Breast cancer classification by proteomic technologies: current state of knowledge

SW Lam
CR Jimenez
E Boven

Cancer Treat Rev. 2014 Feb;40(1):129-38
Abstract

Breast cancer is traditionally considered as a heterogeneous disease. Molecular profiling of breast cancer by gene expression studies has provided us an important tool to discriminate a number of subtypes. These breast cancer subtypes have been shown to be associated with clinical outcome and treatment response. In order to elucidate the functional consequences of altered gene expressions related to each breast cancer subtype, proteomic technologies can provide further insight by identifying quantitative differences at the protein level. In recent years, proteomic technologies have matured to an extent that they can provide proteome-wide expressions in different clinical materials. This technology can be applied for the identification of proteins or protein profiles to further refine breast cancer subtypes or for discovery of novel protein biomarkers pointing towards metastatic potential or therapy resistance in a specific subtype. In this review, we summarize the current state of knowledge of proteomic research on molecular breast cancer classification and discuss important aspects of the potential usefulness of proteomics for discovery of breast cancer-associated protein biomarkers in the clinic.
Introduction

Breast cancer affects more than 1.3 million women worldwide each year and accounts for about 14% of cancer-related deaths. The incidence of breast cancer has increased over the past decades and is expected to rise substantially in the coming years. Hence, breast cancer will remain a considerable health burden.

Work on breast cancer has revealed substantial tumor heterogeneity consisting of different molecular subtypes, each with distinct biological and clinical characteristics. In the pivotal study by Perou et al., it has been shown that differential gene expression patterns account for heterogeneity among breast carcinomas. Based on the so-called intrinsic gene signatures, four major breast cancer subtypes were initially classified: luminal, HER2-enriched, basal-like and normal breast-like subtype. Subsequent studies by Perou et al. and others have expanded these initial findings by providing additional information for further refinements and adjustments of the breast cancer classification. Within the luminal subtype characterized by the expression of luminal epithelial markers, three groups are currently recognized: luminal A, luminal B/HER2-negative and luminal B/HER2-positive. Basal-like breast cancer is a heterogeneous group of tumors comprising different histologies, which express basal epithelial markers. The normal breast-like subtype was located in a cluster containing normal breast and benign tumor samples and showed overexpression of genes related to adipose tissue and non-epithelial cell types in the original and subsequent validation studies. This subtype may also be a technical artifact due to low tumor cellularity. Hence, the normal breast-like subtype was often overlooked and was consequently poorly characterized.

The classification of breast cancer based on gene expression patterns has resulted into attempts to characterize clinically meaningful subgroups showing correlation with survival, disease relapse, site of preference of metastatic spread and chemotherapy response. Since microarray techniques are rather expensive and not readily available, immunohistochemistry (IHC) is an important method to define surrogate protein biomarkers for the classification of breast cancer. The main advantages of IHC are its lower costs and easy implementation into standard pathology workflow. It has been shown that the molecular classification by microarray analysis corresponds reasonably well to IHC classification of different breast carcinomas. Consequently, molecular and IHC classifications are concomitantly used to define the breast cancer subtypes (Table 1).

Ongoing research will identify new subtypes within the designated breast cancer classification. Complementary to the genomic-based approach, proteomics might provide new insights into aberrant processes among breast cancer subtypes and may identify additional proteins or protein profiles to refine current breast cancer classifications. Moreover, proteomics might reveal biological insights and identify protein biomarkers defining differen-
ces in therapy resistance, prognosis and metastatic spread within a specific subtype. The purpose of this review is to discuss the current state of knowledge of proteomic studies conducted in relation with the molecular classification of breast cancer.

A brief overview of proteomic technologies

Definition of proteomics
Proteomics is a term which refers to a large-scale study of proteins encompassing several aspects, such as protein identification, protein ontology, protein-protein interaction, pathways involvement, quantification and functional analysis. In addition, proteomics involves the identification of protein subgroups, such as kinases (kinome), secreted proteins (secretome), phosphorylated proteins (phosphoproteomics), exosomal proteins (exosome) and proteases (degradome). A multitude of complex biological samples can be analyzed using proteomic technologies, such as tissue, serum, plasma, saliva, nipple fluid, urine, cerebrospinal fluid and so on, which makes proteomics an attractive strategy for biomarker discovery. This section aims to provide a brief overview of commonly used proteomic approaches based on the utilization of mass spectrometry (MS) as well as main methods for validation of protein candidates.

MS-based proteomics for biomarker discovery
Mass spectrometers have increasingly been employed as a platform for discovery proteomics or targeted follow-up of proteins in complex biological samples representing different disease conditions. In recent years, remarkable progress towards near complete proteome coverage and high sample throughput has been made by technological and methodological advancements, which is reviewed elsewhere. Additionally, reproducible results can be achieved in an optimized workflow. It requires, however, a large initial investment to acquire a mass spectrometer, specialized staff and highly sophisticated bioinformatics to obtain reliable results.

The experimental design of earlier proteomic discovery studies are those of two-dimensional gel electrophoresis (2D-PAGE) coupled with MS. Proteins are separated in two dimensions based on mass and charge differences and can be visualized with e.g. Coomassie Brilliant Blue or silver staining. The resulting protein spots on a gel can be picked up for protein identification by MS instruments. Although 2D-PAGE has been widely used in proteomic studies, it suffers from inherent problems including the requirement of a large amount of material, gel-to-gel variation, limited dynamic range, low-throughput identification of proteins and bias towards abundant proteins. Two-dimensional difference gel electrophoresis (2D-DIGE) is designed for analysis of multiple samples on one gel after labeling with different fluorescent dyes, while it simplifies the analysis and reduces gel-to-gel variations.
Table 1. Breast carcinoma subtypes: histopathological, molecular and clinical features.

<table>
<thead>
<tr>
<th>Molecular subtype</th>
<th>Prevalence</th>
<th>IHC definition</th>
<th>Additional markers</th>
<th>Genes</th>
<th>Histological grade</th>
<th>TP53 mutation</th>
<th>Prognosis</th>
<th>Consensus recommendation for (Neo)adjuvant systemic treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luminal A</td>
<td>50–60%</td>
<td>ER+ and/or PgR+ HER2-Ki-67 low</td>
<td>CK 8/18 + FOXA1+</td>
<td>ESR1, GATA3, KRT8, KRT18, XBP1, FOXA1, TFF3, CCND1, LIV1</td>
<td>Good differentiation</td>
<td>Low</td>
<td>Good</td>
<td>Endocrine therapy alone**</td>
</tr>
<tr>
<td>Luminal B</td>
<td>10–30%</td>
<td>-</td>
<td>FGFR1 and ZIC3 amplification</td>
<td>ESR1, GATA3, KRT8, KRT18, XBP1, FOXA1, TFF3, SQLE, LAPTMB</td>
<td>Moderate differentiation</td>
<td>Intermediate</td>
<td>Intermediate</td>
<td>-</td>
</tr>
<tr>
<td>Luminal B (HER2 negative)</td>
<td>15–20%</td>
<td>ER+ and/or PgR+ HER2-Ki-67 high</td>
<td>-</td>
<td>Not examined*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Endocrine therapy ± chemotherapy***</td>
</tr>
<tr>
<td>Luminal B (HER2 positive)</td>
<td>6%</td>
<td>ER+ and/or PgR+ Any Ki-67 HER2+</td>
<td>-</td>
<td>Not examined*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Endocrine + cytotoxic + anti-HER2 therapy</td>
</tr>
<tr>
<td>HER2-enriched</td>
<td>10–15%</td>
<td>HER2+ ER- and PgR-</td>
<td>CK 5/6+ GRB7+</td>
<td>ERBB2, GRB7</td>
<td>Poor differentiation</td>
<td>High</td>
<td>Poor</td>
<td>Chemotherapy + anti-HER2 therapy***</td>
</tr>
<tr>
<td>Basal-like</td>
<td>10–20%</td>
<td>ER- and PgR-HER2-</td>
<td>EGFR+, CK 5/6+, CK 14+, CK 17+, HER1+, Cyclin E+, CDKN2A+, RB1: low/–, BRCA1: low/–, FGFR2: amplification</td>
<td>KRT5, CDH3, ID4, FABP7, KRT17, TRIM29, LAMC2, ITGB4</td>
<td>Poor differentiation</td>
<td>High</td>
<td>Poor</td>
<td>Chemotherapy for triple negative breast cancer (ductal)</td>
</tr>
</tbody>
</table>

*Prevalence data as reported in 7,9,12,25
† According to St. Gallen International Expert Consensus 2011 71
‡ Not examined in the original discovery studies by Perou et al. 5 and Sørlie et al. 3
** Some high-risk patients e.g. high nodal status require chemotherapy
*** Inclusion and type of chemotherapeutic agents may depend on level of endocrine receptor expression, perceived risk and patient preference
**** Patients at very low risk e.g. pT1aN0 may be observed without systemic adjuvant treatment

ER, estrogen receptor; IHC, immunohistochemistry; PgR, progesterone receptor
A typical ‘modern’ MS-based proteomic workflow encompasses multidimensional protein separation by gel electrophoresis and nano-liquid chromatography prior to identification and quantification by MS. The selection of appropriate separation steps is crucial, because the number of identified proteins differs substantially. Different MS instruments, including a variety of techniques, can be used for protein identification and quantification and have recently been reviewed. In general, two established MS-based strategies have been widely adopted: stable isotope labels or label-free quantitative proteomics.

The isotope labeling technologies are based on the principle of isotope-induced shifts in mass, which can be detected in the same MS experiment. This allows simultaneous quantification of proteins among disease conditions. The use of isotope labels delivers significantly improved accuracy of protein quantification, but it comes at a cost in terms of expensive isotope labels, the requirement of specialized software and statistical challenges. In the label-free MS approach, all samples are processed and analyzed in parallel allowing the flexibility to conduct multiple comparisons. The total number of identified peptides corresponding to a certain protein (spectral count) or peptide ion intensity are used as a measurement for the relative quantity of a particular protein.

Methods for validation or verification of proteins
Verification of proteomic findings with an alternative assay provides further information about data validity, while validation of proteins in a large patient cohort supports potential clinical application as a biomarker. Commonly used methods suitable for these purposes utilize an antibody-based approach and include enzyme-linked immunosorbent assay (ELISA), Western blot and IHC. IHC combined with tissue microarrays (TMAs) allows broad-scaled analysis of predefined proteins, but requires scoring by experienced pathologists due to the risk of interobserver and intraobserver variability. These methods require prior selection of proteins based on a certain hypothesis or previous experiment and are, therefore, less suitable for biomarker discovery. Sensitive and well-validated antibodies are often not available for novel proteins. Apart from the drawbacks, antibody-based approach is relatively inexpensive, requires less-specialized laboratory facilities.

A novel antibody-based technology is protein microarrays, which allows simultaneous analysis of multiple proteins using capture molecules spotted on a surface. There are different types of protein microarrays based on the type of capture molecules and the question to be answered (reviewed in more detail: ). Antibody arrays are coated with immobilized antibodies to measure the abundance of many specific proteins in complex mixtures and are suitable for biomarker discovery in limited quantities of clinical samples and without fractionation. In reverse-phase protein arrays, the analyte rather than the antibody is immobilized on individual spots. Each analyte spot represents an independent array and can be incubated with different antibodies hereby providing information about the functional state of signaling pathways. Although protein microarray technology is developing rapidly, it still requires high quality antibodies. Furthermore, custom-made arrays are rarely availa-
Targeted MS-based proteomics utilizing selected/multiple reaction monitoring (SRM/MRM) provides an exciting potential to simultaneously detect multiple proteins in complex biological matrices, such as plasma and tissue. Unlike MS-based discovery proteomics, which aims to detect all proteins in an unbiased manner, SRM/MRM is fine-tuned to measure many preselected peptides (up to 100) with greater sensitivity, accuracy and quantification ability. Although challenges still have to be overcome, SRM/MRM is emerging to be an important approach in clinical proteomics for detection of novel proteins for which antibodies are not available. For advantages and limitations of SRM as well as experimental workflow reference is made to recent reviews.23,24.

Tissue proteomics to identify proteins related to breast cancer subtypes

Since tumor tissues are routinely obtained as surgical specimens or biopsies, they represent an attractive source for biomarker discovery to identify proteins that potentially improve current clinical management. Proteins can be extracted from a minimum amount of tissue material and their actual abundance between conditions can be determined by different proteomic methods. In the next section, we will summarize results from tissue-based proteomic studies in breast cancer subtypes.

Luminal A and B subtypes

Luminal A subtype represents 50–60% of breast carcinomas being the most common breast cancer subtype (Table 1).9,12,25. The profile includes overexpression of genes related to cells on the luminal side of mammary ducts, estrogen receptor (ER) and ER transcription factors.4,5 The histological and IHC characteristics include low histological grade, low Ki-67 index, positive staining for ER, progesterone receptor (PgR) and luminal epithelial markers cytokeratin (CK) 8 and 18. This subtype has the lowest rate of TP53 mutation as compared with other subtypes.5,12. Furthermore, luminal A subtype has the best prognosis with a 5-year survival rate of approximately 90% for patients presenting with primary breast cancer.13 At diagnosis, this subtype is also associated with less lymph node involvement,13 lower local recurrent25 and distant recurrent rates9.

The prevalence of the luminal B subtype ranges from 10–30% (Table 1).12,25. Luminal B subtype shares some histological similarities with luminal A subtype showing positive ER and/or PgR, but it is associated with moderate histological grade and higher Ki-67 index.26 Based on cases with simultaneous HER2 overexpression, luminal B carcinomas are now divided into luminal B/HER2-negative and luminal B/HER2-positive subgroups. On the gene expression level, luminal B subtype contains low expression of ER and ER-related genes, high expression of proliferation- and cell cycle-related genes as compared with luminal A.
The luminal B/HER2-positive subtype contains HER2 overexpression. Patients with primary breast cancer of the luminal B subtype carry a slightly less favorable prognosis than those with luminal A subtype with a reported 5-year survival rate of 88%\(^{13,26}\), while it is still better than that of non-luminal subtypes. The local recurrent and distant recurrent rates are lower than those of non-luminal subtypes, but are slightly higher than reported for luminal A subtype.\(^{9,25}\)

A distinction between luminal A and B subtypes has not been made in the vast majority of proteomic studies. Since hormone receptor status is a key determinant of luminal subtypes, the following section will summarize proteomic data derived from hormone receptor-positive breast cancer.

**Differential expression of proteins in hormone receptor-positive breast cancer versus normal breast tissue**

Most newly diagnosed breast carcinomas are hormone receptor-positive (ER and/or PgR) and can be classified as the luminal subtypes. Consequently, proteins specifically related to these hormone receptor pathways might be useful to distinguish malignant tissue from normal tissue and might be candidates for early breast cancer detection. Tissue-based proteomic analysis is an effective strategy to further prioritize candidate proteins.

Weitzel et al.\(^ {27}\) have analyzed 18 ER-positive/HER2-negative breast carcinomas and adjacent benign tissue identifying 40 upregulated proteins in breast cancer. Three proteins (PPIaseB, Rho-GDI α and TPM4) were verified by Western blotting (Table 2). The upregulation of TPM4 in ER-positive breast cancer has also been described in another proteomic analysis comparing ER-positive and ER-negative breast carcinomas.\(^ {28}\) In contrast, in the latter study, the expression of Rho-GDI α was found to be downregulated in ER-positive breast cancer.

The discordant protein expressions among studies may be explained by differences in experimental workflow, but more importantly by tumor heterogeneity and contamination with adherent benign tissue. Consequently, some investigators have employed laser capture microdissection to partly overcome these problems and to enrich for tumor cells. Cha et al.\(^ {29}\) have employed proteomic analysis on laser microdissected tumor cells from nine ER-positive breast carcinomas and nine normal breast tissues. In total, 1,623 proteins were identified, in which 298 proteins were significantly more abundant in ER-positive breast cancer. These proteins were related to biological processes including focal adhesion, lipid metabolism and extracellular matrix receptor interaction. Although no proteins were validated, these results provide further insight into aberrant processes on the protein level in hormone receptor-positive breast cancer as compared with normal tissue.
Differential expression of proteins in hormone receptor-positive versus hormone receptor-negative breast cancer

Hormone receptor status has traditionally been used as a prognostic factor and to select patients likely to have benefit from endocrine therapy. Patients with hormone receptor-positive breast cancer generally have a better prognosis than those with hormone receptor-negative breast cancer. This knowledge has prompted proteomic investigations to better understand the molecular differences between ER-positive and ER-negative breast cancer. Additionally, proteins related to hormone receptor-positive breast cancer may represent potential therapeutic targets.

Traub et al. 30 have identified thymosin α1 as the most overexpressed protein in 39 ER-positive compared with 41 ER-negative breast carcinoma samples. Thymosin α1 has been shown to exhibit antitumor and immunomodulating effects in preclinical models 31. Using pooled samples of laser microdissected tumor cells from eight ER-positive and eight ER-negative breast carcinomas, Neubauer et al. 32 have described that PGRMC1 was more abundant in ER-negative breast tumors, whereas PGRMC1 was more phosphorylated in ER-positive breast tumors. The localization of PGRMC1 overexpression identified by subsequence immunofluorescence appeared to be restricted to ER-negative cells located in the hypoxic zone surrounding necrotic tumor tissue. Further preclinical findings have shown that PGRMC1 has a promoting role in tumor growth partly through binding to EGFR and triggers apoptosis after inhibition suggesting therapeutic potential 33. In the study of Rezaul et al. 34, comparative analysis between laser microdissected tumor cells from three ER-positive and three ER-negative breast carcinomas has revealed a total of 2,995 proteins. Among 236 differential proteins, selected proteins found to be upregulated (liprin-α1, β-arrestin-1) and downregulated (fascin, DAP5) in ER-positive breast cancer were validated by IHC using TMAs of 33 patient samples.

A change in glycosylation of a glycoprotein often causes a change in function and may be associated with cancerous transformation. Therefore, Semaan et al. 35 have compared glycosylated proteins between ER-positive and ER-negative breast carcinomas. Among the selected glycoproteins with differential glycosylation were α1-acid glycoprotein, α1-antitrypsin, calmodulin, and superoxide dismutase mitochondrial precursor. Verification of these glycoproteins by Western blot showed differences in glycosylation of α1-antitrypsin in malignant tissues. An increase in the molecular weight of native α1-antitrypsin was highest in ER-negative tissues. While the native form of α1-antitrypsin is thought to act as a tumor suppressor, glycosylation might point towards a change in its function 35.

The secretome refers to proteins released by tumor cell lines into the conditioned culture medium and provides a valuable source of potentially secreted, cancer-specific biomarkers 36. Pavlou et al. 37 have characterized the secretomes of eight cell lines representing the luminal, basal-like and HER2-enriched subtypes and identified, respectively, four, 23 and four subtype-specific proteins. The validation was, however, primarily focused on four
ER-specific proteins. Higher levels of ABAT and PDZK1 measured in secretomes of luminal (ER-positive) cell lines were also detected in cytosolic fractions of ER-positive breast carcinomas as compared with ER-negative counterparts. The high level of ABAT in ER-positive breast carcinomas was confirmed in an independent set of 20 breast cancer samples using targeted LC-MS/MS. In a further experiment, tissue mRNA expression and survival data from four existing datasets were used and genes were matched with the proteins identified by proteomic analysis. ABAT showed consistently higher expression in patients without recurrence in each of the four datasets suggesting its prognostic potential.

**Differential expression of proteins related to treatment resistance in hormone receptor-positive breast cancer**

Patients with hormone receptor-positive breast cancer have benefit from endocrine therapy, such as tamoxifen, although some tumors do not respond. The ability to distinguish between responsive and unresponsive tumors to tamoxifen by using protein biomarkers would enable the selection of patients eligible for alternative drugs. To identify such proteins, an extensive comparative analysis has been performed of laser microdissected tumor cells pooled from 24 tamoxifen-sensitive and 27 tamoxifen-resistant breast carcinomas. Samples had been taken from primary tumors of patients that had not received adjuvant hormonal therapy. They were treated with tamoxifen as first-line therapy upon detection of advanced disease. Of the 100 differentially abundant proteins identified, the expression of ENPP1, EIF3E and GNB4 was verified in individual tumor samples by independent targeted LC-MS/MS experiments and was found to be closely related to tamoxifen response. Tumors with absence of ENPP1, high abundance of EIF3E, high presence of GNB4 were all associated with benefit from tamoxifen. The availability of an antibody against EMMPRIN enabled IHC on TMA from 130 breast cancer patients that had received tamoxifen, which showed that positive EMMPRIN staining was correlated with a shorter time to progression.

Neoadjuvant chemotherapy is an approach in use for primary breast cancer to reduce the tumor burden before surgery. Similarly to endocrine therapy, it is not possible to predict response to chemotherapy. Hence, Hodgkinson et al. have employed two different proteomic approaches on ER-positive breast carcinomas aiming to identify proteins associated with chemotherapy resistance. In the first study, 2D-PAGE and MALDI-TOF MS were used yielding 57 differentially expressed proteins. Canonical signaling pathway analysis revealed that the top proteins were involved in cell cycle processes including several isoforms of 14-3-3 family proteins. Of interest, in the second study an antibody-based microarray was carried out in lysates from chemotherapy-sensitive or -resistant luminal breast carcinomas identifying 38 differential proteins among which was 14-3-3 theta/tau. Clinical validation in pretreatment biopsies by IHC was possible for three of these proteins (14-3-3 thea/tau, Bcl-XL and tBID). While the presence of tBID was associated with chemotherapy sensitivity, 14-3-3 theta/tau staining was associated with resistance.
### Table 2. Summary of proteomic studies in hormone receptor-positive breast cancer.

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of sample</th>
<th>Comparison</th>
<th>Proteomic approach</th>
<th>Proteins identified</th>
<th>Validation/Verification</th>
<th>Candidate proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hormone receptor positive-breast cancer vs normal tissue</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weitzel et al. 27</td>
<td>Tissue</td>
<td>18 ER+ BC vs adjacent benign tissue</td>
<td>2D-DIGE, MALDI-TOF MS</td>
<td>Total: 50 in total</td>
<td>Verification: WB</td>
<td>PPlaseB, Rho-GDI α, TPM4</td>
</tr>
<tr>
<td>Cha et al. 29</td>
<td>LCM tissue</td>
<td>18 ER+ BC vs adjacent benign tissue</td>
<td>GeLC-MS/MS</td>
<td>Differential: 298</td>
<td>Not done</td>
<td>-</td>
</tr>
<tr>
<td><strong>Hormone receptor-positive vs –negative breast cancer</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traub et al. 30</td>
<td>Tissue</td>
<td>9 ER+ BC vs 9 healthy tissue</td>
<td>HPLC, MALDI-TOF MS</td>
<td>Not stated</td>
<td>Validation: IHC</td>
<td>Thymosin α1</td>
</tr>
<tr>
<td>Neubauer et al. 32</td>
<td>LCM tissue</td>
<td>39 ER+ vs 41 ER- BC</td>
<td>2D-PAGE, MALDI-TOF MS</td>
<td>Differential: 16</td>
<td>Validation: IFA</td>
<td>PGRMC1</td>
</tr>
<tr>
<td>Rezaul et al. 34</td>
<td>LCM tissue</td>
<td>8 ER+ vs 8 ER- BC</td>
<td>GeLC-MS/MS</td>
<td>Differential: 236</td>
<td>Validation: IHC</td>
<td>Liprin-α1, β-arrestin-1, fascin, DAP5</td>
</tr>
<tr>
<td>Sahab et al. 30</td>
<td>LCM tissue</td>
<td>3 ER+ vs 3 ER- BC</td>
<td>2D-PAGE, MALDI-TOF MS</td>
<td>Differential: 22</td>
<td>Verification: WB</td>
<td>α1-antitrypsin</td>
</tr>
<tr>
<td>Semaan et al. 35</td>
<td>Tissue</td>
<td>ER+ vs adherent ER- BC</td>
<td>2D-PAGE, LC-LTQ/FTICR MS</td>
<td>Differential glyco-proteins: 11</td>
<td>Verification: WB</td>
<td>ABAT, PDZK1</td>
</tr>
<tr>
<td>Pavlou et al. 37</td>
<td>Cell line secre-tome</td>
<td>ER+ vs ER- BC</td>
<td>2D LC-MS/MS</td>
<td>Differential: 31</td>
<td>Validation: Targeted MS</td>
<td>EMMPRIN, ENPP1, EIF3E, GNB4</td>
</tr>
<tr>
<td><strong>Hormone receptor-positive breast cancer and treatment response</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Umar et al. 39</td>
<td>LCM tissue</td>
<td>3 luminal vs 3 basal vs 3 HER2+</td>
<td>nanoLC-FTICR MS</td>
<td>Differential: 100</td>
<td>Validation: Targeted MS, IHC</td>
<td>-</td>
</tr>
<tr>
<td>Hodgkinson et al. 40</td>
<td>Tissue</td>
<td>24 tamoxifen-sensitive, ER+ vs 27 tamoxifen-resistant, ER+ BC</td>
<td>2D-PAGE, MALDI-TOF MS</td>
<td>Differential: 57</td>
<td>Not done</td>
<td>14-3-3 theta/tau, tBID</td>
</tr>
<tr>
<td>Hodgkinson et al. 41</td>
<td>Tissue</td>
<td>2 chemosensitive, ER+ vs 2 chemoresistant, ER+ BC</td>
<td>Antibody microarray</td>
<td>Differential: 45</td>
<td>Verification: WB, IHC</td>
<td>α1-antitrypsin</td>
</tr>
</tbody>
</table>

2D-DIGE, two-dimensional differential gel electrophoresis; 2D-PAGE, two-dimensional gel electrophoresis; BC, breast cancer; ER, estrogen receptor; IHC, immunohistochemistry; GeLC-MS/MS, 1D gel liquid chromatography MS/MS; GEM, genetically engineered mouse; MALDI-TOF, matrix-assisted laser desorption ionization time of flight; MS, mass spectrometry; RPPA, reverse phase protein array; TNBC, triple negative breast cancer; WB, western blot
HER2-positive subtype

Approximately 4–7% of the breast carcinomas can be classified as HER2-enriched subtype \(^9,12,25\), which is characterized by amplification or overexpression of HER2 oncogene and genes located at the HER2 amplicon on chromosome 17q22.24 (Table 1) \(^4,5\). Histological and IHC features including HER2 overexpression, negative ER and PgR, and high histological grade. Compared with luminal subtypes, patients with HER2-enriched subtype have a high incidence of advanced disease at the time of diagnosis, an increased risk of metastasis preferentially to brain, lung and liver, high relapse rate and short overall survival \(^9,13,42\). The introduction of the HER2-targeting agent trastuzumab in combination with chemotherapy has improved the outcome of patients with this subtype in the adjuvant and metastatic setting.

Differential expression of proteins in HER2-positive versus HER2-negative breast cancer

In several studies, HER2-positive breast cancer has been compared with HER2-negative breast cancer to characterize proteins potentially involved in HER2-driven tumor progression (Table 3). Durán et al. \(^43\) have conducted proteomic analysis to identify 75 differential proteins in pooled samples of HER2-positive breast cancer with or without lymph node metastasis. Ninety-six differential proteins were identified between non-metastasized HER2-positive and -negative breast cancer. Of the 20 candidates selected for IHC validation, expression of only three proteins (HSP90α, laminin and GSTP1) were consistent with positive HER2 overexpression. No candidates related to lymph node status could be validated. Zhang et al. \(^44\) have analyzed a small set of matched hormone receptor-negative/HER2-positive and -negative breast cancer tissues (3 for each condition). Nine proteins were upregulated in HER2-positive tissues and involved in processes such as lipid and glucose metabolism, stress-mediated chaperonage and detoxification. Of those, FASN, HSP27, PGK1 and GLO were validated by IHC. Using a similar proteomic approach, the same group has subsequently identified CK19 \(^45\) and HNRNPH1 \(^46\) as protein biomarkers positively related to HER2-positive breast cancer, while GRP78/BiP and RKIP \(^46\) demonstrated a negative relation. In previous reports, these proteins have been associated with tumor survival (RKIP, GRP78/BiP) \(^47,48\), chemoresistance (GRP78/BiP) \(^49,50\), and prognosis (RKIP, CK19) \(^51-53\), but the possible relation with HER2 status and clinical relevance await further investigation.

Differential protein phosphorylation in HER2-positive breast cancer

Despite clinical benefit of targeted therapy such as trastuzumab or lapatinib for HER2-positive breast cancer, resistance remains a significant problem and necessitates the discovery of novel targets. To this end, MS-based phosphoproteomics has been employed to identify protein kinases conferring HER2-signaling pathways \(^54,55\). Bose et al. \(^54\) have examined global protein expressions and phosphorylations in HER2-transfected cell lines. The analysis yielded 198 proteins with increased and 81 proteins with decreased phosphorylation. Apart from proteins known to be involved in HER2 and EGFR signaling, novel proteins...
were identified previously not linked to HER2 signaling. Phosphorylation of a subset of HER2- and EGFR-related proteins (PLCγ1, DOK1, STAT1, δ-catenin) was verified by Western blot. *In vitro* pharmacological inhibition by a dual EGFR and HER2 tyrosine kinase blocker (PD168393) decreased phosphorylation of two novel proteins (AXL and FYB) pointing towards involvement of the HER2-signaling pathway. In a recent study, *in vitro* overexpression of AXL has been demonstrated to be involved in acquired resistance to lapatinib, while treatment with a multikinase inhibitor of AXL, MET and VEGFR restored lapatinib and trastuzumab sensitivity. Rexel et al. have compared phosphoproteomics among six lapatinib-resistant breast cancer cell lines and parental cell lines showing increased phosphorylation of SRC family kinases, including LYN, SRC, LCK, FYN, FRK, FGR. When compared with baseline, increased expression of LYN, LCK and FYN was detected in HER2-positive breast cancer biopsies from patients after 6 weeks of treatment with lapatinib. Exposure to a small-molecule inhibitor of SRC kinases (AZD0530) restored lapatinib sensitivity *in vitro* and in xenograft models. Interestingly, increased phosphorylation of EGFR- and SRC-related kinases have been identified by phosphoproteomics comparing two trastuzumab-resistant with two-sensitive breast cancer cell lines. Gene silencing with siRNA for selected proteins restored trastuzumab sensitivity in resistant cell lines. Taken together, these phosphoproteomic studies may indicate novel activated proteins as targets for development of agents in HER2-positive breast cancer refractory to trastuzumab and/or lapatinib.

### Basal-like subtype

Basal-like subtype, with a prevalence of approximately 10–20\%\(^7,9,12\), is characterized by overexpression of genes related to cells in the basal (myoepithelial) layer of mammary ducts including CK5/6, CK17, integrin b4, laminin, fatty acid binding protein 7 and EGFR (Table 1).\(^4,5\) Basal-like breast carcinoma is usually of high histological grade and IHC shows positive staining for CK5/6, CK17 and a high Ki-67 index. This subtype lacks ER and PgR expression as well as HER2 overexpression, hence it is often referred to as triple-negative breast cancer (TNBC). Although basal-like breast cancer and TNBC are not identical, they largely overlap in approximately 70–80\% of cases in terms of morphology (ductal carcinoma of high histological grade), clinical presentation (young women), prognosis (early metastasis), association with BRCA1 gene mutation and therapeutic options (no benefit from endocrine or HER2-based therapy).\(^6,58,59\) Ongoing studies will unravel differences between basal-like subtype and TNBC. This section will summarize proteomic studies on basal-like subtype or TNBC.

### Differential expression of proteins in triple-negative breast cancer

Our group has compared the proteome of breast tumors arising from genetically engineered mouse models of BRCA1-deficient and -proficient TNBC.\(^60\) The majority of proteins found to be upregulated in BRCA1-deficient breast cancer were related to DNA replication, recombination and repair. Importantly, the analysis provided a signature of 45 proteins, which was validated *in silico* on publicly available gene expression data. The protein signature was able to classify BRCA1- and BRCA2-related breast cancer from sporadic breast...
cancer across multiple gene expression data sets and showed prognostic value. A number of proteins are currently subjected to the validation process.

Despite having a poor prognosis, some women with TNBC remain disease-free for a long period of time indicating the need for additional markers to discriminate patients with relatively favorable and poor prognosis. One study stands out in terms of a large sample size including 1,200 breast carcinoma tissues of which 184 samples were TNBC\(^61\). Based on literature findings related to TNBC, 42 proteins were selected for IHC staining in 924 breast carcinomas. A total of 24 proteins showed prognostic value in the univariate analysis and was subsequently tested in the TNBC subset. A signature of 11 proteins was able to correctly classify for disease outcome in 89% of these cases (Table 4). In order to identify proteins for increased risk of locoregional metastasis, Greenwood et al.\(^62\) have analyzed pooled TNBC tissues from patients with and without lymph node metastasis. Bioinformatics revealed enrichment of proteins involved in cytokine and growth factor signaling, regulation of the actin cytoskeleton and mRNA processing in the metastatic samples. Verification by IHC and Western blot showed high STAT1 protein expression in lymph node-positive compared with lymph node-negative TNBC, while CD74 level correlated with STAT1 abundance. Subsequent \textit{in vitro} experiments showed that abundant levels of STAT1 led to increased migration and invasion, while CD74 overexpression led to increased adhesion. Sun et al.\(^63\) have used a mouse metastatic breast cancer model to identify metastasis-associated proteins of TNBC. Comparison between primary breast tumor and metastatic tissues revealed 19 metastasis-associated proteins. IHC validation was performed on 100 human breast cancer tissues (61 TNBC and 39 ER+ and PR+ breast cancer) with or without metastasis to explore the clinical relevance. Among selected proteins, the expression of HSP70 and 14-3-3 was closely associated with metastasis, specifically in TNBC tissues. In the second validation step in 66 TNBC tissue, the combination of HSP70 and 14-3-3 showed the highest sensitivity (93%) to predict metastasis, but the specificity (51%) was rather low.

\textbf{Differential protein phosphorylation in basal-like breast cancer}

Abnormal protein phosphorylation affects a multitude of intracellular processes and is considered an important process for the development of cancer. In the study of Hochgräfe et al.\(^64\), MS-based analysis of phosphoproteins has been performed in multiple breast cancer cell lines to identify potential therapeutic targets and prognostic biomarkers useful in basal-like breast cancer. Different protein phosphorylation and expression patterns among breast cancer subtypes were detected. In particular, basal-like breast cancer cell lines showed elevated tyrosine phosphorylation including EGFR and MET, FAK, EPHA2, LYN and SRC family-associated kinases. Western blotting revealed increased expression of MET, EGFR, LYN and CAV1. Validation by IHC in breast carcinomas showed that elevated LYN expression was significantly related to the basal-like subtype and to decreased survival. Functional knockdown of LYN reduced invasiveness \textit{in vitro}. Besides LYN, overexpression of MET and FAK have previously been related to poor prognosis in two IHC studies\(^61,65\).
Table 3. Summary of proteomic studies in HER2-positive breast cancer.

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of sample</th>
<th>Comparison</th>
<th>Proteomic approach</th>
<th>Proteins identified</th>
<th>Validation/Verification</th>
<th>Candidate proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Durán et al.</td>
<td>Tissue</td>
<td>2 primary HER2+ vs 8 primary HER2- BC</td>
<td>2D-DIGE, MALDI-TOF MS</td>
<td>Differential: 96; 75</td>
<td>Validation: IHC</td>
<td>HSP90α, laminin, GSTP1</td>
</tr>
<tr>
<td>Zhang et al.</td>
<td>LCM tissue</td>
<td>3 HER2+ vs 3 HER2- BC</td>
<td>2D-PAGE, MALDI-TOF MS</td>
<td>Differential: 9</td>
<td>Validation: WB, IHC</td>
<td>FASN, HSP27, PGK1, GLO</td>
</tr>
<tr>
<td>Zhang et al.</td>
<td>LCM tissue</td>
<td>3 HER2+ vs 3 HER2- BC</td>
<td>2D-PAGE, MALDI-TOF MS</td>
<td>Differential: 7</td>
<td>Validation: RT-PCR, RPPA, IHC</td>
<td>CK19</td>
</tr>
<tr>
<td>Zhang et al.</td>
<td>LCM tissue</td>
<td>7 HER2+ vs 7 HER2- BC</td>
<td>2D-PAGE, MALDI-TOF MS</td>
<td>Differential: 21</td>
<td>Validation: WB, IHC</td>
<td>HNRNPH1, GRP78/BiP, RKIP</td>
</tr>
<tr>
<td>Schulz et al.</td>
<td>Tissue</td>
<td>19 HER2+ BC vs 15 TNBC</td>
<td>2D-DIGE, MALDI-TOF MS</td>
<td>Differential: 34</td>
<td>Validation: WB, IHC</td>
<td>CK7, GAPDH, PGK1</td>
</tr>
</tbody>
</table>

Phosphoproteins in HER2-positive breast cancer

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of sample</th>
<th>Comparison</th>
<th>Proteomic approach</th>
<th>Proteins identified</th>
<th>Validation/Verification</th>
<th>Candidate proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bose et al.</td>
<td>Cell lines</td>
<td>HER2 transfected cells vs controls ± EGFR/HER2 inhibitor</td>
<td>Phosphoproteins enrichment, SILAC, nanoLC-MS/MS</td>
<td>Differential: 279 proteins with differential phosphorylations</td>
<td>Verification: WB</td>
<td>PLC1, DOK1, STAT1, δ-catenin</td>
</tr>
<tr>
<td>Rexer et al.</td>
<td>Cell lines</td>
<td>6 lapatinib-resistant HER2+ vs 6 lapatinib-sensitive parental cells</td>
<td>Phosphoproteins enrichment, GeLC-MS/MS</td>
<td>Differential: 20 tyrosin phosphorylation sites from 22 proteins</td>
<td>Validation: Microarray</td>
<td>LYN, LCK, FYN</td>
</tr>
<tr>
<td>Boyer et al.</td>
<td>Cell lines</td>
<td>2 trastuzumab-resistant HER2+ vs 2 trastuzumab-sensitive parental cells</td>
<td>Phosphoproteins enrichment, SILAC, nanoLC-MS/MS</td>
<td>Differential: 62 and 106 proteins with differential phosphorylations per cell line, 3 proteins with increased phosphorylation were found in both cell lines</td>
<td>Verification: WB, siRNA</td>
<td>WB: EGFR, HER4, CDCP3/Trask, FAK, PXN SiRNA: MAPK1, PXN, FAM83A, EFS</td>
</tr>
</tbody>
</table>

2D-DIGE, two-dimensional differential gel electrophoresis; 2D-PAGE, two-dimensional gel electrophoresis; BC, breast cancer; GeLC-MS/MS, 1D gel liquid chromatography MS/MS; EGFR, epidermal growth factor receptor; IHC, immunohistochemistry; LCM, laser capture microdissection; MALDI-TOF, matrix-assisted laser desorption ionization time of flight; MS, mass spectrometry; RPPA, reverse phase protein array; RT-PCR, reverse transcription polymerase chain reaction; SILAC, stable isotope labeling with amino acids in cell culture; TNBC, triple negative breast cancer; WB, western blot
**Table 4. Summary of proteomic studies in basal-like or triple-negative breast cancer.**

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of sample</th>
<th>Comparison</th>
<th>Proteomic approach</th>
<th>Proteins identified</th>
<th>Validation/Verification</th>
<th>Candidate proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warmoes et al.</td>
<td>GEM</td>
<td>3 <em>BRCA1</em>-deficient vs 6 <em>BRCA1</em>-proficient TNBC</td>
<td>GeLC-MS/MS</td>
<td>Differential: 801</td>
<td>In silico validation: microarray datasets</td>
<td>45 protein signature</td>
</tr>
<tr>
<td>Charpin et al.</td>
<td>Tissue</td>
<td>184 TNBC with poor and good prognosis</td>
<td>IHC on TMA</td>
<td>Not applicable</td>
<td>Not applicable</td>
<td>STAT1, PTEN, pMAPK, P38, P27, P21, MAS-PIN, CD10, FAK, EGFR, Caveolin</td>
</tr>
<tr>
<td>Greenwood et al.</td>
<td>Tissue</td>
<td>6 lymph node-positive vs 6 lymph node-negative TNBC</td>
<td>GeLC-MS/MS</td>
<td>Differential: 339</td>
<td>Verification: WB, IHC</td>
<td>STAT1, CD74</td>
</tr>
<tr>
<td>Sun et al.</td>
<td>GEM</td>
<td>Primary TNBC vs metastasized TNBC</td>
<td>2D-PAGE, MALDI-TOF MS</td>
<td>Differential: 19</td>
<td>Validation: IHC</td>
<td>HSP70, 14-3-3</td>
</tr>
<tr>
<td>Schulz et al.</td>
<td>Tissue</td>
<td>15 TNBC vs 19 HER2+ BC</td>
<td>2D-DIGE, MALDI-TOF MS</td>
<td>Differential: 34</td>
<td>Verification: WB, IHC</td>
<td>CK14</td>
</tr>
</tbody>
</table>

Phosphoproteins in basal-like subtype

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of sample</th>
<th>Comparison</th>
<th>Proteomic approach</th>
<th>Total: 544 tyrosin phosphorylation sites from 295 non-redundant proteins</th>
<th>Validation/Verification</th>
<th>Candidate proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hochgräfe et al.</td>
<td>Cell lines</td>
<td>8 luminal vs 7 basal</td>
<td>Phosphoproteins enrichment, LC-MS/MS</td>
<td>RPPA, IHC</td>
<td>EGFR Y1173, MET Y1234/35, LYN</td>
<td></td>
</tr>
</tbody>
</table>

2D-DIGE, two-dimensional differential gel electrophoresis; 2D-PAGE, two-dimensional gel electrophoresis; BC, breast cancer; GeLC-MS/MS, 1D gel liquid chromatography MS/MS; GEM, genetically engineered mouse; IHC, immunohistochemistry; MALDI-TOF, matrix-assisted laser desorption ionization time of flight; MS, mass spectrometry; RPPA, reverse phase protein array; TNBC, triple negative breast cancer; WB, western blot
### Table 5. Summary of proteomic studies in plasma to identify proteins related to breast cancer subtypes.

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of sample</th>
<th>Comparison</th>
<th>Proteomic approach</th>
<th>Proteins identified</th>
<th>Validation/Verification</th>
<th>Candidate proteins</th>
</tr>
</thead>
</table>
| Nakshatri et al.\(^66\) | Plasma         | 6 luminal A vs 18 controls  
5 luminal B vs 18 controls  
10 TNBC vs 18 controls | Depletion, nanoLC-MS/MS | Differential – Luminal A: 0  
Differential – Luminal B: 9  
Differential – TNBC: 2 | Not done | - |
| Pitteri et al.\(^67\) | Plasma         | 420 prediagnostic plasma from ER+ BC vs 420 controls | Depletion, nanoLC-MS/MS | Differential: 57 | Validation: ELISA | EGFR |
| Amon et al.\(^68\) | Plasma         | 420 prediagnostic plasma from ER+ BC vs 420 controls | Depletion, nanoLC-MS/MS | Total: 467 | Not done | - |
| Gonzalez et al.\(^69\) | Plasma         | 20 ER+/HER2- vs 18 controls  
24 ER+/HER2+ vs 21 controls  
18 ER-/HER2+ BC vs 18 controls  
24 ER-/HER2- vs 21 controls | Antibody microarray | Differential – ER+/HER2-: 4  
Differential – ER+/HER2+: 2  
Differential – ER-/HER2+: 3  
Differential – ER-/HER2-: 2 | Not done | - |
| Li et al.\(^70\) | Plasma         | 14 TNBC vs 14 controls | Antibody microarray | Differential: 93 | Validation: antibody microarray | DUSP9, EED, EFNA5, ITGB1, PPMT1 |

BC, breast cancer; ELISA, enzyme-linked immunosorbent assay; ER, estrogen receptor; MS, mass spectrometry; TNBC, triple-negative breast cancer; WB, western blot
Plasma proteomics to identify circulating proteins related to breast cancer subtypes

Although human plasma may represent an attractive, easily accessible source for biomarker discovery, one will encounter technical challenges for MS-based proteomic analysis due to its complexity and large dynamic range of protein concentrations. Consequently, few studies are available on MS-based proteomics of plasma obtained from breast cancer patients to characterize circulating proteins (Table 5). Nakshatri et al. have employed this technique to analyze plasma obtained from 18 healthy controls and 29 breast cancer patients. The vast majority of patients had metastatic disease and had received prior chemotherapy at the time of blood collection. Highly abundant proteins, such as albumin, were immunodepleted from plasma prior to MS analysis. No significant changes in the proteome were observed in six patients with luminal A subtype as compared with that of control plasma. The investigators reasoned that the lack of plasma proteome variation in the luminal A subtype reflected its low aggressive nature, since measuring low-abundant proteins in plasma is difficult. In the luminal B subtype, eight proteins involved in immune response were significantly increased, whereas 12 proteins involved in free radical scavenging were significantly decreased. Two complement factors (CFHR3 and C4BPB) were identified being elevated in TNBC compared with healthy controls. In a recent study of Pitteri et al., attempts have been described to identify protein biomarkers for early detection of ER-positive breast cancer. In the initial discovery phase, pooled plasma samples from 490 patients with ER-positive breast cancer ≤17 months before diagnosis versus pooled samples from 490 age- and race-matched healthy controls were analyzed by MS-based proteomics. The validation work was focused on ER-positive breast cancer; of the 503 total proteins quantified in the discovery experiments, there were 57 candidate proteins associated with ≥1.15 fold-change and a \( P \) value of <0.10. Seven out of 57 were selected for which an ELISA assay was readily available. After validation in two independent cohorts, the pre-diagnostic level of plasma EGFR was significant elevated in women who developed breast cancer. This single marker, however, showed moderate sensitivity and specificity for early diagnosis. In another study from the same group, similar MS-based strategy has been applied in pooled plasma samples from 420 patients with ER-positive breast cancer ≤17 months before diagnosis versus pooled samples from 420 age- and race-matched healthy controls aiming to illuminate molecular pathways vital for breast cancer development. The investigators detected an increased level of glycolysis-related proteins in plasma of breast cancer patients compared with that in controls. In particular, the increase of these proteins was more pronounced in plasma samples taken close to the diagnosis. However, the origin of these glycolysis proteins as well as their role in ER-positive breast cancer remains to be elucidated.

An antibody microarray offers an alternative opportunity to screen a panel of defined proteins. In the study of Gonzalez et al., a panel of 23 proteins was screened in plasma taken from patients with an actual diagnosis of breast cancer of different subtypes (\( n = 85 \)) or benign breast disease (\( n = 39 \)). The breast cancer subtype was defined by ER and HER2
status. The investigators detected 10 proteins that were significantly altered in at least one breast cancer subtype. These included EGFR ligands, metalloproteases, cytokines and pro-angiogenic factors. Only RANTES was increased in all four subtypes. In comparison with women with benign breast disease, four proteins were increased in plasma from patients with ER-positive/HER2-negative breast cancer (amphiregulin, heparin-binding EGF, RANTES, TGFα) and three proteins (EGF, heparin-binding EGF, RANTES) in plasma from those with ER-negative/HER2-positive breast cancer. Furthermore, plasma samples from patients with ER-positive/HER2-positive breast cancer contained higher levels of platelet-derived growth factor (PDGF) and RANTES, while plasma concentrations of VEGF and RANTES were significantly increased in ER-negative/HER2-negative samples. In another study using antibody-based assay, Li et al. 70 have analyzed plasma samples from patients with TNBC at diagnosis and healthy controls. In the initial discovery phase, 93 differential proteins were identified. By validation in a separate cohort, 29 proteins were confirmed. Using stringent criteria, five proteins (DUSP9, EED, EFNA5, ITGB1 and PTPMT1) were retained and might serve as markers for early detection of TNBC.

Proteomic studies in breast cancer subtyping: current status and considerations

Breast cancer subtyping by IHC and gene expression profiling has contributed to a clinically relevant classification providing information about prognosis 9,13,42 and treatment response 8,10. As a complementary approach, proteomic analysis may identify novel protein biomarkers related to specific breast carcinomas with distinct underlying gene aberrations to further refine the existing molecular classification. In addition, this technique may provide readily applicable protein biomarkers for use in IHC classification as well as point towards possible targets for treatment. In this review, we have summarized the current state of proteomic research in breast cancer subtypes with main focus on proteomic discovery. Few studies have already indicated particular proteins potentially useful in further classification of breast carcinoma subtypes, as diagnostic or prognostic biomarkers, as biomarkers for prediction of drug resistance, as well as their possible association with metastases formation. Based on current knowledge of proteomics in breast cancer subtypes, we will discuss some considerations and directions for future research.

Selection of methods for accurate and sensitive protein quantification and identification is of critical importance, because the method of choice will influence the quality of data. The majority of proteomic studies discussed in this review has been published in the past five years and the main technology used was a traditional 2D-PAGE-based method for protein fractionation and quantification. This method typically detects up to a thousand protein spots, differential spots are picked up separately and analyzed by an MS-based instrument usually providing tens to few hundreds of identified proteins (Tables 2–4). Current MS-
based proteomics is capable to identify thousands of proteins in a limited amount of sample providing far more detailed information about altered molecular processes. Typically, our experience and that of many others \textsuperscript{16,60,62} have shown that MS-based proteomics in complex biological samples is highly reproducible. Therefore, MS-based proteomics has emerged as the method of choice for protein profiling in breast cancer studies.

Although breast cancer classification has gained broad interest in different research fields over the past decade, this has not yet been widely adopted in the field of breast cancer proteomic research. At present, discrimination of breast cancer subtypes has been addressed in only a few studies reviewed here \textsuperscript{64,66,69}. Consequently, little is known concerning proteins related to breast cancer subtypes and differences in proteome among subtypes. Few studies have made a distinction between luminal A and B subtypes in hormone receptor-positive breast cancer \textsuperscript{66,69}, while basal-like subtype remains unexplored given the fact that most studies focused on TNBC. The paucity of proteomic studies in breast cancer subtypes may partly be due to the lack of a uniform definition. Although both gene signatures and IHC detection of protein markers are accepted methods for definition, the latter is used more commonly for clinical and practical convenience. It is our expectation that breast cancer subtypes will become increasingly investigated using proteomic technologies. Recognizing the need for a uniform definition, a standardized IHC-based discrimination of breast cancer subtypes has recently been proposed at the 12\textsuperscript{th} St. Gallen International Breast Cancer Conference (Table 1) \textsuperscript{71}. It is expected that the adoption of the proposed definition will enhance consistency of the results from future studies and allows comparison across different studies employing proteomic and genomic technologies.

Most breast cancer proteomic studies are based on a small sample size in the initial stage of biomarker discovery. The candidate biomarkers selected may not fully reflect the biological heterogeneity of breast cancer. Moreover, few studies have validated potential candidate biomarkers. In general, findings obtained from an \textit{in vitro} model system should be verified in representative clinical samples and followed by validation in an independent cohort. Validation of protein biomarkers is preferably performed in clinical samples with patients’ follow-up data to assess their potential clinical relevance. However, it also requires the availability of alternative protein detection assays. As outlined earlier, most frequently used validation methods employ an antibody-based approach resulting in a bias towards proteins for which antibodies are available while ignoring other promising novel proteins. Recent MS instruments offer targeted follow-up of novel proteins of interest, hereby allowing accurate, independent validation. Furthermore, gene expression levels as surrogate measures of protein abundance have previously been used as a valuable source for \textit{in silico} validation \textsuperscript{60}. The large amount of publicly available gene expression datasets with clinical information may facilitate the first assessment of candidate proteins in a large population. It has to be taken in mind that discordance can occur between protein abundance and gene expression levels. Taken together, factors concerning sample size, validation and new technological developments should be taken into consideration in order to exploit the full
potential of proteomics.

The tumor microenvironment represents the cross-talk between tumor and host and plays a crucial role in promoting the process of tumor progression. In this context, observations from cell lines or laser microdissected tumor cells merely reflect processes within a tumor, while the equally important host-related factors are neglected. Of advantage, analysis of whole tumor biopsies will capture features of tumor and host. Although rarely explored in studies reviewed here, insight gained in differences in tumor microenvironment among breast cancer subtypes by proteomic technologies may advance the current breast cancer classification.

The potential benefits of proteomic technologies for characterizing proteomic profiles related to breast cancer subtypes and for defining potential biomarkers within a breast cancer subtype have been explored in several studies reviewed here. Tissue-based protein biomarkers might be incorporated into the traditional IHC panel to improve stratification of therapeutic response in ER-positive breast cancer \(^{39,41}\) or to assess prognosis of patients with TNBC \(^{60,63}\). Circulating protein biomarkers might be used to diagnose ER-positive breast cancer \(^{67}\) or TNBC \(^{70}\). The application of proteomics can also contribute to our understanding of multiple pathways modulated by overexpression of ER \(^{28,30,32,34,35}\) or HER2 \(^{43,46,72}\). Moreover, proteomics can yield important information about resistance mechanisms to tamoxifen for ER-positive breast cancer \(^{39}\) and trastuzumab or lapatinib for HER2-positive breast cancer \(^{55,57}\). Phosphoproteomics is a relatively new, but rapidly developing field of proteomic research that allows identification of phosphorylated proteins. The abundance of certain phosphorylated proteins may reveal novel activated pathways representing potential targets for development of drugs. Few studies have pointed towards potential novel targets in basal-like \(^{64}\) or HER2-positive breast cancer \(^{54,55,57}\). Furthermore, phosphoproteomics in clinical specimen can guide clinicians in selecting patients eligible for additional kinase inhibitors. Nonetheless, the potential of phosphoproteomics remains to be explored.

Concluding remarks

Much progress has been made in proteomic technologies allowing in-depth analysis of protein expression patterns in a multitude of clinical samples. Proteomics in breast cancer classification, however, is still in its infancy. Hence, there are already a few data on protein biomarkers specifically related to each breast cancer subtype. Considerations with regard to analytical method, definition of breast cancer subtypes and validation test should be taken into account in further proteomic studies.

Supplementary material is available at http://dx.doi.org/10.1016/j.ctrv.2013.06.006
REFERENCE


27. Weitzel LR, Byers T, Allen J, et al. Discovery and verification of protein differences between Er po-
PROTEOMICS AND BREAST CANCER SUBTYPES


1;128(3):726-31.


