelin, K., Rank, F., Gromova, I., 2004. Proteomic characterization of the interstitial fluid perfusing the breast tumor microenvironment: a novel resource for biomarker and therapeutic target discovery. *Mol. Cell. Proteomics* 3, 327–344.
- Celis, J.E., Moreira, J.M., Cabezon, T., Gromov, P., Friis, E., Rank, F., Gromova, I., 2005a. Identification of extracellular and intracellular signaling components of the mammary adipose tissue and its interstitial fluid in high risk breast cancer patients: toward dissecting the molecular circuitry of epithelial-adipocyte stromal cell interactions. *Mol. Cell. Proteomics* 4, 492–522.
- Celis, J.E., Trentemølle, S., Gromov, P., 2005b. Gel-based proteomics: high-resolution two-dimensional gel electrophoresis of proteins. Isoelectric focusing (IEF) and nonequilibrium pH gradient electrophoresis (NEPHGE). In: Celis, J.E., Carter, N., Hunter, T., Shotton, D., Simons, K., Small, J.V. (Eds.), *Cell Biology. A Laboratory Handbook*, 4. Elsevier, San Diego, CA, pp. 165–177.
- Celis, J.E., Gromova, I., Cabezon, T., Gromov, P., Shen, T., Timmermans-Wielenga, V., Rank, F., Moreira, J.M., 2007a. Identification of a subset of breast carcinomas characterized

- by expression of cytokeratin 15: relationship between CK15+ progenitor/amplified cells and pre-malignant lesions and invasive disease. *Mol. Oncol* 1, 321–349.
- Celis, J.E., Moreira, J.M., Gromova, I., Cabezón, T., Gromov, P., Shen, T., Timmermans, V., Rank, F., 2007b. Characterization of breast precancerous lesions and myoepithelial hyperplasia in sclerosing adenosis with apocrine metaplasia. *Mol. Oncol* 1, 97–119.
- Celis, J.E., Gromov, P., Cabezón, T., Moreira, J.M., Friis, E., Jirstrom, K., Llombart-Bosch, A., Timmermans-Wielenga, V., Rank, F., Gromova, I., 2008. 15-prostaglandin dehydrogenase expression alone or in combination with ACSM1 defines a subgroup of the apocrine molecular subtype of breast carcinoma. *Mol. Cell. Proteomics* 7, 1795–1809.
- Celis, J.E., Cabezón, T., Moreira, J.M., Gromov, P., Gromova, I., Timmermans-Wielenga, V., Iwase, T., Akiyama, F., Honma, N., Rank, F., 2009. Molecular characterization of apocrine carcinoma of the breast: Validation of an apocrine protein signature in a well-defined cohort. *Mol. Oncol.* Feb 3. [Epub ahead of print].
- Chahed, K., Kabbage, M., Hamrita, B., Guillier, C.L., Trimeche, M., Remadi, S., Ehret-Sabatier, L., Chouchane, L., 2008. Detection of protein alterations in male breast cancer using two dimensional gel electrophoresis and mass spectrometry: the involvement of several pathways in tumorigenesis. *Clin. Chim. Acta* 388, 106–114.
- Chapman, C., Murray, A., Chakrabarti, J., Thorpe, A., Woolston, C., Sahin, U., Barnes, A., Robertson, J., 2007. Autoantibodies in breast cancer: their use as an aid to early diagnosis. *Ann. Oncol* 18, 868–873.
- Chen, E.I., Yates 3rd, J.R., 2007. Cancer proteomics by quantitative shotgun proteomics. *Mol. Oncol* 1, 144–159.
- Conrotto, P., Souchelnytskyi, S., 2008. Proteomic approaches in biological and medical sciences: principles and applications. *Exp. Oncol* 30, 171–180.
- Decramer, S., Gonzalez de Peredo, A., Breuil, B., Mischak, H., Monsarrat, B., Bascands, J.L., Schanstra, J.P., 2008. Urine in clinical proteomics. *Mol. Cell. Proteomics* 7, 1850–1862.
- Dede, D.S., Arslan, C., Altundag, K., 2009. Serum levels of CEA and CA 15-3 in triple-negative breast cancer at the time of diagnosis. *Med. Oncol. Sep* 23. [Epub ahead of print].
- Deng, S.S., Xing, T.Y., Zhou, H.Y., Xiong, R.H., Lu, Y.G., Wen, B., Liu, S.Q., Yang, H.J., 2006. Comparative proteome analysis of breast cancer and adjacent normal breast tissues in human. *Genom. Proteom. Bioinform* 4, 165–172.
- Desmetz, C., Bascoul-Mollevi, C., Rochaix, P., Lamy, P.-J., Kramar, A., Rouanet, P., Maudelonde, T., Mange, A., Solassol, J., 2009. Identification of a new panel of serum autoantibodies associated with the presence of in situ carcinoma of the breast in younger women. *Clin. Cancer Res.* 15, 4733–4741.
- Doustjalali, S.R., Yusof, R., Yip, C.H., Looi, L.M., Pillay, B., Hashim, O.H., 2004. Aberrant expression of acute-phase reactant proteins in sera and breast lesions of patients with malignant and benign breast tumors. *Electrophoresis* 25, 2392–2401.
- Duffy, M.J., 2006. Serum tumor markers in breast cancer: are they of clinical value? *Clin. Chem.* 52, 345–351.
- Elmore, J.G., Armstrong, K., Lehman, C.D., Fletcher, S.W., 2005. Screening for breast cancer. *JAMA* 293, 1245–1256.
- Elston, C.W., Ellis, I.O., 1991. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 19, 403–410.
- Faça, V.M., Ventura, A.P., Fitzgibbon, M.P., Pereira-Faça, S.R., Pitteri, S.J., Green, A.E., Ireton, R.C., Zhang, Q., Wang, H., O'Brian, K.C., Drescher, C.W., Schummer, M., McIntosh, M.W., Knudsen, B.S., Hanash, S.M., 2008. Proteomic analysis of ovarian cancer cells reveals dynamic processes of protein secretion and shedding of extra-cellular domains. *PLoS One* 3 (6), e2425.
- Fass, L., 2008. Imaging and cancer: a review. *Mol. Oncol* 2, 115–152.
- Ferlay, J., Autier, P., Boniol, M., Heanue, M., Colombet, M., Boyle, P., 2007. Estimates of the cancer incidence and mortality in Europe in 2006. *Ann. Oncol* 18, 581–592.
- Franzén, B., Linder, S., Alaiya, A.A., Eriksson, E., Fujioka, K., Bergman, A.C., Jörnval, H., Auer, G., 1997. Analysis of polypeptide expression in benign and malignant human breast lesions. *Electrophoresis* 18, 582–587.
- Gagnon, A., Shi, Q., Ye, B., 2008. Surface-enhanced laser desorption/ionization mass spectrometry for protein and Peptide profiling of body fluids. *Methods Mol. Biol.* 441, 41–56.
- Garrisi, V.M., Abbate, I., Quaranta, M., Mangia, A., Tommasi, S., Paradiso, A., 2008. SELDI-TOF serum proteomics and breast cancer: which perspective? *Expert Rev. Proteomics* 5, 779–785.
- Gast, M.C., van Dulken, E.J., van Loenen, T.K., Kingma-Vegter, F., Westerga, J., Flohil, C.C., Knol, J.C., Jimenez, C.R., van Gils, C.H., Wessels, L.F., Schellens, J.H., Beijnen, J.H., 2009a. Detection of breast cancer by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry tissue and serum protein profiling. *Int. J. Biol. Markers* 24, 130–141.
- Gast, M.C., Van Gils, C.H., Wessels, L.F., Harris, N., Bonfrer, J.M., Rutgers, E.J., Schellens, J.H., Beijnen, J.H., 2009b. Serum protein profiling for diagnosis of breast cancer using SELDI-TOF MS. *Oncol Rep* 22, 205–213.
- Gillet, D., Kennedy, C., Carmalt, H., 1997. Breast cancer in young women. *Aust. N.Z.J. Surg* 67, 761–764.
- Gion, M., Mione, R., Leon, A.E., Lüftner, D., Molina, R., Possinger, K., Robertson, J.F., 2001. CA27.29: a valuable marker for breast cancer management. A confirmatory multicentric study on 603 cases. *Eur. J. Cancer* 37, 355–363.
- Gonçalves, A., Charafe-Jauffret, E., Bertucci, F., Audebert, S., Toiron, Y., Esterni, B., Monville, F., Tarpin, C., Jacquemier, J., Houvenaeghel, G., Chabannon, C., Extra, J.M., Viens, P., Borg, J.P., Birnbaum, D., 2008. Protein profiling of human breast tumor cells identifies novel biomarkers associated with molecular subtypes. *Mol. Cell. Proteomics* 7, 1420–1433.
- Gromov, P., Celis, J.E., Gromova, I., Rank, F., Timmermans-Wielenga, V., Moreira, J.M., 2008. A single lysis solution for the analysis of tissue samples by different proteomic technologies. *Mol. Oncol* 2, 368–379.
- Gromova, I., Celis, J.E., 2006. Protein detection in gels by silver staining: a procedure compatible with mass-spectrometry. In: Celis, J.E., Carter, N., Hunter, T., Shotton, D., Simons, K., Small, J.V. (Eds.), *Cell Biology. A Laboratory Handbook*, vol. 4. Academic Press, San Diego, pp. 219–223.
- Groves, M.D., Hess, K.R., Puduvalli, V.K., Colman, H., Conrad, C.A., Gilbert, M.R., Weinberg, J., Cristofanilli, M., Yung, W.K., Liu, T.J., 2009. Biomarkers of disease: cerebrospinal fluid vascular endothelial growth factor (VEGF) and stromal cell derived factor (SDF)-1 levels in patients with neoplastic meningitis (NM) due to breast cancer, lung cancer and melanoma. *J. Neurooncol* 94, 229–234.
- Haab, B.B., 2005. Antibody arrays in cancer research. *Mol. Cell. Proteomics* 4, 377–383.
- Hamrita, B., Chahed, K., Kabbage, M., Guillier, C.L., Trimeche, M., Chaïeb, A., Chouchane, L., 2008. Identification of tumor antigens that elicit a humoral immune response in breast cancer patients' sera by serological proteome analysis (SERPA). *Clin. Chim. Acta* 393, 95–102.
- Hanash, S.M., Pitteri, S.J., Faca, V.M., 2008. Mining the plasma proteome for cancer biomarkers. *Nature* 452, 571–579.
- Harter, C., Reinhard, C., 2000. The secretory pathway from history to the state of the art. *Subcell Biochem* 34, 1–38.

- Hayes, D.F., 1996. Serum (circulating) tumor markers for breast cancer. *Recent Results Cancer Res.* 140, 101–113.
- Hegmans, J.P., Bard, M.P., Hemmes, A., Luider, T.M., Willemsen, R., Severijnen, L.A., van Meerbeeck, J.P., Burgers, S.A., Hoogsteden, H.C., Lambrecht, B.N., 2004. Proteomic analysis of exosomes isolated from human malignant pleural effusions. *Am. J. Pathol.* 164, 1807–1815.
- Honma, N., Takubo, K., Akiyama, F., Sawabe, M., Arai, T., Younes, M., Kasumi, F., Sakamoto, G., 2005. Expression of GCDPF-15 and AR decreases in larger or node-positive apocrine carcinomas of the breast. *Histopathology* 47, 195–201.
- Hu, S., Loo, J.A., Wong, D.T., 2006. Human body fluid proteome analysis. *Proteomics* 6, 6326–6353.
- Huang, R., Lin, Y., Shi, Q., Flowers, L., Ramachandran, S., Horowitz, I.R., Parthasarathy, S., Huang, R.P., 2004. Enhanced protein profiling arrays with ELISA-based amplification for high-throughput molecular changes of tumor patients' plasma. *Clin. Cancer Res.* 10, 598–609.
- Humphrey, L.L., Helfand, M., Chan, B.K., Woolf, S.H., 2002. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann. Intern. Med.* 137, 347–360.
- Issaq, H.J., Xiao, Z., Veenstra, T.D., 2007. Serum and plasma proteomics. *Chem. Rev.* 107, 3601–3620.
- Jacobs, J.M., Waters, K.M., Kathmann, L.E., Camp 2nd, D.G., Wiley, H.S., Smith, R.D., Thrall, B.D., 2008. The mammary epithelial cell secretome and its regulation by signal transduction pathways. *J. Proteome Res.* 7, 558–569.
- Jeffery, J., Murphy, M.J., 2001. Ascitic fluid analysis: the role of biochemistry and haematology. *Hosp. Med* 62, 282–286.
- Kenny, P.A., Lee, G.Y., Myers, C.A., Neve, R.M., Semeiks, J.R., Spellman, P.T., Lorenz, K., Lee, E.H., Barcellos-Hoff, M.H., Petersen, O.W., Gray, J.W., Bissell, M.J., 2007. The morphologies of breast cancer cell lines in three-dimensional assays correlate with their profiles of gene expression. *Mol. Oncol* 1, 84–96.
- Kilic, A., Schuchert, M.J., Luketich, J.D., Landreneau, R.J., Lokshin, A.E., Bigbee, W.L., El-Hefnawy, T.J., 2008. Use of novel autoantibody and cancer-related protein arrays for the detection of esophageal adenocarcinoma in serum. *Thorac. Cardiovasc. Surg* 136, 199–204.
- Kim, H.S., Park, Y.H., Park, M.J., Chang, M.H., Jun, H.J., Kim, K.H., Ahn, J.S., Kang, W.K., Park, K., Im, Y.H., 2009. Clinical significance of a serum CA15-3 surge and the usefulness of CA15-3 kinetics in monitoring chemotherapy response in patients with metastatic breast cancer. *Breast Cancer Res. Treat* 118, 89–97.
- Koga, K., Matsumoto, K., Akiyoshi, T., Kubo, M., Yamanaka, N., Tasaki, A., Nakashima, H., Nakamura, M., Kuroki, S., Tanaka, M., Katano, M., 2005. Purification, characterization and biological significance of tumor-derived exosomes. *Anticancer Res.* 25, 3703–3707.
- Kulasingam, V., Diamandis, E.P., 2007. Proteomics analysis of conditioned media from three breast cancer cell lines: a mine for biomarkers and therapeutic targets. *Mol. Cell. Proteomics* 6, 1997–2011.
- Kulasingam, V., Diamandis, E.P., 2008. Tissue culture-based breast cancer biomarker discovery platform. *Int. J. Cancer* 123, 2007–2012.
- Liang, X., Huuskonen, J., Hajivandi, M., Manzanedo, R., Predki, P., Amshey, J.R., Pope, R.M., 2009. Identification and quantification of proteins differentially secreted by a pair of normal and malignant breast-cancer cell lines. *Proteomics* 9, 182–193.
- Liotta, L.A., Kohn, E.C., 2001. The microenvironment of the tumour-host interface. *Nature* 411, 375–379.
- Liu, X., Feng, Q., Chen, Y., Zuo, J., Gupta, N., Chang, Y., Fang, F., 2009. Proteomics-based identification of differentially-expressed proteins including galectin-1 in the blood plasma of type 2 diabetic patients. *J. Proteome Res.* 8, 1255–1262.
- Loch, C.M., Ramirez, A.B., Liu, Y., Sather, C.L., Delrow, J.J., Scholler, N., Garvik, B.M., Urban, N.D., McIntosh, M.W., Lampe, P.D., 2007. Use of high density antibody arrays to validate and discover cancer serum biomarkers. *Mol. Oncol.* 1, 313–320.
- Lu, H., Goodell, V., Disis, M.L., 2008. Humoral immunity directed against tumor-associated antigens as potential biomarkers for the early diagnosis of cancer. *J. Proteome Res.* 7, 1388–1394.
- Lumachi, F., Basso, S.M., 2004. Serum tumor markers in patients with breast cancer. *Expert Rev. Anticancer Ther.* 4, 921–931.
- Lunt, S.J., Fyles, A., Hill, R.P., Milosevic, M., 2008. Interstitial fluid pressure in tumors: therapeutic barrier and biomarker of angiogenesis. *Future Oncol* 4, 793–802.
- Mannello, F., Medda, V., Tonti, G.A., 2009. Protein profile analysis of the breast microenvironment to differentiate healthy women from breast cancer patients. *Expert Rev. Proteomics* 6, 43–60.
- Martín, A., Corte, M.D., Alvarez, A.M., Rodriguez, J.C., Andicochea, A., Bongera, M., Junquera, S., Pidal, D., Allende, T., Muñoz, J.L., Vizoso, F., 2006. Prognostic value of pre-operative serum CA 15.3 levels in breast cancer. *Anticancer Res.* 26, 3965–3971.
- Maurya, P., Meleady, P., Dowling, P., Clynes, M., 2007. Proteomic approaches for serum biomarker discovery in cancer. *Anticancer Res.* 27, 1247–1255.
- Mbeunkui, F., Metge, B.J., Shevde, L.A., Pannell, L.K., 2007. Identification of differentially secreted biomarkers using LC-MS/MS in isogenic cell lines representing a progression of breast cancer. *J. Proteome Res.* 6, 2993–3002.
- Moreira, J.M., Ohlsson, G., Rank, F.E., Celis, J.E., 2005. Down-regulation of the tumor suppressor protein 14-3-3sigma is a sporadic event in cancer of the breast. *Mol. Cell. Proteomics* 4, 555–569.
- Moll, R., 1998. Cytokeratins as markers of differentiation in the diagnosis of epithelial tumors. *Sub-cellular Biochem* 31, 205–262.
- Musselman, I., Speicher, D.W., 2005. Human serum and plasma proteomics. *Curr. Protoc. Protein Sci.* Chapter 24, unit 24.1.
- Nickel, W., 2003. The mystery of nonclassical protein secretion. A current view on cargo proteins and potential export routes. *Eur. J. Biochem* 270, 2109–2119.
- Nothacker, M., Duda, V., Hahn, M., Warm, M., Degenhardt, F., Madjar, H., Weinbrenner, S., Albert, U.S., 2009. Early detection of breast cancer: benefits and risks of supplemental breast ultrasound in asymptomatic women with mammographically dense breast tissue. A systematic review. *BMC Cancer* 9, 335.
- O'Farrell, P.H., 1975. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250, 4007–4021.
- Okano, T., Kondo, T., Kakisaka, T., Fujii, K., Yamada, M., Kato, H., Nishimura, T., Gemma, A., Kudoh, S., Hirohashi, S., 2006. Plasma proteomics of lung cancer by a linkage of multi-dimensional liquid chromatography and two-dimensional difference gel electrophoresis. *Proteomics* 6, 3938–3948.
- Oliver, C., Vidal, M., 2007. Proteomic analysis of secreted exosomes. *Subcell. Biochem* 43, 99–131.
- Omenn, G.S., 2004. Advancement of biomarker discovery and validation through the HUPO plasma proteome project. *Dis. Markers* 20, 131–134.
- Ong, S.E., Mann, M., 2005. Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* 1, 252–262.
- Paliouras, M., Diamandis, E.P., 2008. Androgens act synergistically to enhance estrogen-induced upregulation of human tissue kallikreins 10, 11, and 14 in breast cancer cells via a membrane bound androgen receptor. *Mol. Oncol* 1, 413–424.
- Pan, C., Kumar, C., Bohl, S., Klingmueller, U., Mann, M., 2009. Comparative proteomic phenotyping of cell lines and primary cells to assess preservation of cell type-specific functions. *Mol. Cell. Proteomics* 8, 443–450.
- Pisano, E., 2005. Issues in breast cancer screening. *Technol. Cancer Res. Treat* 4, 5–9.

- Polanski, M., Anderson, N.L., 2006. A list of candidate cancer biomarkers for targeted proteomics. *Biomarker Insights* 1, 1–48.
- Puntmann, V.O., 2009. How-to guide on biomarkers: biomarker definitions, validation and applications with examples from cardiovascular disease. *Postgrad. Med. J.* 85, 538–545.
- Qiu, J., Choi, G., Li, L., Wang, H., Pitteri, S.J., Pereira-Faca, S.R., Krasnoselsky, A.L., Randolph, T.W., Omenn, G.S., Edelstein, C., Barnett, M.J., Thornquist, M.D., Goodman, G.E., Brenner, D.E., Feng, Z., Hanash, S.M., 2008. Occurrence of autoantibodies to annexin I, 14-3-3 θ and LAMR1 in prediagnostic lung cancer sera. *J. Clin. Oncol* 26, 5060–5066.
- Rajcevic, U., Niclou, S.P., Jimenez, C.R., 2009. Proteomics strategies for target identification and biomarker discovery in cancer. *Front. Biosci* 14, 3292–3303.
- Rakha, E.A., Ellis, I.O., 2009. Triple-negative/basal-like breast cancer: review. *Pathology* 41, 40–47.
- Rifai, N., Gillette, M.A., Carr, S.A., 2006. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat. Biotechnol* 24, 971–983.
- Rose, K., Bougueleret, L., Baussant, T., Böhm, G., Botti, P., Colinge, J., Cusin, I., Gaertner, H., Gleizes, A., Heller, M., Jimenez, S., Johnson, A., Kussmann, M., Menin, L., Menzel, C., Ranno, F., Rodriguez-Tomé, P., Rogers, J., Saudrais, C., Villain, M., Wetmore, D., Bairoch, A., Hochstrasser, D., 2004. Industrial-scale proteomics: from liters of plasma to chemically synthesized proteins. *Proteomics* 4, 2125–2150.
- Ross, J.S., Slodkowska, E.A., Symmans, W.F., Puztai, L., Ravdin, P.M., Hortobagyi, G.N., 2009. The HER-2 receptor and breast cancer: ten years of targeted anti-HER-2 therapy and personalized medicine. *Oncologist* 14, 320–368.
- Rui, Z., Jian-Guo, J., Yuan-Peng, T., Hai, P., Bing-Gen, R., 2003. Use of serological proteomic methods to find biomarkers associated with breast cancer. *Proteomics* 3, 433–439.
- Schenk, S., Schoenhals, G.J., de Souza, G., Mann, M., 2008. A high confidence, manually validated human blood plasma protein reference set. *BMC Med. Genomics* 1, 41.
- Schiess, R., Wollscheid, B., Aebersold, R., 2009. Targeted proteomic strategy for clinical biomarker discovery. *Mol. Oncol* 3, 33–44.
- Schmidt, A., Aebersold, R., 2006. High-accuracy proteome maps of human body fluids. *Genome Biol.* 7, 242.
- Shen, Y., Jacobs, J.M., Camp 2nd, D.G., Fang, R., Moore, R.J., Smith, R.D., Xiao, W., Davis, R.W., Tompkins, R.G., 2004. Ultra-high-efficiency strong cation exchange LC/RPLC/MS/MS for high dynamic range characterization of the human plasma proteome. *Anal. Chem.* 76, 1134–1344.
- Shevchenko, A., Wilm, M., Vorm, O., Mann, M., 1996. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* 68, 850–858.
- Simpson, R.J., Lim, J.W., Moritz, R.L., Mathivanan, S., 2009. Exosomes: proteomic insights and diagnostic potential. *Expert. Rev. Proteomics* 6, 267–283.
- Slamon, D.J., Clark, G.M., Wong, S.G., Levin, W.J., Ullrich, A., McGuire, W.L., 1987. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* 235, 177–182.
- Sørli, T., Perou, C.M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Thorsen, T., Quist, H., Matese, J.C., Brown, P.O., Botstein, D., Eystein Lønning, P., Børresen-Dale, A.L., 2001. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. USA* 98, 10869–10874.
- Souchelnytskyi, S., Lomnytska, M., Dubrovskaya, A., Hellman, U., Volodko, N., 2006. Proteomics success story. Towards early detection of breast and ovarian cancer: plasma proteomics as a tool to find novel markers. *Proteomics* 6 (Suppl. 2), 65–68.
- Stahl-Zeng, J., Lange, V., Ossola, R., Eckhardt, K., Krek, W., Aebersold, R., Domon, B., 2007. High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. *Mol. Cell. Proteomics* 6, 1809–1817.
- Stefani, M.M., Guerra, J.G., Sousa, A.L., Costa, M.B., Oliveira, M.L., Martelli, C.T., Scollard, D.M., 2009. Potential plasma markers of Type 1 and Type 2 leprosy reactions: a preliminary report. *BMC Infect. Dis.* 9, 75.
- Takahashi, T., Takasaki, Y., Takeuchi, K., Yamanaka, K., Oshimi, K., Hashimoto, H., 1997. Detection of proliferating cell nuclear antigen (PCNA) in peripheral blood mononuclear cells and sera of patients with malignant lymphoma. *Leuk. Lymphoma* 28, 113–125.
- Timins, J.K., 2005. Controversies in mammography. *N.J. Med.* 102, 45–49.
- Tlsty, T.D., Coussens, L.M., 2006. Tumor stroma and regulation of cancer development. *Annu. Rev. Pathol* 1, 119–150.
- Tracy, R.P., Currie, R.M., Young, D.S., 1982. Two-dimensional gel electrophoresis of serum specimens from a normal population. *Clin. Chem.* 28, 890–899.
- Uehara, M., Kinoshita, T., Hojo, T., Akashi-Tanaka, S., Iwamoto, E., Fukutomi, T., 2008. Long-term prognostic study of carcinoembryonic antigen (CEA) and carbohydrate antigen 15-3 (CA 15-3) in breast cancer. *Int. J. Clin. Oncol* 13, 447–451.
- Veenstra, T.D., Conrads, T.P., Hood, B.L., Avellino, A.M., Ellenbogen, R.G., Morrison, R.S., 2005. Biomarkers: mining the biofluid proteome. *Mol. Cell. Proteomics* 4, 409–418.
- Witz, I.P., 2009. The tumor microenvironment: the making of a paradigm. *Cancer Microenviron.* 2 (Suppl. 1), 9–17.
- Xue, H., Lu, B., Lai, M., 2008. The cancer secretome: a reservoir of biomarkers. *J. Transl. Med* 6, 52.
- Yom, C.K., Woo, H.Y., Min, S.Y., Kang, S.Y., Kim, H.S., 2009. Clusterin overexpression and relapse-free survival in breast cancer. *Anticancer Res.* 29, 3909–3912.
- Zeng, Z., Hincapie, M., Haab, B.B., Hanash, S., Pitteri, S.J., Kluck, S., Hogan, J.M., Kennedy, J., Hancock, W.S., 2009. The development of an integrated platform to identify breast cancer glycoproteome changes in human serum. *J. Chromatogr. A*. Epub ahead of print.
- Zhang, D., Tai, L.K., Wong, L.L., Chiu, L.L., Sethi, S.K., Koay, E.S., 2005. Proteomic study reveals that proteins involved in metabolic and detoxification pathways are highly expressed in HER-2/neu-positive breast cancer. *Mol. Cell. Proteomics* 4, 1686–1696.