Clinical proteomics in breast cancer: a review

Marie-Christine W. Gast, Jan H. M. Schellens, Jos H. Beijnen

Breast Cancer Res Treat (2009) 116:17–29
DOI 10.1007/s10549-008-0263-3

REVIEW

Clinical proteomics in breast cancer: a review

Marie-Christine W. Gast · Jan H. M. Schellens · Jos H. Beijnen

Received: 24 November 2008 / Accepted: 24 November 2008 / Published online: 11 December 2008

Abstract Breast cancer imposes a significant healthcare burden on women worldwide. Early detection is of paramount importance in reducing mortality, yet the diagnosis of breast cancer is hampered by the lack of an adequate detection method. In addition, better breast cancer prognostication may improve selection of patients eligible for adjuvant therapy. Hence, new markers for early diagnosis, accurate prognosis and prediction of response to treatment are warranted to improve breast cancer care. Since proteomics can bridge the gap between the genetic alterations underlying cancer and cellular physiology, much is expected from proteome analyses for the detection of better protein biomarkers. Recent technical advances in mass spectrometry, such as matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS) and its variant surface-enhanced laser desorption/ionisation (SELDI-) TOF MS, have enabled high-throughput proteome analysis. In the current review, we give a comprehensive overview of the results of expression proteomics (i.e. protein profiling) research performed in breast cancer using these two platforms. Many protein peaks have been reported to bear significant diagnostic, prognostic or predictive value, however, only few candidate markers have been structurally identified yet. In addition, although of pivotal importance in preventing overfitting of data and systematic bias by pre-analytical parameters, validation of biomarker candidates by other, quantitative, methods and/or in new populations is very limited. Moreover, none of the identified candidate biomarkers has been investigated for their utility as breast cancer markers in large, prospective, clinical settings. As such, the candidate biomarkers discussed in this overview have not been validated sufficiently to be used for clinical patient care. Nonetheless, regarding the promising results up to now, MALDI- and SELDI-TOF MS protein profiling studies could eventually fulfil the great promise that protein biomarkers have for improving cancer patient outcome, provided that these studies are performed with adequate statistical power and analytical rigour.

Keywords Breast cancer · Biomarkers · MALDI-TOF MS · SELDI-TOF MS

M.-C. W. Gast (✉) · J. H. Beijnen
Department of Pharmacy and Pharmacology, The Netherlands Cancer Institute/Slotervaart Hospital, P.O. Box 90440, 1006 BK Amsterdam, The Netherlands
e-mail: Marie-Christine.Gast@slz.nl

J. H. M. Schellens
Department of Medical Oncology, Antoni van Leeuwenhoek Hospital/The Netherlands Cancer Institute, P.O. Box 90203, 1006 BE Amsterdam, The Netherlands

J. H. M. Schellens · J. H. Beijnen
Faculty of Science, Department of Pharmaceutical Sciences, Division of Biomedical Analysis, Utrecht University, P.O. Box 80082, 3508 TB Utrecht, The Netherlands

Introduction

Breast cancer imposes a significant healthcare burden to women worldwide. For example, in the USA, breast cancer is estimated to be the most commonly diagnosed neoplasm in women in 2008, as it will account for 26% of all new female cancer cases [1]. In addition, preceded only by lung cancer, breast cancer is expected to be the second leading cause of USA cancer deaths in 2008 [1]. The 5-year survival rates of breast cancer decrease from 98% for localised disease to 26% for late stage disease [2]. Hence, short of prevention, detection of breast cancer at an early, still

curable stage would offer the best route to decrease its mortality rates. However, since only 63% of breast cancers are still confined to the breast at the time of diagnosis [1], the currently applied diagnostic screening tools (e.g. mammography) obviously do not suffice for adequate breast cancer diagnosis. In addition, despite the survival benefit achieved by locoregional treatment and adjuvant systemic therapy, 30–50% of breast cancer patients will eventually develop metastatic relapse and die [3], while a small percentage of patients would have survived without these treatment modalities. Evidently, the currently applied prognostic and predictive markers (e.g. age, hormone receptor status) lack adequate performance as well. Hence, better markers for early diagnosis, accurate prognosis and prediction of response to treatment are warranted to improve breast cancer care.

We now comprehend that cancer arises from successive genetic changes, by which a number of cellular processes, including growth control, senescence, apoptosis, angiogenesis, and metastasis, are altered [4, 5]. Consequently, researchers initially searched for markers by employing genomic and transcriptomic approaches, providing new biomarkers (e.g. [6–9]) and expanding our insight into the genetic basis of cancer. It is, however, currently understood that gene analysis by itself provides an incomplete picture. Due to alternative splicing of both mRNA and proteins, combined with more than 100 unique post-translational modifications, one gene can give rise to multiple protein species [10]. Hence, compared to the genome, the proteome can provide a more dynamic and accurate reflection of both the intrinsic genetic programme of the cell and the impact of its immediate environment [11]. Since proteome analysis can provide the link between gene sequence and cellular physiology [12], proteomics is expected to complement gene analyses for evaluating disease development, prognosis, and response to treatment [13].

Until recently, the search for novel protein biomarkers has been dominated by two-dimensional gel electrophoresis [14], a major disadvantage of which is its lack of real high-throughput capability. However, recent advances in analytical technologies, such as protein microarrays and mass spectrometry (MS), have enabled large-scale proteomic analyses [15]. Due to their relative simplicity of sample preparation, high analytical sensitivity and speed of data acquisition, two MS-based technologies in particular, i.e. matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) MS [16] and its variant surface-enhanced laser desorption/ionisation (SELDI-) TOF MS [17, 18] have been widely deployed for cancer biomarker discovery [19]. In both laser desorption/ionisation (LDI) platforms, biological samples (e.g. serum, tissue lysate) are co-crystallised with an energy absorbing matrix on a sample probe surface. Subsequent irradiation with brief laser pulses sublimates and ionises the proteins out of their crystalline matrix, after which an electric field migrates the charged proteins to the time-of-flight mass analyser. Herein, proteins are separated based on their mass, as the time to detector impact (TOF) is proportional to protein mass per charge (m/z). The two LDI platforms differ in their sample probe surfaces. In MALDI, the probe surface merely presents the sample to the mass spectrometer, warranting off-line sample fractionation and clean-up to produce usable MS signals. In contrast, the probe surfaces utilised by SELDI are comprised of various chromatographic surfaces, enabling their active role in sample fractionation (Fig. 1).

In the current overview, we focus on the expression proteomics (i.e. protein profiling) studies performed using the two LDI platforms in the search for novel breast cancer biomarkers. We will discuss the studies performed thus far for discovery of diagnostic, prognostic and predictive biomarkers, and evaluate the potential of the discriminating proteins identified in this research for clinical use as breast cancer biomarkers.

Diagnostic protein profiling studies

Short of prevention, detection at an early stage remains the best route to decrease breast cancer related mortality. Hence, the majority of MALDI/SELDI protein profiling studies performed in breast cancer have searched for novel diagnostic markers (Table 1). All diagnostic protein profiling studies were performed in vivo, investigating various biological matrices, including serum, plasma and tissue, but also nipple aspirate fluid, ductal lavage fluid and saliva.

Protein profiling of tissue

As tissue proteins will reflect the earliest changes caused by the successive genetic mutations that lead to breast cancer, it has been hypothesised that the concentration of potential biomarkers is highest in the tumour and its immediate microenvironment [19]. Although tissue provides an invaluable sample source, tissue sampling through biopsies is highly invasive, thereby limiting the number of diagnostic tissue protein profiling studies performed thus far.

Analysis of tumour tissue lysates by SELDI-TOF MS revealed several peaks that were significantly associated with lymph node status [20] or cancer subtype (i.e. lobular and ductal carcinoma) [21]. However, the search for tumour originating proteins can be complicated by the high cellular heterogeneity of whole tumour tissue specimens. This can be reduced by laser capture microdissection (LCM), enabling selective capture of a specific subset of cells [22]. Following microdissection, captured cells can be

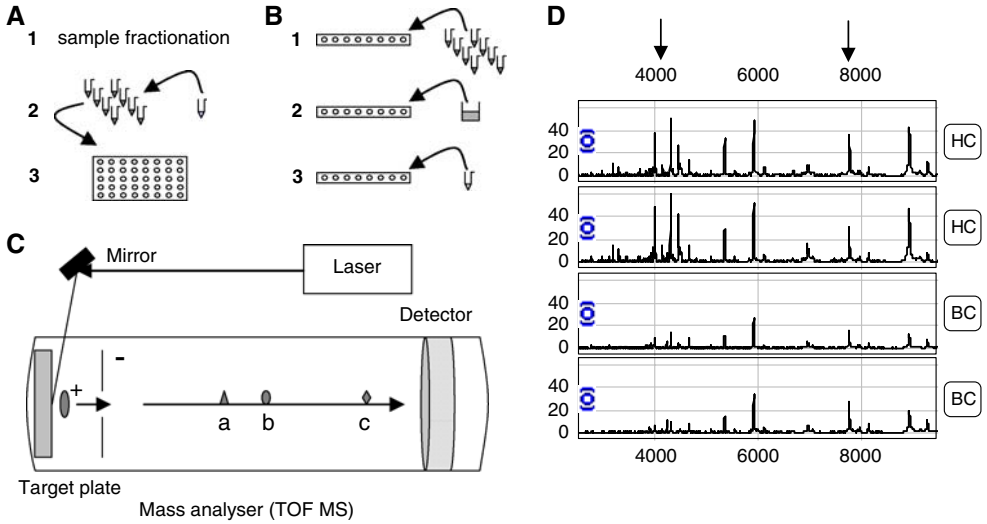

Fig. 1 Schematic representation of the MALDI- and SELDI-TOF MS principle (adapted from [15]). **a** Protein profiling by MALDI-TOF MS: 1. samples (μl volume) are fractionated off-line using for instance magnetic beads coated with a chromatographic surface (e.g. hydrophilic, hydrophobic, cationic, anionic, or immobilised metal affinity capture moiety), 2. addition of energy absorbing matrix (e.g. α-cyano-4-hydroxy-cinnamic acid) to (fractionated) samples, 3. application of mixed specimen to inert target plate for laser irradiation in (**c**). **b** Protein profiling by SELDI-TOF MS: 1. application of sample (μl volume) from, for example, cancer and control patients to an 8-spot array with a chromatographic surface (e.g. hydrophilic, hydrophobic, cationic, anionic, or immobilised metal affinity capture moiety) in appropriate binding buffer, 2. on-chip sample clean-up using various wash-buffers, 3. application of energy absorbing matrix (e.g. sinapinic acid) for desorption/ionisation of proteins by laser irradiation in a laser desorption/ionisation time-of-flight analyser (**c**). **c** Schematic representation of laser desorption/ionisation (LDI) time-of-flight (TOF) analyser: the MALDI target plate or SELDI array is inserted in the MALDI or SELDI instrument. Subsequent laser irradiation desorbs and ionises bound proteins, after which an electric field migrates the charged proteins to the TOF analyser. Herein, proteins are separated based on their mass, as the time to detector impact (TOF) is proportional to the protein mass per charge ($m/z = \text{constant} * t^2$). Thus, small proteins (c) fly faster than large ones (a), and multiple charged ones (b) faster than single-charged ones (a). **d** Representative example of SELDI-TOF mass spectra of sera from female healthy controls (HC) and breast cancer patients (BC). On the *x*-axis the protein m/z is displayed, and the *y*-axis depicts its abundance. Expression differences are visible between breast cancer and control sera at m/z 3980 and m/z 4292 (first arrow, ITIH4 fragments), and m/z 8939 (second arrow, $C3a_{desArg}$)

mounted directly on a MALDI target, thereby preserving their spatial conformation for imaging MS [23, 24]. Using LCM, Umar et al. [24] detected 9 differentially expressed tryptic peptides (not structurally identified) following analysis of stromal and tumour cells collected from five tissue specimens. In addition, Sanders et al. [23] identified ubiquitin and S100-A8 to be decreased in tumour ($n = 122$) compared to normal tissue ($n = 167$), whereas S100-A6 was found increased. Their split-sample approach allowed a successful within-study validation of the three potential markers [23]. As both ubiquitin and S100-A6 were also found to decrease in lysates of human breast cancer cell lines following chemotherapy induced apoptosis [25], these proteins may provide insight into the pathogenesis of breast cancer upon further investigation.

Despite the clear potential of (tumour) tissue to yield cancer-specific diagnostic biomarkers, their routine clinical application is seriously hampered by the intricacies associated with tissue sampling. Although this can be avoided by assessment of tumour-derived markers in easier accessible biological matrices such as serum, this type of validation has not been performed in breast cancer yet.

Protein profiling of serum and plasma

Since whole blood is considered to provide a dynamic reflection of physiological and pathological status, human plasma and serum represent the most extensively studied biological matrices in the quest for (breast) cancer biomarkers [26]. Constantly perfusing and percolating the human body, the blood compartment endows a protein-rich information archive [27]. Besides the expected circulatory proteins, this archive also contains specific tumour-secreted proteins, normal tissue- and plasma-proteins digested by tumour-secreted proteases, and proteins produced by local and distant responses to the tumour [11, 28, 29]. Moreover, whole blood is an easy to sample, readily accessible matrix that allows repeated collection, thereby augmenting the clinical relevance of candidate blood-borne biomarkers [28, 30].

Several MALDI-TOF MS and SELDI-TOF MS peaks (not structurally identified) have been reported to differentiate between serum or plasma of breast cancer patients, patients with benign breast disease and/or healthy controls [31–36]. Since a small percentage (7–10%) of breast

Table 1 Protein profiling studies performed in breast cancer by MALDI- and SELDI-TOF MS

Matrix	Training			Samples (*n*)			Validation	Samples (*n*)			ID	Ref.
	Platform	Pretreatment	Condition	BC	BD	HC	Platform	BC	BD	HC		
Diagnostic studies												
Tissue	MALDI	LCM, IMS	–	5	–	3[a]	n.p.				No	[24]
	MALDI	LCM, IMS	–	62	–	84	MALDI	60	0	83	Yes	[23]
	SELDI	LCM lysate	IMAC Cu	65	–	–	n.p.				No	[20]
	SELDI	Lysis	IMAC, WCX, SAX	20	–	–	n.p.				No	[21]
Serum	MALDI	C8 fractionation	–	78	–	29	n.p.				No	[32]
	MALDI	C18 fractionation	–	21	–	33	n.p.				Yes	[47]
	MALDI	WCX fractionation	–	46	46	–	n.p.				No	[36]
	MALDI	IMAC fractionation	–	48	–	28	n.p.				No	[31]
	MALDI	WCX fractionation	–	76	–	77	n.p.				No	[51]
	SELDI	Albumin depletion	H4	49	–	33	n.p.				No	[33]
	SELDI	–	IMAC Cu, SAX	45	42	47	n.p.				No	[39]
	SELDI	–	IMAC Cu, SAX	16	–	15	n.p.				No	[34]
	SELDI	–	IMAC Cu, SAX	–	–	–	SELDI	47	0	48	No	[38]
	SELDI	–	IMAC Cu	15	–	15	n.p.				No	[37]
	SELDI	–	IMAC Ni	155	–	155	SELDI				No	[40]
	SELDI	–	IMAC Ni	103	25	41	n.p.				No	[41]
	SELDI	–	IMAC Ni	–	–	–	SELDI	93	37	46	Yes	[42]
	SELDI	–	IMAC Ni	–	–	–	SELDI	49	13	27	No	[43]
	SELDI	–	IMAC Ni	–	–	–	SELDI	48	0	48	No	[44]
	SELDI	–	Immunoassay	20	–	41	n.p.				Yes	[48]
	SELDI	–	Immunoassay	19	–	40	n.p.				Yes	[46]
Plasma	SELDI	SAX fractionation	IMAC Cu	61	–	61	SELDI	28	0	0	Yes	[50]
	SELDI	–	IMAC Cu	29	–	15	n.p.				No	[35]
NAF	SELDI	–	H4	12	–	15	n.p.				No	[65]
	SELDI	–	NP20, H4, SAX	20	–	13	n.p.				Yes	[66]
	SELDI	–	NP20, H4, SAX	25	83	–	n.p.				No	[67]
	SELDI	–	IMAC Cu, WCX	23	–	23[b]	n.p.				No	[68]
	SELDI	–	IMAC Cu, WCX	23	–	23[b],5	n.p.				No	[70]
	SELDI	–	IMAC Cu	5	–	5	SELDI	9	0	7[b]	Yes	[64]
	SELDI	–	NP20	38	–	63	n.p.				No	[62]
	SELDI	–	IMAC Cu, WCX	21	–	21[b],44	n.p.				No	[69]
DLF	SELDI	–	SAX	16	–	16[b]	n.p.				No	[63]
Saliva	SELDI	–	WCX	3	–	3	n.p.				No	[71]
Prognostic studies												
Cell line	SELDI	Lysis	IMAC Cu, WCX	27	–	–	IHC on TMA	547[a]	0	0	Yes	[74]
Tissue	SELDI	Lysis	IMAC Cu	105	–	–	n.p.				Yes	[80]
	SELDI	Lysis	IMAC Cu	60	–	–	n.p.				Yes	[73]
Serum	SELDI	SAX fractionation	IMAC Cu, SAX	81	–	–	n.p.				Yes	[45]
	SELDI		SAX	63	–	–	1D GE	371	0	0	Yes	[72]
CSF	MALDI	Tryptic digestion	–	87	–	–	n.p.				Yes	[75, 76]
Predictive studies												
Cell line	SELDI	Lysis	H4	3	–	–	n.p.				No	[82]
	SELDI	(Medium)	IMAC Cu	2	–	–	n.p.				Yes	[83]
	SELDI	Lysis	WCX	3	–	–	n.p.				Yes	[25]
Serum	SELDI		WCX	6	–	–	n.p.				Yes	[84]

Table 1 continued

Matrix	Training			Samples (*n*)			Validation	Samples (*n*)			ID	Ref.
	Platform	Pretreatment	Condition	BC	BD	HC	Platform	BC	BD	HC		
Plasma	SELDI		IMAC Cu	24	–	–	n.p.				No	[35]

Abbreviations. 1D GE: 1 dimensional gel electrophoresis, BC: breast cancer, BD: benign breast disease patient, CSF: cerebrospinal fluid, DLF: ductal lavage fluid, H4: reversed phase array, HC: healthy control, ID: structural identification of candidate biomarkers, IHC: immunohistochemistry, IMAC: immobilised metal affinity capture (fractionation or array), IMS: imaging mass spectrometry, LCM: laser capture microdissection, NAF: nipple aspirate fluid, n.p.: not performed, NP20: normal phase array, SAX: strong anion exchange (fractionation or array), TMA: tissue microarray, WCX: weak cation exchange (fractionation or array)

[a] Tissue sample obtained from tissue adjacent to tumourous tissue, [b] NAF/DLF sample obtained from non-cancerous contralateral breast

cancers is attributable to hereditary syndromes (e.g. BRCA-1, -2 mutations), Becker et al. [37] investigated whether the BRCA-1 mutation was reflected by the serum proteome. Multiple SELDI-TOF MS peaks were significantly different in expression between breast cancer patients with and without the BRCA-1 mutation [37]. However, as none of these peaks were structurally identified, their association to the BRCA-1 gene remains unclear. Moreover, none of the peaks reported by these studies have been validated by analysis of an independent sample set. Yet validation is of utmost importance to ascertain reproducibility and prevent systematic bias and overfitting of data. This is highlighted by a study of our group [38], in which the potential markers for breast cancer and lymph node status, reported by Vlahou et al. [39] and Laronga et al. [34], respectively, could not be confirmed following analysis of an independent sample set. In contrast, Belluco et al. [40] report excellent performance of their seven-peak classifier (not structurally identified) following validation by an independent sample set analysed 14 months after their initial discovery study.

Li et al. [41] observed three serum peaks to distinguish patients from controls: one (4.3 kDa) decreased and two (8.1 and 8.9 kDa) increased in patients. These peaks were structurally identified as a fragment of inter-alpha-trypsin inhibitor heavy chain H4 (ITIH4, 4.3 kDa), C3a des-arginine ($C3a_{desArg}$, 8.9 kDa) and a C-terminal truncated form thereof ($C3a_{desArg\Delta8}$, 8.1 kDa) [42]. Subsequent analysis of an independent sample set could only confirm the increased 8.1 and 8.9 kDa C3a fragments [42]. However, the 8.1 kDa $C3a_{desArg\Delta8}$ was found to lack significance in a second [43] and third validation study [44]. The latter study also reported a decreased 8.9 kDa $C3a_{desArg}$ expression in breast cancer [44], whereas in all previous studies, this fragment was found increased [41–43]. Beyond these four studies, $C3a_{desArg}$ has been found associated with survival, as its expression decreased in metastatic relapse [45]. In addition, the 4.3 kDa ITIH4 fragment was one of the several ITIH4 fragments found increased in breast cancer by Song et al. [46]. Similar ITIH4 fragments, observed by Villanueva et al. [47] and Fung et al. [48], were found either increased in cancer [47], or devoid of discriminative power [48]. Regarding the inconsistent regulation observed across multiple studies, the definitive value of the different ITIH4 fragments, $C3a_{desArg}$, and $C3a_{desArg\Delta8}$ in the diagnosis of breast cancer cannot be determined yet.

In addition to the various ITIH4 fragments, several fragments of fibrinopeptide A, fibrinogen alpha, C3f, C4a, apolipoprotein A-IV, bradykinin, factor XIII, and transthyretin were found to provide accurate class discrimination [47]. Generated by exoprotease activities superimposed on the ex vivo coagulation and complement-degradation pathways, these fragments are proposed to bear cancer-type specificity. It has, however, been argued that this "peptidome signature" merely reflects the hypercoagulable state of the blood of cancer patients [49] and not necessarily a cancer-specific signature [19]. Although the peptidome signature has not been validated yet, two fragments thereof (i.e. the ITIH4 fragment discussed above, and a fibrinogen fragment) have been encountered in other studies as well [41, 46, 48]. The fibrinogen fragment, though increased in the breast cancer serum peptidome, was found decreased in breast cancer plasma and reverted to normal values after surgical extirpation of the tumour [50]. The difference between study results most likely originates from the biological matrix investigated, as plasma differs from serum by inhibition of the coagulation cascade, by which fibrinogen is generated.

Also of interest are the results of the "Classification competition on clinical mass spectrometry proteomic diagnosis data" [51]. For this competition, sera of breast cancer patients ($n = 76$) and healthy controls ($n = 77$) were analysed by MALDI-TOF MS. Data were subsequently analysed by ten competition participants for construction of diagnostic classifiers [52–61]. Surprisingly, though the various bioinformatic methods applied resulted in highly divergent classification models, reported performances (ranging from 83% to 89%) were very similar. However, as these results are based on a single dataset, validation by analysis of an independent study population most likely will reveal differences between the various bioinformatic methods and their resulting classification models.

Serum and plasma protein profiling studies by MALDI-TOF or SELDI-TOF MS have yielded numerous protein peaks with a significantly different expression between breast cancer and healthy control. However, although elucidation of protein identities is essential for insights into the molecular mechanisms involved in breast cancer, thus far, only a small percentage of reported peaks has been structurally identified. Moreover, since most studies did not investigate other cancer types or patients with benign breast disease, the specificity of reported markers for breast cancer still has to be addressed. Furthermore, although of pivotal importance, only few potential markers have been validated by analysis of independent sample sets. As these studies generally yielded contradictory results, further research is needed to determine the potential of identified markers in breast cancer diagnosis.

Protein profiling of nipple aspirate fluid and ductal lavage fluid

Most breast cancers (70–80%) are thought to arise from the epithelial cells lining the mammary ducts [13]. The breast epithelium exfoliates cells as a renewal of tissue and secretes fluid into the ductal-lobular system of the breast. While this fluid exits each breast through six to nine orifices at the nipple, it can be collected by either of two non-invasive methods; aspiration or ductal lavage. Both nipple aspirate fluid (NAF) and ductal lavage fluid (DLF) are traditionally used for cytological assessment [62, 63], but their vicinity to the breast epithelium renders them valuable matrices for diagnostic protein profiling studies as well.

Although several discriminating protein peaks were detected when comparing equal volumes of NAF or DLF from breast cancer patients and healthy controls by SELDI-TOF MS [62–65], large variations in the spectra between different samples within one diagnostic group were observed [63–65]. Likely originating from the wide protein content range of NAF (1–90 mg/ml [62]), further studies have normalised the protein content prior to analysis. Nonetheless, despite normalisation, Sauter et al. [66, 67] could not confirm the initially observed diagnostic potential of three SELDI-TOF MS peaks (identified as hemoglobin beta chain isoforms) in a second, larger, study population. In contrast, despite the very limited sample size, the differential expression of human neutrophil peptides 1–3 observed in NAF ($n = 10$) was confirmed by analyses of pooled DLF samples from cancerous ($n = 9$) and unrelated healthy ($n = 7$) breasts [64].

As the breasts are a paired organ system, NAF samples from both the cancerous and non-cancerous breast of patients with unilateral breast cancer have been compared as well. Surprisingly, although different between patients, protein expression patterns were highly similar in both breasts of each patient [68–70]. Comparison of either the cancerous or the contralateral breast to unrelated healthy controls, however, yielded several significantly different peaks [69, 70].

Despite limited sample sizes and lack of validation studies, NAF protein profiling did distinguish between women with and without breast cancer. However, as identification of the cancer-bearing breast was not possible, protein profiling of NAF may have more value in breast cancer risk assessment and disease monitoring than as a diagnostic tool [69]. Evidently, further research is needed to assess the value of the intraductal approach in breast cancer diagnosis.

Protein profiling of saliva

The use of saliva in diagnosis of systemic diseases such as breast cancer has been demonstrated by the detection of increased levels of solubilised c-erbB-2 and CA15.3 in breast cancer patients compared to healthy controls [33]. Investigating saliva for diagnostic purposes has several key advantages, including its noninvasive collection, the possibility of repeated sampling, and the ease of sample handling and processing. Nonetheless, thus far, only one feasibility study has been performed in saliva. Using SELDI-TOF MS, five high molecular weight peaks were found to be overexpressed in breast cancer ($n = 3$) compared to control ($n = 3$) [71]. Although these peaks were neither structurally identified nor validated in larger sample sets, this study does show the potential of using saliva for diagnostic purposes.

Prognostic protein profiling studies

Compared to diagnostic studies, protein profiling studies aimed at discovering novel markers to improve breast cancer prognostication are rather limited (Table 1). Investigating post-operative sera of 83 high-risk breast cancer patients by SELDI-TOF MS, Goncalves et al. [45] constructed a 40-protein signature that correctly predicted outcome in 83% of patients. Major components of this signature included haptoglobin alpha-1, complement component C3a, transferrin, and apolipoprotein A-I and C-I (Table 2). These results should be interpreted cautiously, as the number of proteins used for classification is rather high in comparison with the limited study population, indicating probable over-fitting of the data. Moreover, results have not been validated in independent sample sets. The importance of validation is emphasised by a study performed by our group [72]. Using SELDI-TOF MS, we discovered a strong association between haptoglobin phenotype and recurrence free survival in sera of 63 high-risk

Table 2 Candidate biomarkers in breast cancer identified by MALDI- and SELDI-TOF MS protein profiling studies

Biomarker identity	Platform (m/z)		Matrix	Expression		Function	Ref.
	MALDI	SELDI		±	in		
α-1-antichymotrypsin	t.p.		CSF	+	Metastasis	Acute phase protein, serine protease inhibitor	[75, 76]
Annexin V		33327	Tissue	+	Luminal subtype	Tumour proliferation/ metastasis?, anticoagulant protein	[80]
Apolipoprotein A-I		28284	Serum	−	Relapse	Lipid metabolism	[45]
	t.p.		CSF	+	Metastasis		[75, 76]
Apolipoprotein A-II		9285	Serum	−	Shock	Lipid metabolism	[84]
Apolipoprotein A-IV	2508		Serum	+	Cancer	Lipid metabolism	[47]
Apolipoprotein C-I		6647	Serum	−	Relapse	Acute phase protein, lipid metabolism	[45]
Apolipoprotein E	t.p.		CSF	−	Metastasis	Lipid metabolism	[75, 76]
Bradykinin (fragments)		904, 1061	Serum	+	Cancer	Inflammation mediator	[47]
C3a C-terminal fragment ($C3a_{desArg\Delta 8}$)		8100	Serum	+	Cancer	Complement activation	[41]
		8116	Serum	+	Cancer		[42]
		8129	Serum	n.s.	Cancer		[43]
		8129	Serum	n.s.	Cancer		[44]
C3a des-R anaphylatoxin ($C3a_{desArg}$)		8900	Serum	+	Cancer	Complement activation	[41]
		8926	Serum	+	Cancer		[42]
		8919	Serum	+	Cancer		[43]
		8941	Serum	−	Cancer		[44]
		8936	Serum	−	Relapse		[45]
C3f (fragments)	942–1865		Serum	−	Cancer	Complement activation	[47]
C4a (fragments)	1627–2704		Serum	+	Cancer	Complement activation	[47]
Factor XIII	2602		Serum	+	Cancer	Blood coagulation	[47]
Ferritin light chain		19809	Tissue	+	Relapse	Acute phase protein, iron homeostasis	[73]
Fibrinogen alpha (fragments)	2379, 2659		Serum	+	Cancer	Blood coagulation	[47]
		2661	Plasma	−	Cancer		[50]
Fibrinopeptide A (fragments)	905–1537		Serum	±*	Cancer	Blood coagulation	[47]
Haptoglobin (alpha 1)	t.p.		CSF	+	Metastasis	Acute phase protein, haemoglobin binding	[75, 76]
		9192	Serum	+	Relapse		[45]
		9192	Serum	−	Relapse		[72]
Heat shock protein 27		27152	Tissue	+	Luminal subtype	Stress resistance, actin organisation	[80]
Hemoglobin beta chain (isoforms)		15940	NAF	+	Cancer	Oxygen transport	[66]
Hemopexin	t.p.		CSF	+	Metastasis	Haeme binding and transport, acute phase protein	[75, 76]
Human neutrophil peptide 1–3		3375–3490	NAF	+	Cancer	Antibiotic, fungicide and antiviral	[64]
ITIH4		4300	Serum	−	Cancer	Acute phase reactant?	[41]
		4300	Serum	+	Cancer		[42]
		4286	Serum	−	Cancer		[43]
		4276	Serum	−	Cancer		[44]
		2271–3272	Serum	n.s.	Cancer		[48]
		2271–4293	Serum	+	Cancer		[46]
	998–2358		Serum	+	Cancer		[47]

Table 2 continued

Biomarker identity	Platform (m/z)		Matrix	Expression		Function	Ref.
	MALDI	SELDI		±	in		
Kininogen HMW		7790	Serum	−	Shock	Blood coagulation, bradykinin release	[84]
Prostaglandin D2 synthase	t.p.		CSF	−	Metastasis	Catalyses prostaglandin conversion	[75, 76]
S100-A6 (isoforms)		10900	Cell line	−	Apoptosis	Ca^{2+}-binding protein, growth factor	[25]
	10094		Tissue	−	Cancer		[23]
S100-A8	10842		Tissue	+	Cancer	Ca^{2+}-binding protein, inflammation (dimer with S100-A9)	[23]
S100-A9		13300	Cell line	−	Basal-like subtype	Ca^{2+}-binding protein, inflammation (dimer with S100-A8)	[74]
Transferrin (human)		81763	Serum	+	Relapse	Acute phase reactant, iron binding and transport, cell proliferation	[45]
	t.p.		CSF	+	Metastasis		[75, 76]
Transferrin (bovine)		7600	Medium	+	Resistance	Iron binding and transport, cell proliferation	[83]
Transthyretin (fragment)	2451		Serum	+	Cancer	Thyroid hormone-binding protein, acute phase reactant	[47]
	t.p.		CSF	−	Metastasis		[75, 76]
Ubiquitin		8445	Cell line	+	Basal-like subtype	Protein modifier	[74]
		8507	Tissue	−	Metastasis		[73]
		8560	Cell line	−	Apoptosis		[25]
	8568		Tissue	+	Cancer		[23]

Abbreviations. C: complement component, ITIH4: inter-α-trypsin inhibitor heavy chain 4, m/z: mass-to-charge ratio, detected by MALDI- or SELDI-TOF MS, t.p.: tryptic digest peptides

* ± expression: one fragment was found increased (m/z 905), while the other fragments were found decreased (m/z 1264, 1351, and 1537) in breast cancer vs. control sera

primary breast cancer patients. However, as results were not confirmed following validation by haptoglobin phenotyping of a six-fold larger sample set ($n = 371$), this observation most likely resulted from a type I error (i.e. false positive) [72].

In a third SELDI-TOF MS study, performed in breast cancer tissue ($n = 60$), high levels of ubiquitin and/or low levels of ferritin light chain were found associated with a good prognosis [73]. Although results have not been confirmed by analysis of independent sample sets, ubiquitin has also been found differently expressed in breast cancer by three other studies investigating tissue specimens [23] and cell lines [25, 74].

Lastly, cerebrospinal fluid (CSF) has also been explored for prognostic markers [75, 76]. CSF is specific for the central nervous system [77], contains less total protein than serum and provides a low fluid-volume-to-organ ratio, thereby augmenting biomarker discovery [30]. As collection of CSF by invasive lumbar puncture is not applicable to healthy controls, this matrix has thus far only been investigated for prognostic purposes. In search for markers indicative of leptomeningeal metastases (LM), whole CSF samples of 106 breast cancer patients were digested with trypsin [75]. Following MALDI-TOF MS analysis of the resulting peptides, a classifier with 77% accuracy in determining LM status was constructed [75]. The discriminative tryptic peptides were derived of several proteins [76]. Three of these proteins (i.e. apolipoprotein A-I, haptoglobin and transferrin) have also been found associated to clinical outcome in serum [45].

Currently, breast cancer prognosis is assessed by a.o. TNM classification, assigning breast tumours to different stages based on depth of tumour invasion and presence of metastases. However, considering the heterogeneity in outcome of patients diagnosed with equivalent TNM stage, this classification system is suboptimal in tumour characterisation. Instead, tumour staging on the molecular level could be more accurate. Indeed, microarray-based gene expression profiling studies have identified five major molecular breast cancer subtypes (i.e. luminal A and B,

ERBB2-overexpressing, basal-like, and normal-like), showing distinct clinical courses and responses to therapeutic agents [78, 79]. Hence, in search for prognostic markers, two studies have investigated the correlation between SELDI-TOF MS protein profiles of tumour tissue lysates ($n = 105$) [80] and breast cancer cell lines ($n = 27$) [74] with the previously reported breast cancer subtypes. Although discrepancies between cells grown in vivo and in vitro exist due to adaptation to cell culture conditions, breast cancer cell lines have been shown to accurately reflect the genomic, transcriptional, and biological heterogeneity found in primary tumours [81]. As such, they appear to be a good surrogate matrix for tumour tissues, enabling proteome comparisons without introducing interfering factors. Indeed, in both studies, patient subgroups identified by hierarchical clustering of SELDI-TOF MS protein profiles were analogous to the molecular breast cancer subtypes [74, 80]. Of the several differentially expressed protein peaks detected, heat shock protein (HSP) 27 and annexin V were identified as over-expressed in the luminal A type tumour tissue lysates [80], while S100-A9 and a C-terminal truncated form of ubiquitin were found differentially expressed between the luminal-like and basal-like cell lines [74]. Of note, subsequent immunohistochemical analysis of S100-A9 in tumour specimens of 547 early breast cancer patients confirmed its association with basal subtypes, as well as its value as an indicator of poor prognosis [74]. The in vivo prognostic potential of HSP 27 and annexin V should be assessed by validation in clinical samples.

Similar to the diagnostic studies, the prognostic studies published thus far generally investigated only a limited number of samples. Combined with the large number of features generated by the resulting protein profiles, datasets are frequently subjected to multiple testing. Hence, candidate biomarkers are prone to be false positive, rendering validation of pivotal importance to assess their true clinical performance. Nonetheless, thus far, only two validation studies have been performed. All studies have, however, structurally identified (part of) the candidate prognostic markers. The markers identified across serum and CSF (e.g. apolipoprotein A-I, haptoglobin and transferrin) were highly abundant, non-specific, host-response generated proteins. In addition, many of the proteins identified in tissue and cell lines (e.g. annexin V, S100-A9) are in fact normal cellular proteins. However, as their precise role in breast cancer remain to be elucidated, further research is needed to determine their value for breast cancer prognostication.

Predictive protein profiling studies

Although accurate prediction of chemosensitivity in cancer therapy would enable individualised therapy, thus avoiding toxic side effects and eliminating the use of ineffective agents, protein profiling studies searching for markers for response prediction and treatment monitoring of breast cancer are scarce. Several SELDI-TOF MS peaks (not structurally identified) were found indicative of treatment regimen for chemosensitive and -resistant breast cancer cell lines following exposure to doxorubicin or paclitaxel [82]. In addition, Dowling et al. [83] found an increase of a 7.6 kDa bovine transferrin fragment in serum-free conditioned medium of paclitaxel resistant human breast cancer cell lines, corresponding to the increased expression of the transferrin receptor they observed in whole cell lysates. Although these results were not translated to a human in vivo setting, other studies have indeed reported an association between increased serum and CSF transferrin levels and poor clinical outcome [45, 75, 76]. Similarly, while ubiquitin and S100-A6 were found to decrease in lysates of human breast cancer cell lines following chemotherapy induced apoptosis [25], an aberrant expression of both proteins has also been reported in breast cancer tissue [23, 73]. Nonetheless, regarding the very limited number of samples investigated in the various studies, screening of larger cohorts and validation of the preclinical data in clinical samples is warranted before these potential markers can be used to improve therapeutic accuracy in clinical practice.

In vivo studies have been performed as well [35, 84]. In serum, both high molecular weight kininogen and apolipoprotein A-II were found significantly decreased in expression following docetaxel-induced shock [84]. Likewise, in plasma, a SELDI-TOF MS peak at m/z 2790 (not structurally identified) was found to significantly increase following (neo)adjuvant paclitaxel infusion [35]. As it remains to be elucidated whether identified proteins are treatment-responsive, originate from micrometastatic carcinoma, or merely result from a general host-response to cytotoxic therapy, the definitive value of identified proteins as predictive markers can not be established yet.

Discussion and conclusion

Thus far, the majority of LDI protein profiling studies performed in breast cancer have searched for novel diagnostic markers, while the search for new prognostic and predictive biomarkers is limited to only few studies. The studies discussed in the current overview together have reported hundreds of mass-to-charge values, intensities of which were found to contain significant diagnostic, prognostic or predictive value. However, although indispensable for providing insight into the pathophysiological mechanisms associated with, or underlying, breast cancer, and development of absolute quantitative assays, only very few

of these mass-to-charge ratios have been structurally identified yet. Moreover, the candidate markers that have been identified constitute of normal cellular proteins and high abundant blood proteins involved in coagulation and the acute phase response. Since their biology cannot be linked directly to tumour biochemistry, one of the ultimate aims of (LDI) protein profiling studies, i.e. increasing knowledge of the molecular mechanisms involved in cancer by identification of discriminative (full-length) proteins generated exclusively by cancer cells, has not been fulfilled yet.

Moreover, many of the identified candidate breast cancer markers have been found to bear diagnostic potential for other cancer types as well (e.g. $C3a_{desArg}$ in colorectal cancer [85], apolipoprotein A-I in ovarian cancer [86]), indicating a general lack of tumour-specificity. However, as cancer cells are deranged host cells, and most cancers of epithelial origin share similar molecular features [77], it may prove difficult to find a true cancer-specific protein that is expressed exclusively by one type of malignant cells. On the other hand, as such proteins are expected to be among the least abundant proteins, they could well be below the detection limit of the current (LDI) methods. Hence, these specific tumour-secreted proteins might actually exist, but could simply have eluded detection thus far.

Nonetheless, identification of specific tumour-secreted proteins is no prerequisite for improving breast cancer care, as better breast cancer diagnosis, prognosis, and prediction can also be accomplished by surrogate biomarkers of disease. A class of proteins currently recognised for their surrogate biomarker potential is the (proteolytic fragments of) high-abundant circulatory proteins. These fragments are hypothesised to be generated by cancer type-specific exoprotease activity, superimposed on the ex vivo coagulation and complement degradation proteolytic pathways. In addition, these fragments can also result from the proteases specifically expressed by malignant cells within the tumour microenvironment for tumour invasion and metastasis [87, 88], as they proteolytically process the acute phase proteins that are generated by the host response to the tumour. Since these modified host response proteins generally are present at substantially higher circulatory concentrations than the enzymes that process them upon their exposure to the tumour microenvironment, they can be detected in blood by current (LDI) methods for diagnostic purposes [48]. Although in breast cancer, this concept has been investigated for a.o. serum ITIH4, the various studies have reported contradictory results, a finding not entirely unforeseen regarding the biological matrix commonly investigated (i.e. serum). Since serum is generated by coagulation, its proteome is prone to the proteases involved in this cascade, as well as to those involved in the complement cascade, activated upon clotting. Various pre-analytical parameters, such as sampling device, clotting temperature, and storage time, can thus all exert a distinct influence on the serum proteome. Hence, the concept of cancer type-specific (host response) protein fragments generated by tumour-secreted proteases still awaits confirmation by validation studies that adhere to rigorous sample handling protocols.

The need for such validation studies is, however, not limited to the reported host response protein fragments. Regardless of their identity, the majority of markers has been reported by single breast cancer studies, in which only limited numbers of samples were investigated, thereby compromising the generalisibility of results. Moreover, as the number of generated features (i.e. protein MS peaks) usually by far exceeds the number of samples investigated, proteomic (LDI) datasets are frequently subjected to multiple testing. As such, many candidate biomarkers are prone to be false positive. Hence, to prevent overfitting of data, as well as systematic bias by above-mentioned pre-analytical parameters, validation of biomarker candidates by other, quantitative, methods and/or in new study populations is of pivotal importance. Yet, thus far, such validation studies have been performed for only few of the candidate biomarkers detected in LDI studies (i.e. serum $C3a_{desArg}$, $C3a_{desArg\Delta 8}$, ITIH4 fragments, haptoglobin alpha-1, plasma m/z 2660 fibrinogen, and tissue S100-A9). As these studies generally yielded contradictory results (except for the m/z 2660 fibrinogen fragment and S100-A9), further research is needed to determine the true value of these markers in breast cancer management.

The few validation studies performed thus far are all of retrospective nature. In fact, none of the identified candidate markers has been investigated for their utility as breast cancer biomarkers in a larger, prospective, clinical setting. As such, none of the candidate biomarkers discussed in this overview has been validated sufficiently to be used for clinical patient care. Yet, the move from discovery phase to the pre-clinical and subsequent clinical validation phase is mandatory, as the sole purpose of a biomarker lies in its application. Nonetheless, overseeing the results of MALDI- and SELDI-TOF MS protein profiling studies up to now, the two platforms hold promise as high-throughput screening tools for discovery of novel breast cancer markers. Provided that these studies are performed with adequate statistical power and analytical rigour, they could eventually fulfil the great promise that protein biomarkers have for improving cancer patient outcome.

References

1. Jemal A, Siegel R, Ward E et al (2008) Cancer statistics, 2008. CA Cancer J Clin 58:71–96. doi:10.3322/CA.2007.0010

2. Ries LA, Melbert D, Krapcho M et al (2008) SEER Cancer Statistics Review, 1975–2004. http://seer.cancer.gov/csr/1975_2004/
3. Early Breast Cancer Trialists' Collaborative Group (1998) Polychemotherapy for early breast cancer: an overview of the randomised trials. Lancet 352:930–942
4. Mommers EC, van Diest PJ, Leonhart AM et al (1999) Balance of cell proliferation and apoptosis in breast carcinogenesis. Breast Cancer Res Treat 58:163–169. doi:10.1023/A:1006396103777
5. Reis-Filho JS, Lakhani SR (2003) The diagnosis and management of pre-invasive breast disease: genetic alterations in pre-invasive lesions. Breast Cancer Res 5:313–319. doi:10.1186/bcr650
6. 't Veer LJ, Dai H, van de Vijver MJ et al (2002) Gene expression profiling predicts clinical outcome of breast cancer. Nature 415:530–536. doi:10.1038/415530a
7. Wang Y, Klijn JG, Zhang Y et al (2005) Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 365:671–679
8. Miki Y, Swensen J, Shattuck-Eidens D et al (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 266:66–71
9. Vogelstein B, Kinzler KW (1994) Has the breast cancer gene been found? Cell 79:1–3. doi:10.1016/0092-8674(94)90393-X
10. Banks RE, Dunn MJ, Hochstrasser DF et al (2000) Proteomics: new perspectives, new biomedical opportunities. Lancet 356:1749–1756. doi:10.1016/S0140-6736(00)03214-1
11. Aebersold R, Anderson L, Caprioli R et al (2005) Perspective: a program to improve protein biomarker discovery for cancer. J Proteome Res 4:1104–1109. doi:10.1021/pr050027n
12. Dove A (1999) Proteomics: translating genomics into products? Nat Biotechnol 17:233–236. doi:10.1038/6972
13. Clarke W, Zhang Z, Chan DW (2003) The application of clinical proteomics to cancer and other diseases. Clin Chem Lab Med 41:1562–1570. doi:10.1515/CCLM.2003.239
14. Hondermarck H, Vercoutter-Edouart AS, Revillion F et al (2001) Proteomics of breast cancer for marker discovery and signal pathway profiling. Proteomics 1:1216–1232. doi:10.1002/1615-9861(200110)1:10≤1216::AID-PROT1216≥3.0.CO;2-P
15. Engwegen JY, Gast MC, Schellens JH et al (2006) Clinical proteomics: searching for better tumour markers with SELDI-TOF mass spectrometry. Trends Pharmacol Sci 27:251–259. doi:10.1016/j.tips.2006.03.003
16. Hortin GL (2006) The MALDI-TOF mass spectrometric view of the plasma proteome and peptidome. Clin Chem 52:1223–1237. doi:10.1373/clinchem.2006.069252
17. Hutchens TW, Yip TT (1993) New desorption strategies for the mass spectrometric analysis of macromolecules. Rapid Commun Mass Spectrom 7:576–580. doi:10.1002/rcm.1290070703
18. Merchant M, Weinberger SR (2000) Recent advancements in surface-enhanced laser desorption/ionization-time of flight-mass spectrometry. Electrophoresis 21:1164–1177. doi:10.1002/(SICI)1522-2683(20000401)21:6≤1164::AID-ELPS1164≥3.0.CO;2-0
19. Simpson RJ, Bernhard OK, Greening DW et al (2008) Proteomics-driven cancer biomarker discovery: looking to the future. Curr Opin Chem Biol 12:72–77. doi:10.1016/j.cbpa.2008.02.010
20. Nakagawa T, Huang SK, Martinez SR et al (2006) Proteomic profiling of primary breast cancer predicts axillary lymph node metastasis. Cancer Res 66:11825–11830. doi:10.1158/0008-5472.CAN-06-2337
21. Traub F, Feist H, Kreipe HH et al (2005) SELDI-MS-based expression profiling of ductal invasive and lobular invasive human breast carcinomas. Pathol Res Pract 201:763–770. doi:10.1016/j.prp.2005.08.011
22. Cowherd SM, Espina VA, Petricoin EFIII et al (2004) Proteomic analysis of human breast cancer tissue with laser-capture microdissection and reverse-phase protein microarrays. Clin Breast Cancer 5:385–392. doi:10.3816/CBC.2004.n.046
23. Sanders ME, Dias EC, Xu BJ et al (2008) Differentiating proteomic biomarkers in breast cancer by laser capture microdissection and MALDI MS. J Proteome Res 7:1500–1507. doi:10.1021/pr7008109
24. Umar A, Dalebout JC, Timmermans AM et al (2005) Method optimisation for peptide profiling of microdissected breast carcinoma tissue by matrix-assisted laser desorption/ionisation-time of flight and matrix-assisted laser desorption/ionisation-time of flight/time of flight-mass spectrometry. Proteomics 5:2680–2688. doi:10.1002/pmic.200400128
25. Leong S, Christopherson RI, Baxter RC (2007) Profiling of apoptotic changes in human breast cancer cells using SELDI-TOF mass spectrometry. Cell Physiol Biochem 20:579–590. doi:10.1159/000107541
26. Hanash SM, Pitteri SJ, Faca VM (2008) Mining the plasma proteome for cancer biomarkers. Nature 452:571–579. doi:10.1038/nature06916
27. Petricoin EF, Zoon KC, Kohn EC et al (2002) Clinical proteomics: translating benchside promise into bedside reality. Nat Rev Drug Discov 1:683–695. doi:10.1038/nrd891
28. Conrads TP, Zhou M, Petricoin EFIII et al (2003) Cancer diagnosis using proteomic patterns. Expert Rev Mol Diagn 3:411–420. doi:10.1586/14737159.3.4.411
29. Grizzle WE, Semmes OJ, Bigbee WL et al (2005) The need from review and understanding of SELDI/MALDI mass spectroscopy data prior to analysis. Cancer Inform 1:86–97
30. Good DM, Thongboonkerd V, Novak J et al (2007) Body fluid proteomics for biomarker discovery: lessons from the past hold the key to success in the future. J Proteome Res 6:4549–4555. doi:10.1021/pr070529w
31. Callesen AK, Vach W, Jorgensen PE et al (2008) Combined experimental and statistical strategy for mass spectrometry based serum protein profiling for diagnosis of breast cancer: a case-control study. J Proteome Res 7:1419–1426. doi:10.1021/pr7007576
32. de Noo ME, Deelder A, van der Werff WM et al (2006) MALDI-TOF serum protein profiling for the detection of breast cancer. Onkologie 29:501–506. doi:10.1159/000095933
33. Hu Y, Zhang S, Yu J et al (2005) SELDI-TOF-MS: the proteomics and bioinformatics approaches in the diagnosis of breast cancer. Breast 14:250–255. doi:10.1016/j.breast.2005.01.008
34. Laronga C, Becker S, Watson P et al (2003) SELDI-TOF serum profiling for prognostic and diagnostic classification of breast cancers. Dis Markers 19:229–238
35. Pusztai L, Gregory BW, Baggerly KA et al (2004) Pharmacoproteomic analysis of prechemotherapy and postchemotherapy plasma samples from patients receiving neoadjuvant or adjuvant chemotherapy for breast carcinoma. Cancer 100:1814–1822. doi:10.1002/cncr.20203
36. Shin S, Cazares L, Schneider H et al (2007) Serum biomarkers to differentiate benign and malignant mammographic lesions. J Am Coll Surg 204:1065–1071. doi:10.1016/j.jamcollsurg.2007.01.036
37. Becker S, Cazares LH, Watson P et al (2004) Surfaced-enhanced laser desorption/ionization time-of-flight (SELDI-TOF) differentiation of serum protein profiles of BRCA-1 and sporadic breast cancer. Ann Surg Oncol 11:907–914. doi:10.1245/ASO.2004.03.557
38. Gast MC, Bonfrer JM, van Dulken EJ et al (2006) SELDI-TOF MS serum protein profiles in breast cancer: assessment of robustness and validity. Cancer Biomark 2:235–248
39. Vlahou A, Laronga C, Wilson L et al (2003) A novel approach toward development of a rapid blood test for breast cancer. Clin Breast Cancer 4:203–209. doi:10.3816/CBC.2003.n.026
40. Belluco C, Petricoin EF, Mammano E et al (2007) Serum proteomic analysis identifies a highly sensitive and specific discriminatory pattern in stage 1 breast cancer. Ann Surg Oncol 14:2470–2476

41. Li J, Zhang Z, Rosenzweig J et al (2002) Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. Clin Chem 48:1296–1304
42. Li J, Orlandi R, White CN et al (2005) Independent validation of candidate breast cancer serum biomarkers identified by mass spectrometry. Clin Chem (Kyoto) 51:2229–2235
43. Mathelin C, Cromer A, Wendling C et al (2006) Serum biomarkers for detection of breast cancers: a prospective study. Breast Cancer Res Treat 96:83–90
44. van Winden AWJ, Gast MCW, Beijnen JH et al (2008) Validation of previously identified serum biomarkers for breast cancer with SELDI-TOF MS: a case control study. BMC Med Genomics (accepted for publication)
45. Goncalves A, Esterni B, Bertucci F et al (2006) Postoperative serum proteomic profiles may predict metastatic relapse in high-risk primary breast cancer patients receiving adjuvant chemotherapy. Oncogene 25:981–989. doi:10.1038/sj.onc.1209131
46. Song J, Patel M, Rosenzweig CN et al (2006) Quantification of fragments of human serum inter-alpha-trypsin inhibitor heavy chain 4 by a surface-enhanced laser desorption/ionization-based immunoassay. Clin Chem 52:1045–1053. doi:10.1373/clinchem.2005.065722
47. Villanueva J, Shaffer DR, Philip J et al (2006) Differential exoprotease activities confer tumor-specific serum peptidome patterns. J Clin Invest 116:271–284. doi:10.1172/JCI26022
48. Fung ET, Yip TT, Lomas L et al (2005) Classification of cancer types by measuring variants of host response proteins using SELDI serum assays. Int J Cancer 115:783–789. doi:10.1002/ijc.20928
49. Goldsmith GH (2001) Hemostatic changes in patients with malignancy. Int J Hematol 73:151–156. doi:10.1007/BF02981931
50. Shi Q, Harris LN, Lu X et al (2006) Declining plasma fibrinogen alpha fragment identifies HER2-positive breast cancer patients and reverts to normal levels after surgery. J Proteome Res 5:2947–2955. doi:10.1021/pr060099u
51. van der Werff MP, Mertens B, de Noo ME et al (2008) Case-control breast cancer study of MALDI-TOF proteomic mass spectrometry data on serum samples. Stat Appl Genet Mol Biol 7:Article 2
52. Barrett JH, Cairns DA (2008) Application of the random forest classification method to peaks detected from mass spectrometric proteomic profiles of cancer patients and controls. Stat Appl Genet Mol Biol 7:Article 4
53. Heidema AG, Nagelkerke N (2008) Developing a discrimination rule between breast cancer patients and controls using proteomics mass spectrometric data: a three-step approach. Stat Appl Genet Mol Biol 7:Article 5
54. Fearn T (2008) Principal component discriminant analysis. Stat Appl Genet Mol Biol 7:Article 6
55. Datta S (2008) Classification of breast cancer versus normal samples from mass spectrometry profiles using linear discriminant analysis of important features selected by random forest. Stat Appl Genet Mol Biol 7:Article 7
56. Hoefsloot HC, Smit S, Smilde AK (2008) A classification model for the Leiden proteomics competition. Stat Appl Genet Mol Biol 7:Article 8
57. Strimenopoulou F, Brown PJ (2008) Empirical Bayes logistic regression. Stat Appl Genet Mol Biol 7:Article 9
58. Goeman JJ (2008) Autocorrelated logistic ridge regression for prediction based on proteomics spectra. Stat Appl Genet Mol Biol 7:Article 10
59. Pham TV, van de Wiel MA, Jimenez CR (2008) Support vector machine approach to separate control and breast cancer serum samples. Stat Appl Genet Mol Biol 7:Article 11
60. Valkenborg D, Van Sanden S, Lin D et al (2008) A cross-validation study to select a classification procedure for clinical diagnosis based on proteomic mass spectrometry. Stat Appl Genet Mol Biol 7:Article 12
61. Gammerman A, Nouretdinov I, Burford B et al (2008) Clinical mass spectrometry proteomic diagnosis by conformal predictors. Stat Appl Genet Mol Biol 7:Article 13
62. He J, Gornbein J, Shen D et al (2007) Detection of breast cancer biomarkers in nipple aspirate fluid by SELDI-TOF and their identification by combined liquid chromatography-tandem mass spectrometry. Int J Oncol 30:145–154
63. Mendrinos S, Nolen JD, Styblo T et al (2005) Cytologic findings and protein expression profiles associated with ductal carcinoma of the breast in ductal lavage specimens using surface-enhanced laser desorption and ionization-time of flight mass spectrometry. Cancer 105:178–183. doi:10.1002/cncr.21052
64. Li J, Zhao J, Yu X et al (2005) Identification of biomarkers for breast cancer in nipple aspiration and ductal lavage fluid. Clin Cancer Res 11:8312–8320. doi:10.1158/1078-0432.CCR-05-1538
65. Paweletz CP, Trock B, Pennanen M et al (2001) Proteomic patterns of nipple aspirate fluids obtained by SELDI-TOF: potential for new biomarkers to aid in the diagnosis of breast cancer. Dis Markers 17:301–307
66. Sauter ER, Zhu W, Fan XJ et al (2002) Proteomic analysis of nipple aspirate fluid to detect biologic markers of breast cancer. Br J Cancer 86:1440–1443. doi:10.1038/sj.bjc.6600285
67. Sauter ER, Shan S, Hewett JE et al (2005) Proteomic analysis of nipple aspirate fluid using SELDI-TOF-MS. Int J Cancer 114:791–796. doi:10.1002/ijc.20742
68. Kuerer HM, Coombes KR, Chen JN et al (2004) Association between ductal fluid proteomic expression profiles and the presence of lymph node metastases in women with breast cancer. Surgery 136:1061–1069. doi:10.1016/j.surg.2004.04.011
69. Noble JL, Dua RS, Coulton GR et al (2007) A comparative proteinomic analysis of nipple aspiration fluid from healthy women and women with breast cancer. Eur J Cancer 43:2315–2320. doi:10.1016/j.ejca.2007.08.009
70. Pawlik TM, Fritsche H, Coombes KR et al (2005) Significant differences in nipple aspirate fluid protein expression between healthy women and those with breast cancer demonstrated by time-of-flight mass spectrometry. Breast Cancer Res Treat 89:149–157. doi:10.1007/s10549-004-1710-4
71. Streckfus CF, Bigler LR, Zwick M (2006) The use of surface-enhanced laser desorption/ionization time-of-flight mass spectrometry to detect putative breast cancer markers in saliva: a feasibility study. J Oral Pathol Med 35:292–300. doi:10.1111/j.1600-0714.2006.00427.x
72. Gast MC, van Tinteren H, Bontenbal M et al (2008) Haptoglobin phenotype is not a predictor of recurrence free survival in high-risk primary breast cancer patients. BMC Cancer (accepted for publication)
73. Ricolleau G, Charbonnel C, Lode L et al (2006) Surface-enhanced laser desorption/ionization time of flight mass spectrometry protein profiling identifies ubiquitin and ferritin light chain as prognostic biomarkers in node-negative breast cancer tumors. Proteomics 6:1963–1975. doi:10.1002/pmic.200500283
74. Goncalves A, Charafe-Jauffret E, Bertucci F et al (2008) Protein profiling of human breast tumor cells identifies novel biomarkers associated with molecular subtypes. Mol Cell Proteomics 7:1420–1433. doi:10.1074/mcp.M700487-MCP200
75. Dekker LJ, Boogerd W, Stockhammer G et al (2005) MALDI-TOF mass spectrometry analysis of cerebrospinal fluid tryptic peptide profiles to diagnose leptomeningeal metastases in patients with breast cancer. Mol Cell Proteomics 4:1341–1349. doi:10.1074/mcp.M500081-MCP200
76. Rompp A, Dekker L, Taban I et al (2007) Identification of leptomeningeal metastasis-related proteins in cerebrospinal fluid of

patients with breast cancer by a combination of MALDI-TOF, MALDI-FTICR and nanoLC-FTICR MS. Proteomics 7:474–481
77. Hu S, Loo JA, Wong DT (2006) Human body fluid proteome analysis. Proteomics 6:6326–6353. doi:10.1002/pmic.200600284
78. Sorlie T, Perou CM, Tibshirani R et al (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA 98:10869–10874. doi:10.1073/pnas.191367098
79. Perou CM, Sorlie T, Eisen MB et al (2000) Molecular portraits of human breast tumours. Nature 406:747–752. doi:10.1038/35021093
80. Brozkova K, Budinska E, Bouchal P et al (2008) Surface-enhanced laser desorption/ionization time-of-flight proteomic profiling of breast carcinomas identifies clinicopathologically relevant groups of patients similar to previously defined clusters from cDNA expression. Breast Cancer Res 10:R48
81. Neve RM, Chin K, Fridlyand J et al (2006) A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell 10:515–527. doi:10.1016/j.ccr.2006.10.008
82. Mian S, Ball G, Hornbuckle J et al (2003) A prototype methodology combining surface-enhanced laser desorption/ionization protein chip technology and artificial neural network algorithms to predict the chemoresponsiveness of breast cancer cell lines exposed to Paclitaxel and Doxorubicin under in vitro conditions. Proteomics 3:1725–1737. doi:10.1002/pmic.200300526
83. Dowling P, Maurya P, Meleady P et al (2007) Purification and identification of a 7.6-kDa protein in media conditioned by superinvasive cancer cells. Anticancer Res 27:1309–1317
84. Heike Y, Hosokawa M, Osumi S et al (2005) Identification of serum proteins related to adverse effects induced by docetaxel infusion from protein expression profiles of serum using SELDI ProteinChip system. Anticancer Res 25:1197–1203
85. Habermann JK, Roblick UJ, Luke BT et al (2006) Increased serum levels of complement C3a anaphylatoxin indicate the presence of colorectal tumors. Gastroenterology 131:1020–1029. doi:10.1053/j.gastro.2006.07.011
86. Zhang Z, Bast RC Jr, Yu Y et al (2004) Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. Cancer Res 64:5882–5890. doi:10.1158/0008-5472.CAN-04-0746
87. Bank U, Kruger S, Langner J et al (2000) Review: peptidases and peptidase inhibitors in the pathogenesis of diseases. Disturbances in the ubiquitin-mediated proteolytic system. Protease–antiprotease imbalance in inflammatory reactions. Role of cathepsins in tumour progression. Adv Exp Med Biol 477:349–378. doi:10.1007/0-306-46826-3_38
88. Blasi F (1999) Proteolysis, cell adhesion, chemotaxis, and invasiveness are regulated by the u-PA-u-PAR-PAI-1 system. Thromb Haemost 82:298–304