Published in final edited form as:

Expert Rev Proteomics. 2008 June ; 5(3): 457–467. doi:10.1586/14789450.5.3.457.

Breast tumor metastasis: analysis via proteomic profiling

Steve Goodison, PhD[†] and

Department of Surgery, University of Florida, 653 West 8th Street, Jacksonville, FL 32209, USA,
Tel.: +1 904 633 0978, Fax: +1 904 633 0979

Virginia Urquidi, PhD

Department of Medicine, University of Florida, 655 West 11th Street, Jacksonville, FL
32206-3516, USA, Tel.: +1 904 633 0977, Fax: +1 904 633 0979

Steve Goodison: steve.goodison@jax.ufl.edu; Virginia Urquidi: virginia.urquidi@jax.ufl.edu

Abstract

The ability to predict the metastatic behavior of a patient's cancer, as well as to detect and eradicate such recurrences, remain major clinical challenges in oncology. While many potential molecular biomarkers have been identified and tested previously, none have greatly improved the accuracy of specimen evaluation over routine histopathological criteria and, to date, they predict individual outcomes poorly. The ongoing development of high-throughput proteomic profiling technologies is opening new avenues for the investigation of cancer and, through application in tissue-based studies and animal models, will facilitate the identification of molecular signatures that are associated with breast tumor cell phenotype. The appropriate use of these approaches has the potential to provide efficient biomarkers, and to improve our knowledge of tumor biology. This, in turn, will enable the development of targeted therapeutics aimed at ameliorating the lethal dissemination of breast cancer. In this review, we focus on the accumulating proteomic signatures of breast tumor progression, particularly those that correlate with the occurrence of distant metastases, and discuss some of the expected future developments in the field.

Keywords

2D separation; biomarkers; breast cancer; laser-capture microdissection; molecular prognostics; oncoproteomics; tissue microarray

Breast cancer is the second most common cause of death from cancer among women in the USA. In 2007, it is estimated that approximately 178,480 new cases of breast cancer will be diagnosed, and 40,460 women are expected to die from this disease. The lethal aspect of breast cancer is the recurrence of therapeutically resistant disseminated disease, which in many patients has already occurred by the time the primary tumor is diagnosed.

© 2008 Expert Reviews Ltd

[†]Author for correspondence: Department of Surgery, University of Florida, 653 West 8th Street, Jacksonville, FL 32209, USA, Tel.: +1 904 633 0978, Fax: +1 904 633 0979, steve.goodison@jax.ufl.edu.

For reprint orders, please contact: reprints@expert-reviews.com

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Currently used postoperative treatment guidelines use consensus criteria, which determine whether a patient is at a high risk of distant metastases based on a panel of clinical (age), histopathological (lymph node status, tumor size, histological grade) and cell biological markers, including the presence of estrogen and progesterone receptors. Due to low specificity, these criteria are generally considered to be insufficient for accurately predicting metastatic behavior; hence, more accurate prognostic factors are urgently needed for the assessment of newly diagnosed patients. While a number of potential molecular biomarkers have been individually identified and tested, none have improved the stratification of disease and they predict individual outcomes poorly.

The concept that the presence or absence of one molecular marker will aid prognostic evaluation has not proved to be the case. This makes sense when one considers the complex interactions between various molecules within a single pathway, the cross-talk between molecular pathways, and the redundancy of cellular signaling pathways. There must be an evolution from single-marker/single-pathway research to a more global assessment of breast cancer. The search for molecular signatures or fingerprints associated with specific cancer phenotypes requires high-throughput profiling techniques coupled with sophisticated bioinformatics tools for complex data analysis and pattern recognition. Accordingly, recently developed high-performance screening technologies are now revolutionizing the study of disease pathogenesis. The surge in such studies has been led by the genomics field, where oligonucleotide microarrays have been employed to identify gross chromosomal changes [1] and gene expression patterns that distinguish molecular subtypes of cancer (including breast cancer) and are predictive of metastatic relapse [2–5]. Indeed, a 70-gene signature that is reported to be predictive of breast cancer disease recurrence is the subject of a large prospective trial known as MINDACT [6]. These studies have the potential to improve our knowledge of tumor molecular biology and to provide pivotal information for clinical evaluation of breast cancer progression. However, although the pattern of gene expression may be abnormal in a tissue with a pathological lesion, there can be a poor correlation between the level of the transcription of different genes and the relative abundance within the tissue of the corresponding proteins. Consequently, the information about a pathological process that can be derived at the level of gene transcription is incomplete.

There have been parallel improvements in technologies to detect, identify and characterize proteins in biological samples. Newly developed, high-throughput techniques, such as TOF mass spectrometry (MS), SDS-PAGE with MALDI-MS and liquid-phase 2D separation techniques greatly facilitate the analysis of proteins in complex biological tissues and fluids. The application of multiple proteomic approaches to the analysis of patient's tissues and appropriate models of disease has the potential to greatly advance our understanding of the critical role specific proteins play in cancer development and progression. In this review, we focus on the proteomic analysis of breast cancer progression (particularly those that correlate with the occurrence of distant metastases) and discuss some of the key issues and expected future developments in the field.

Proteomic profiling of human tissues & fluids

High-throughput proteomics has been developed in order to attain the comprehensive display, quantitation and identification of the proteins expressed by the genome of an organism. The goal of clinical proteomics is to characterize protein pathways, networks and signaling events that are relevant in disease. Proteomic profiling can be target specific or of a global, unbiased nature. The targeted approach requires protein-specific reagents, most often antibodies, and thus is limited to the availability of such reagents. Unbiased, global protein profiling screens the unidentified protein component of a sample, detecting as many

signals as possible. Individual proteins may be subsequently identified if the corresponding signal is found to be associated with a phenotype in multisample analyses. Global profiling is the most challenging, but potentially the most informative, of proteomic approaches because the direct analysis and identification of proteins present in a tissue sample or specimen can reveal completely unexpected correlations of specific proteins with disease states.

Numerous proteomics technologies, including 2D-PAGE, 2D-DIGE, SELDI-TOF and MALDI-TOF, have been used in attempts to uncover molecular mechanisms associated with breast carcinoma at the global level. Other relatively high-throughput techniques, such as protein arrays and tissue microarrays (TMA), have, to date, played more of a role in the validation of both genomic and proteomic profiling studies. To date, in the majority of proteomic studies in which breast cancer was under investigation, the focus has been to identify biomarkers for diagnosis. This has typically been approached as a differential analysis comparing the expressed proteome in two or more distinct groups; for example, normal breast tissue and breast tumor tissue, or the serum from patients with or without tumor burden. In studies on the progression of breast cancer, the analysis requires the comparison of tissue samples and correlation with clinical outcome, such as disease-free survival or disease recurrence as distant metastasis.

The standard method for protein separation in complex mixtures has been 2D gel electrophoresis (2D-GE), in which proteins are separated by isoelectric focusing and according to molecular weight [7]. After separation, proteins of interest are isolated from the gel, digested and subjected to MS for identification. A limited number of studies have used 2D-GE methods for the analysis of breast cancer [8–11]. Although 2D-GE has excellent resolution, the requirement of large amounts of material is limiting in the clinical realm. A study by Nimeus *et al.* used 2D-GE of solid tissue samples to identify candidate proteins that distinguished breast cancer patients with or without disease recurrence within 4 years after adjuvant post-surgical treatment. Although the study only profiled 20 patients, several proteins associated with distant recurrence were revealed, including thioredoxin domain-containing protein 5 and ferritin [12]. As reported in numerous studies, the investigators in this study found it easier to predict clinical outcome for the estrogen-receptor (ER)-positive than for the ER-negative cohort, based on gene expression data [13]. A possible explanation for this could be that the ER-positive tumors are a more homogenous group than ER-negative tumors. It is noteworthy that the study by Nimeus *et al.* also successfully adapted a protein extraction protocol from tumor samples that allowed purification of both mRNA and proteins. This enables the comparison of proteomic and genomic profiling analyses from the exact same sample, and must occur more often as techniques evolve in both fields.

A 2D-DIGE technique has been developed to overcome issues of sensitivity and reproducibility associated with conventional 2D-GE [14]. This approach involves covalently labeling protein extracts with a selection of fluorescent dyes prior to separation. Differentially labeled protein samples can then be mixed and co-separated on the same gel. Digital detection of the fluorescent gel patterns enables the relative quantification of individual proteins in the test sample(s) by comparison with an in-gel reference sample. In a study of 12 breast tumors and the matched, associated lymph nodes, 2D-DIGE analysis was performed to determine whether the global pattern of protein expression in primary tumors compared with the associated lymph nodes [15]. Examples of proteins that were differentially expressed in the breast carcinomas relative to associated lymph nodes were ubiquitin-activating enzyme E1, transferrin, annexin VI, L-plastin, glutathione *S*-transferase A1 and protein disulfide isomerase.

The advent of versatile MS instrumentation has provided new proteomic approaches to unravel the complexities of clinical specimens. MS separates peptides or proteins according to their mass-to-charge ratio, and can be used to screen for differential expression directly, or by coupling to upstream protein separation methods [16]. MALDI-TOF-MS analysis is now a routine tool for the identification of proteins separated by 2D methods [17]. MALDI-TOF has also been applied directly to identify peptides sequestered from the serum of breast cancer patients. Using magnetic beads to perform peptide extraction and a MALDI-TOF-MS-based approach, Villanueva *et al.* demonstrated that specific subsets of serum peptides can provide accurate class discrimination between patients with and without breast cancer [18]. SELDI-TOF-MS is a relatively straightforward proteomic technology that can quantitatively analyze protein mixtures. Sample complexity is reduced by selectively capturing proteins of different classes on supports coated with chemical or biological baits. Examples of selection surfaces include those with hydrophobic, hydrophilic, anionic or cationic properties, antibodies, nucleic acids or ligands. In breast cancer, SELDI-TOF-MS has been used to investigate serum/plasma [19–21], nipple aspirate fluid [22,23] and solid tumor tissues [24]. Once again, the focus to date has been on the identification of diagnostic biomarkers, but recent studies have used the technique as a potential tool to predict outcome [24,25] and to monitor cancer treatment [26].

The most publicized application of SELDI-TOF-MS in clinical studies has been in the analysis of serum. Serum and/or plasma is readily available for clinical proteomic studies, and it is possible that this can act as a reliable surrogate tissue for pathologic processes, including cancer. Using SELDI-TOF-MS analysis of sera from patients with early breast cancer, Goncalves *et al.* identified a multiprotein index of serum proteins that were differentially expressed according to metastatic outcome [25]. Included in the components of the predictive multiprotein index were haptoglobin-1, Apo A-I and C-I, and transferrin. In a study applying SELDI-TOF-MS technology to breast tissue extracts, Ricolleau *et al.* identified ubiquitin and ferritin light chain as providing prognostic information with respect to metastatic disease recurrence [24]. There are major criticisms of this technique regarding lack of sensitivity and ability to detect disease-specific protein traces in a background of nonspecific proteins [27], but this may have been compounded by choosing serum, which is a particularly complex biological specimen, as the source material. Despite current skepticism, the SELDI-TOF approach is relatively easy to use and inexpensive, and thus it will very likely have a place in the proteomic armamentarium as instrumentation evolves. The lack of detection sensitivity of current techniques means that the majority of current biomarkers are high-abundance proteins. Once specific targets are identified, the sensitivity of detection can be greatly enhanced with specific reagents such as antibodies or peptide ligands, but increasing the sensitivity of profiling techniques within complex biological samples is the next required technical advance.

While the aforementioned techniques are able to discover new biomarkers, there is also the need for proteomics technologies to validate potential biomarkers in larger numbers of appropriate human specimens. The availability of TMA has greatly accelerated the analysis of large cohorts of patient samples, albeit querying usually only single or a handful of markers. The technique enables the analysis of DNA or RNA by *in situ* hybridization, or proteins by immunohistochemistry of hundreds or even thousands of tumor samples arrayed onto a glass slide. Cylindrical cores as small as 1 mm in diameter are selected from formalin-fixed paraffin-embedded (FFPE) archived tumor blocks and arrayed into a new paraffin block. Sections of this array block can then be interrogated by specific antibodies or nucleic acid probes. Limitations of the technique include the representation of the whole tissue specimen using small cores, and the qualitative, subjective nature of immunohistochemical evaluation. Thus, to date, this technique is seen as a tool for the validation of genomic and/or proteomic profiling analyses, rather than for unbiased

screening. Numerous studies have used TMA to validate promising results from diagnostic and prognostic biomarker discovery studies. Examples of validation of prognostic biomarkers include an association of elevated expression of COX2 [28], telomerase [29], Ep-CAM/TACSTD1 [30], and EPIL [31] with unfavorable outcome. Proteins associated with a favorable clinical outcome include PHLDA1 [32], and BCL2 [33].

To date, only a couple of studies have used TMA to evaluate multiple protein signatures comprised of putative prognostic markers proposed from individual studies. Using a supervised method of data analysis, Jacquemier *et al.* were able to identify a set of 21 proteins whose combined expression significantly correlated to metastasis-free survival and, in multivariate analysis, the 21-protein set was the strongest independent predictor of clinical outcome, irrespective of ER status [34]. Another study evaluated the predictive value of a nine-protein profile for outcome in 324 breast cancer patients treated with adjuvant tamoxifen [35]. Through the application of semiquantitative immunohistochemistry and sophisticated bioinformatics, the study confirmed the utility of BCL2, ERBB2/HER-2/neu, MYC and TP53 in predicting outcome. The latter two targets are established biomarkers for many cancer types. Whenever new biomarkers are identified in unbiased studies, the expression of those biomarkers must be clearly validated in independent studies of large cohorts, but they must also be evaluated with respect to established biomarkers in multivariate analyses. In breast cancer, this would necessarily include ER, progesterone receptor and ERBB2/HER-2/neu proteins. This strategy identifies correlations of the new biomarker with established biomarkers, and offers the best approach for identifying more accurate panels of biomarkers for clinical decision-making.

Among the tissue-based studies of breast cancer progression, some trends in functional class of protein consistently emerge. Proteins involved in translation/protein folding and degradation through ubiquitination have been highlighted, but there is also consistent detection of increases in proteins involved in iron transport such as ferritin and transferrin [12,24,25]. Ferritin appears to be important for proliferation in many different neoplasms [36], and a proposed role for transferrin signaling has been suggested in regulating the metastatic capacity of various solid tumors, including breast cancer [37,38]. Transferrin has also been shown to promote the angiogenic phenotype [39].

Accurate analysis of clinical samples is a particularly difficult task, which is complicated by multiple clinical and technical factors, but compounded by tissue heterogeneity. A solid breast tumor is, in essence, a distinct organ with a complex architecture made up of neoplastic epithelial cells, vascular endothelial cells and a supporting stromal matrix containing fibroblasts and circulating cells of the immune system. Thus, expression differences derived from tissues with categorical labels, such as tumor and nontumor, may primarily reflect varying proportions of the neoplastic and non-neoplastic components. The importance of the cell type contribution has been more fully investigated in gene expression microarray analyses [40,41]. In a study of breast cancer xenograft tissues and lymph node metastases, comparison of the metastasis-associated gene lists produced by oligonucleotide microarrays from whole and microdissected tumors showed remarkable differences, with only 1% of genes common to both techniques [41]. This problem has been largely glossed over by stating that the tissue contained over 50% tumor cells, but it remains a confounding problem when highly sensitive detection techniques are applied to tissue specimens.

Laser-capture microdissection provides a method for the extraction of distinct cell types from complex tissue specimens. Selected cells are identified through an inverted microscope and retrieved from a tissue section by laser-beam activation of a transfer film placed in contact with the tissue. In this way, epithelia can be separated from stromal tissue, and malignant cells can be separated from benign cells based on cytomorphology or by antibody

flagging protocols [42]. The successful extraction of DNA, RNA and proteins from the microdissected material has been achieved [43,44]. Limitations of the method include the yield obtainable from minimal amounts of material and the fragmentation of the target molecules through fixation and tissue processing. This fragmentation is particularly problematic for quantitative RNA analyses. Although it is a time-consuming procedure, laser capture microdissection (LCM) coupled to molecular profiling analysis allows us to understand how specific cell types contribute to the total cancer expression signature.

While many studies have used LCM on breast tissue for analysis of selected transcripts or proteins, it is only relatively recently that LCM has been coupled to high-throughput proteomic techniques. In breast tissue, LCM procurement of sample has been coupled with MALDI-TOF [45] and with nano-liquid chromatography (LC)-Fourier transform ion cyclotron resonance (FTICR)-MS. Optimization of the latter technique enabled the identification of 2282 peptides in protein digests obtained from as few as 3000 tumor cells (300 ng of protein) procured from breast carcinoma tissue [46]. Coupling LCM with 2D-GE has revealed proteomic patterns characteristic of malignant breast epithelium. LCM was used to procure epithelial cells from malignant breast epithelia and corresponding adjacent normal breast epithelia from patients with invasive breast carcinoma. Differentially expressed proteins were identified by MALDI-TOF and findings were validated by immunohistochemistry in an independent cohort of 50 breast cancer specimens [47]. With respect to aggressive breast cancer phenotype, Zhang *et al.* used 2D-GE and MALDI tandem TOF (TOF/TOF) MS/MS to identify four proteins that are associated with ERBB2/HER-2/neu expression by procuring proteins from ERBB2/HER-2/neu-positive and ERBB2/HER-2/neu-negative breast tumor specimens through LCM [48]. Thus, as with other proteomics platforms, the techniques continue to evolve and obtain high levels of coverage and accuracy. The coupling of LCM with high-throughput, in-depth proteomic analysis of clinical samples is very promising.

Histopathological evaluation of clinical specimens is based upon FFPE tissues. This procedure retains the tissue in an excellent format for histological inspection and long-term archiving. Due to its cross-linking effects, it has long been accepted that formalin fixation of routinely processed tissues in the clinic prevents protein extraction and profiling. The only routinely used method currently available for protein analysis in FFPE tissues is immunohistochemistry, which is notoriously difficult to quantify. However, investigators are now attempting to make this vast archival tissue resource available to advanced proteomics methods by developing standardized, economical and easy to use techniques for the solubilization of immunoreactive proteins from formalin-fixed tissues for western blot and protein microarray analysis. Optimized protein extraction buffer systems have allowed material from sections of routinely processed formalin-fixed and paraffin-embedded tissues to be solubilized and subsequently analyzed by western blot and reverse-phase protein microarrays [49] and MS [50]. Reported protein yields and abundance from fresh-frozen or formalin-fixed tissues are similar in some cases. If these solubilization methods can be successfully combined with LCM, then the source material available for high-throughput proteomics will exponentially expand. Most importantly for clinical prognostic studies, it is the archival material that enables the association of proteomic data with the essential, long-term clinical follow-up data.

Proteomic profiling of models of metastasis

The multistep nature of metastasis poses difficulties in both design and interpretation of experiments designed to unveil the mechanisms causing the process. Studies on excised human tissues are complicated by the variance of genetic background between individuals and by the cellular heterogeneity of a complex tissue mass. Furthermore, a major limitation

in tissues is the inability to identify those cells in a tumor mass that are truly capable of metastasis. An advanced carcinoma is a mixture of genotypically and phenotypically distinct cells, and only a tiny fraction of those cells (perhaps as few as 0.1% [51]) may possess the ability to disseminate from the primary lesion and form secondary deposits in a distal organ. While genetic studies of clinical specimens will continue to be informative, they provide only a snapshot of a complicated disease state, and there are few experimental opportunities in such analyses. Thus, the study of breast cancer progression requires experimental models for the investigation of the links between molecular profiles and a more aggressive tumor phenotype. Animal xenograft models provide a powerful resource for the identification and investigation of genes essential in determining the metastatic cellular phenotype, and enable the identification of targets that are optimal for therapeutic perturbation. While there are certainly limitations regarding extrapolation from studies in a murine host to the human clinical situation, investigators can alter the expression or activity of single or multiplexes of candidate genes in cell lines with known metastatic characteristics, and monitor which specific mechanisms are perturbed by comparison with isogenic controls. Many of these candidate genes will come from tissue-based analyses, such as those described earlier.

Several established human breast cancer cell lines with varying documented abilities of invasiveness and/or migration *in vitro* are available, and some are capable of spontaneous metastasis *in vivo* (i.e., dissemination from growth in the mammary gland and proliferation in a distal site) [52]. However, most of these are polyclonal and composed of cell populations that are heterogeneous in the metastatic phenotype, making them difficult to use as models in studies seeking to define genes causing metastasis. Several laboratories have obtained cell lines with increasing metastatic phenotype by single-cell cloning or by recycling cells through several rounds of orthotopic inoculation [53,54]. The most informative models for metastasis studies are those that are derived from a single tumor source and are composed of a series of clonal cell lines of varying phenotype, from premalignant to aggressively metastatic. Proteomic profiling approaches are now being performed on such models, a summary of which is presented later.

In order to facilitate more pertinent metastasis studies, we have utilized a pair of monoclonal tumor cell lines, M4A4 and NM2C5, derived from the breast carcinoma cell line MDA-MB-435 [55,56]. When orthotopically inoculated into athymic mice, both cell lines form primary tumors, but only M4A4 is capable of metastasis to the lungs and lymph nodes [55,57]. These isogenic cell lines constitute a stable and accessible model for the identification of genes involved in the process of tumor metastasis. We have performed multiple comparative analyses of these paired cell lines, including cytogenetic analyses and evaluation of global gene expression [55,57–59]. To further elucidate the extent of the molecular changes associated with acquisition of the metastatic phenotype in this model, we have recently employed a panel of proteomic profiling approaches to the model.

Pivotal to our proteomic studies on metastatic models has been the use of 2D liquid-phase separation strategies. As with early tissue studies, 2D-GE has been the most widely used method of separating large numbers of proteins from cell line sources. As described earlier, to overcome the limitations of 2D-GE, 2D LC/MS techniques were developed [17]. The analytical approach fractionates proteins based upon isoelectric point (pI) in the first dimension by chromatofocusing (CF), and each fraction is subsequently separated based upon polarity of the protein in the second dimension by nonporous silica (NPS) reversed-phase (RP) high-performance LC (HPLC). Purified proteins fractionated by means of the 2D liquid separation are collected and stored for further analysis by a range of orthogonal techniques. In our studies, the protein eluants from RP-HPLC were directed online to ESI-TOF-MS to obtain an accurate and reproducible molecular weight of the intact protein, and peptide mapping was achieved using MALDI-TOF-MS. The combination of such methods

enables high accuracy protein profile comparisons between biological samples, and provides information on abundance, protein isoforms and on specific post-translational modification of proteins in complex biological samples. A major advantage of liquid-based 2D separations is the ability to automate the complete workflow (Figure 1).

We have applied these methods to the profiling of first the secreted proteome, and second, the total cellular protein component of the MDA-MB-435 metastasis model [60–62]. Within the serum-free conditioned media obtained from the cell lines, we were able to detect over 400 unique, intact proteins and plot them as a 2D map of pI versus accurate molecular mass (Figure 2). ESI-TOF-MS data and analysis using MALDI-TOF MS confirmed and identified 27 differentially expressed proteins. Proteins associated with the metastatic phenotype included osteopontin and ECM-1, whereas MMP-1 and annexin I were associated with the nonmetastatic phenotype [60]. In the analysis of the more complex cell lysates, we used mass maps to select 89 proteins for digestion by trypsin and analysis by MALDI-TOF-MS, MALDI-QIT-TOF-MS and monolith-based HPLC-MS/MS. A total of 43 proteins were confirmed as differentially expressed between the two cell lines; most notable was the relative overexpression of galectin-1 in the metastatic cells and annexins I and II in the nonmetastatic phenotype [62].

An added advantage of using cell line models is the fact that they are a renewable resource. This provides ample material and multiple sampling for validation and for development and optimization of proteomic techniques. For example, in a recent study, we used the MDA-MB-435 cells to optimize the profiling of the ribosomal protein component of the proteome. There are some 80 ribosomal proteins in the human proteome, which are essential to the translation of mRNA into protein as part of the ribosome. Given the pivotal role of these proteins in the translation of genetic expression into functional phenotype, it is not surprising that recent studies have implicated ribosomal proteins, not only as targets of tumor suppressors or proto-oncogenes [63], but also as more direct mediators of aspects of tumor progression [64]. The detailed characterization of this class of proteins in mammalian cells of malignant phenotype is lacking because, due to their small size and basic nature, ribosomal proteins are particularly difficult to profile within complex biological samples. Using a combination of 2D-LC, ESI-TOF, monolithic capillary LC-MS/MS, MALDI-MS and MS/MS, we were able to identify 45 unique ribosomal proteins, several of which were differentially expressed in metastatic M4A4 cells [61]. Our proteomic studies both confirmed and added to the genomic profiling data we have obtained from the same model [65]. The differential abundance of proteins, such as osteopontin and the annexins, was detected with both approaches, although the proteomic approaches identified several unique post-translational modifications and protein isoforms associated with the metastatic phenotype. Further investigation into the functional role of specific proteins identified in these comparative proteomic studies is under way in our laboratory. We are also using MDA-MB-435 and MCF10A metastasis models to develop the coupling of 2D liquid separation methods to protein microarray technology, particularly the glycoprotein and phosphoprotein subproteomes, in order to apply them to cancer models and human tissue samples [66].

Another human cell line model of breast cancer progression, based on the MCF10A cell line, has been the subject of considerable proteomic profiling efforts. The MCF10A cell line was one of a pair of lines that spontaneously immortalized during the long-term culture of human breast epithelium [67]. MCF10A does not exhibit tumorigenicity in immunocompromised mice, but a tumorigenic derivative, MCF10AT, was obtained by transfection of MCF10A with the HRAS T24-oncogene [68], and more aggressive sublines of MCF10AT were produced by cycling through murine hosts [69]. One of these, MCF10CA1a, is also metastatic, and thus the series includes related cell lines with benign,

pre-malignant and metastatic phenotypes [70]. Through a combination of the NPS-RP-HPLC and 2D-LC methods described earlier, several studies have identified proteins associated with the malignant phenotype in this series [71–73]. The combination of CF and NPS-RP-HPLC with mass mapping revealed 110 unique proteins, 22 of which were confirmed to be differentially expressed in the malignant MCF10CA1a or MCF10CA1d cell lines relative to the pre-malignant MCF10AT line [71]. An approach that coupled RP-HPLC with a second dimension using SDS-PAGE enabled the analysis of the membrane proteins within the microsomal fractions of the different stages of neoplastic progression. Differentially expressed proteins were identified using MALDI-MS and MALDI-qTOF-MS/MS peptide sequencing [74]. A more comprehensive proteomic characterization of the MCF10 model was recently performed in which ESI-TOF-MS was used to resolve and detect intact proteins. Of 133 identified proteins, 67 were shown to have expression consistently associated with the malignant phenotype [73]. The findings indicated significant changes to the cytoskeleton, cellular metabolism and adaptation to the environment in malignant MCF10 cell lines. A number of proteins known to be involved in cancer (including TP53, MYC, SRC and ERBB2/HER-2/neu) appear to be increasingly expressed in the tumorigenic and metastatic counterparts. Complete lists of identified and differentially expressed proteins can be found in each publication, but we have compiled a list of proteins that have been found to be expressed relative to the malignant, metastatic phenotype in more than one model, or in both human tissue studies and model studies (Table 1).

Expert commentary

High-throughput profiling approaches of the cellular transcriptome and proteome have followed technologically separate paths, and perhaps genomics has something of a lead at this time. However, it is the combination of data from both of these sources that will yield the greatest advances in our understanding of cellular function, and ultimately disease pathology. DNA-based analyses can detect mutations, polymorphisms and gross chromosomal aberrations in the human genome. Transcriptomics provides information on relative mRNA abundance and alternative splicing patterns, but proteomic profiling enables the evaluation of global changes in gene expression that result from both transcriptional and post-transcriptional processing of mRNA, as well as translation and post-translation modifications. Given that phenotypic changes can only manifest themselves through altered protein expression, it is clearly preferable to profile this component whenever possible. Furthermore, the ability to identify protein factors involved in the progression of disease will most readily lead to the development of biomarkers that have clinical utility.

Proteins revealed in the analysis of multiple models, and those that are part of any overlap between models and tissue-based studies, will be a logical focus for functional investigation. As described earlier, relevant models for such investigations into breast cancer progression do exist, but it would be beneficial to develop more such models in order to identify those genes and pathways that are implicated across a number of platforms, and are thus likely to be pivotal to metastatic efficiency.

The advent of high-throughput proteomic and genomic technology has spawned a whole new field of bioinformatics (reviewed in [75,76]). However, the majority of proteomic and genomic profiling studies to date have attempted to develop genetic marker-based prognostic systems that might replace the existing clinical criteria, rather than incorporating the valuable clinical information contained in established clinical markers. Given the complexity of breast cancer prognosis, a more promising strategy may be to combine both clinical and genetic marker information that may be complementary. In order to address this, we have recently performed computational studies using genomic prognosis signatures and associated clinical information. Through the combination of both molecular and clinical

markers, the application of our novel I-RELIEF algorithm to breast cancer profiles has identified hybrid signatures that perform significantly better than the molecular signature or clinical criteria alone in the prediction of breast disease recurrence [5]. Ongoing computational developments that enable interstudy comparison and the incorporation of distinct forms of data, including proteomic, genomic and clinical data, will provide platforms for the refinement of cancer-related molecular signatures and lead to more accurate prognostic systems that may facilitate personalized patient evaluation and treatment decisions. These integrated analyses also have the potential to highlight pivotal genes and pathways that are likely to be a part of the biological driving force of metastasis.

Five-year view

A survey of the breast cancer literature reveals the promise of proteomics in unraveling the molecular complexities associated with breast carcinoma. The next few years will see the continued rapid evolution of high-throughput proteomic technologies, including the development of special approaches for the comprehensive analysis of complex biological samples and distinct classes of proteins. If proteomic profiling techniques can approach the coverage currently achieved in the genomics field, proteomics will become an essential tool in the investigation of the pathology of human disease. The field of bioinformatics and integrative systems biology will evolve to optimize the use of the vast amounts of data that are accumulating and, encouragingly, that are being made publicly available through management of standardized databases by the National Cancer Institute (NCI) and others. It is hoped that advanced computational systems will be able to mine these data and reveal patterns of expression that are associated with disease processes, including breast cancer metastasis. The combination of data from multiple sources offers the best chance for identifying biomarkers for early detection, diagnosis and prognosis, and for revealing the most promising therapeutic targets.

Perhaps the most confounding obstacle to accurate protein expression profiling in human tumor specimens is the degree of cellular heterogeneity. The unknown variance that this heterogeneity introduces to global profiling analyses can seriously affect the value of comparative proteomics, and this phenomenon is likely to be the major reason why breast cancer prognosis has not progressed beyond the use of histological criteria. LCM can largely overcome this problem, but the routine application of this tissue processing technique for high-throughput proteomics will require the development of approaches that can use the minimal amounts of preserved material that are available via LCM.

References

Papers of special note have been highlighted as:

- of interest
- of considerable interest

1. Albertson DG. Profiling breast cancer by array CGH. *Breast Cancer Res Treat.* 2003; 78(3):289–298. [PubMed: 12755488]
2. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA.* 2003; 100(14):8418–8423. [PubMed: 12829800]
3. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med.* 2002; 347(25):1999–2009. [PubMed: 12490681]

4. Weigelt B, Hu Z, He X, et al. Molecular portraits and 70-gene prognosis signature are preserved throughout the metastatic process of breast cancer. *Cancer Res.* 2005; 65(20):9155–9158. [PubMed: 16230372]
5. Sun Y, Goodison S, Li J, Liu L, Farmerie W. Improved breast cancer prognosis through the combination of clinical and genetic markers. *Bioinformatics.* 2007; 23(1):30–37. [PubMed: 17130137]
6. Bogaerts J, Cardoso F, Buyse M, et al. Gene signature evaluation as a prognostic tool: challenges in the design of the MINDACT trial. *Nat Clin Pract Oncol.* 2006; 3(10):540–551. [PubMed: 17019432]
7. O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. *J Biol Chem.* 1975; 250(10):4007–4021. [PubMed: 236308]
8. Wulfschlegel JD, Sgroi DC, Krutzsch H, et al. Proteomics of human breast ductal carcinoma *in situ*. *Cancer Res.* 2002; 62(22):6740–6749. [PubMed: 12438275]
9. Vercoutter-Edouart AS, Lemoine J, Le Bourhis X, et al. Proteomic analysis reveals that 14-3-3 σ is down-regulated in human breast cancer cells. *Cancer Res.* 2001; 61(1):76–80. [PubMed: 11196201]
10. Hondermarck H, Vercoutter-Edouart AS, Revillion F, et al. Proteomics of breast cancer for marker discovery and signal pathway profiling. *Proteomics.* 2001; 1(10):1216–1232. [PubMed: 11721634]
11. Huang HL, Stasyk T, Morandell S, et al. Biomarker discovery in breast cancer serum using 2-D differential gel electrophoresis/MALDI-TOF/TOF and data validation by routine clinical assays. *Electrophoresis.* 2006; 27(8):1641–1650. [PubMed: 16550499]
12. Nimeus E, Malmstrom J, Johnsson A, Marko-Varga G, Ferno M. Proteomic analysis identifies candidate proteins associated with distant recurrences in breast cancer after adjuvant chemotherapy. *J Pharm Biomed Anal.* 2007; 43(3):1086–1093. [PubMed: 17085005]
13. Eden P, Ritz C, Rose C, Ferno M, Peterson C. “Good old” clinical markers have similar power in breast cancer prognosis as microarray gene expression profilers. *Eur J Cancer.* 2004; 40(12):1837–1841. [PubMed: 15288284]
14. Unlu M, Morgan ME, Minden JS. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis.* 1997; 18(11):2071–2077. [PubMed: 9420172]
15. Somiari RI, Somiari S, Russell S, Shriver CD. Proteomics of breast carcinoma. *J Chromatogr B Analyt Technol Biomed Life Sci.* 2005; 815(1–2):215–225.
16. Schwartz SA, Weil RJ, Thompson RC, et al. Proteomic-based prognosis of brain tumor patients using direct-tissue matrix-assisted laser desorption ionization mass spectrometry. *Cancer Res.* 2005; 65(17):7674–7681. [PubMed: 16140934]
17. Wall DB, Kachman MT, Gong S, et al. Isoelectric focusing nonporous RP HPLC: a two-dimensional liquid-phase separation method for mapping of cellular proteins with identification using MALDI-TOF mass spectrometry. *Anal Chem.* 2000; 72(6):1099–1111. [PubMed: 10740846]
18. Villanueva J, Shaffer DR, Philip J, et al. Differential exoprotease activities confer tumor-specific serum peptidome patterns. *J Clin Invest.* 2006; 116(1):271–284. [PubMed: 16395409]
19. Li J, Zhang Z, Rosenzweig J, Wang YY, Chan DW. Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. *Clin Chem.* 2002; 48(8):1296–1304. [PubMed: 12142387]
20. Vlahou A, Laronga C, Wilson L, et al. A novel approach toward development of a rapid blood test for breast cancer. *Clin Breast Cancer.* 2003; 4(3):203–209. [PubMed: 14499014]
21. Hu Y, Zhang S, Yu J, Liu J, Zheng S. SELDI-TOF-MS: the proteomics and bioinformatics approaches in the diagnosis of breast cancer. *Breast.* 2005; 14(4):250–255. [PubMed: 16085230]
22. Mendrinos S, Nolen JD, Styblo T, et al. 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.* 2005; 105(3):178–183. [PubMed: 15822128]
23. Fowler LJ, Lovell MO, Izbicka E. Fine-needle aspiration in PreservCyt: a novel and reproducible method for possible ancillary proteomic pattern expression of breast neoplasms by SELDI-TOF. *Mod Pathol.* 2004; 17(8):1012–1020. [PubMed: 15195106]

24. Ricolleau G, Charbonnel C, Lode L, et al. 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*. 2006; 6(6):1963–1975. [PubMed: 16470659]
25. Goncalves A, Esterni B, Bertucci F, et al. Postoperative serum proteomic profiles may predict metastatic relapse in high-risk primary breast cancer patients receiving adjuvant chemotherapy. *Oncogene*. 2006; 25(7):981–989. [PubMed: 16186794]
26. Heike Y, Hosokawa M, Osumi S, et al. Identification of serum proteins related to adverse effects induced by docetaxel infusion from protein expression profiles of serum using SELDI ProteinChip system. *Anticancer Res*. 2005; 25(2B):1197–1203. [PubMed: 15865066]
27. Diamandis EP. Mass spectrometry as a diagnostic and a cancer biomarker discovery tool: opportunities and potential limitations. *Mol Cell Proteomics*. 2004; 3(4):367–378. [PubMed: 14990683]
28. Zerkowski MP, Camp RL, Burtneis BA, Rimm DL, Chung GG. Quantitative analysis of breast cancer tissue microarrays shows high cox-2 expression is associated with poor outcome. *Cancer Invest*. 2007; 25(1):19–26. [PubMed: 17364553]
29. Poremba C, Heine B, Diallo R, et al. Telomerase as a prognostic marker in breast cancer: high-throughput tissue microarray analysis of hTERT and hTR. *J Pathol*. 2002; 198(2):181–189. [PubMed: 12237877]
30. Spizzo G, Went P, Dirnhofer S, et al. High Ep-CAM expression is associated with poor prognosis in node-positive breast cancer. *Breast Cancer Res Treat*. 2004; 86(3):207–213. [PubMed: 15567937]
31. Brandt B, Kemming D, Packeisen J, et al. Expression of early placenta insulin-like growth factor in breast cancer cells provides an autocrine loop that predominantly enhances invasiveness and motility. *Endocr Relat Cancer*. 2005; 12(4):823–837. [PubMed: 16322324]
32. Nagai MA, Fregnani JH, Netto MM, Brentani MM, Soares FA. Down-regulation of PHLDA1 gene expression is associated with breast cancer progression. *Breast Cancer Res Treat*. 2007; 106(1):49–56. [PubMed: 17211533]
33. Callagy GM, Pharoah PD, Pinder SE, et al. Bcl-2 is a prognostic marker in breast cancer independently of the Nottingham Prognostic Index. *Clin Cancer Res*. 2006; 12(8):2468–2475. [PubMed: 16638854]
34. Jacquemier J, Ginestier C, Rougemont J, et al. Protein expression profiling identifies subclasses of breast cancer and predicts prognosis. *Cancer Res*. 2005; 65(3):767–779. [PubMed: 15705873]
35. Linke SP, Bremer TM, Herold CD, Sauter G, Diamond C. A multimarker model to predict outcome in tamoxifen-treated breast cancer patients. *Clin Cancer Res*. 2006; 12(4):1175–1183. [PubMed: 16489071]
36. Kwok JC, Richardson DR. The iron metabolism of neoplastic cells: alterations that facilitate proliferation? *Crit Rev Oncol Hematol*. 2002; 42(1):65–78. [PubMed: 11923069]
37. Cavanaugh PG, Nicolson GL. Selection of highly metastatic rat MTLn2 mammary adenocarcinoma cell variants using *in vitro* growth response to transferrin. *J Cell Physiol*. 1998; 174(1):48–57. [PubMed: 9397155]
38. Cavanaugh PG, Jia L, Zou Y, Nicolson GL. Transferrin receptor overexpression enhances transferrin responsiveness and the metastatic growth of a rat mammary adenocarcinoma cell line. *Breast Cancer Res Treat*. 1999; 56(3):203–217. [PubMed: 10573112]
39. Carlevaro MF, Albini A, Ribatti D, et al. Transferrin promotes endothelial cell migration and invasion: implication in cartilage neovascularization. *J Cell Biol*. 1997; 136(6):1375–1384. [PubMed: 9087450]
40. Stuart RO, Wachsman W, Berry CC, et al. *In silico* dissection of cell-type-associated patterns of gene expression in prostate cancer. *Proc Natl Acad Sci USA*. 2004; 101(2):615–620. [PubMed: 14722351]
41. Harrell JC, Dye WW, Harvell DM, Sartorius CA, Horwitz KB. Contaminating cells alter gene signatures in whole organ versus laser capture microdissected tumors: a comparison of experimental breast cancers and their lymph node metastases. *Clin Exp Metastasis*. 2008; 25(1):81–88. [PubMed: 17932773]

42. Rupp C, Dolznig H, Puri C, et al. Laser capture microdissection of epithelial cancers guided by antibodies against fibroblast activation protein and endosialin. *Diagn Mol Pathol*. 2006; 15(1):35–42. [PubMed: 16531767]
43. Wang-Rodriguez J, Urquidi V, Rivard A, Goodison S. Elevated osteopontin and thrombospondin expression identifies malignant human breast carcinoma but is not indicative of metastatic status. *Breast Cancer Res*. 2003; 5(5):R136–R143. [PubMed: 12927044]
44. Mustafa D, Kros JM, Luider T. Combining laser capture microdissection and proteomics techniques. *Methods Mol Biol*. 2007; 428:159–178. [PubMed: 18287773]
45. Umar A, Dalebout JC, Timmermans AM, Foekens JA, Luider TM. 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*. 2005; 5(10):2680–2688. [PubMed: 15892168]
46. Umar A, Luider TM, Foekens JA, Pasa-Tolic L. NanoLC-FT-ICR MS improves proteome coverage attainable for approximately 3000 laser-microdissected breast carcinoma cells. *Proteomics*. 2007; 7(2):323–329. [PubMed: 17163580]
47. Hudelist G, Singer CF, Pischinger KI, et al. Proteomic analysis in human breast cancer: identification of a characteristic protein expression profile of malignant breast epithelium. *Proteomics*. 2006; 6(6):1989–2002. [PubMed: 16470630]
48. Zhang D, Tai LK, Wong LL, et al. Proteomic study reveals that proteins involved in metabolic and detoxification pathways are highly expressed in HER-2/neu-positive breast cancer. *Mol Cell Proteomics*. 2005; 4(11):1686–1696. [PubMed: 16048908]
49. Becker KF, Metzger V, Hipp S, Hofler H. Clinical proteomics: new trends for protein microarrays. *Curr Med Chem*. 2006; 13(15):1831–1837. [PubMed: 16787224]
50. Chu WS, Liang Q, Liu J, et al. A nondestructive molecule extraction method allowing morphological and molecular analyses using a single tissue section. *Lab Invest*. 2005; 85(11):1416–1428. [PubMed: 16127423]
51. Fidler IJ, Kim SJ, Langley RR. The role of the organ microenvironment in the biology and therapy of cancer metastasis. *J Cell Biochem*. 2007; 101(4):927–936. [PubMed: 17177290]
52. Price JE, Polyzos A, Zhang RD, Daniels LM. Tumorigenicity and metastasis of human breast carcinoma cell lines in nude mice. *Cancer Res*. 1990; 50(3):717–721. [PubMed: 2297709]
53. Bruns CJ, Harbison MT, Kuniyasu H, Eue I, Fidler IJ. *In vivo* selection and characterization of metastatic variants from human pancreatic adenocarcinoma by using orthotopic implantation in nude mice. *Neoplasia*. 1999; 1(1):50–62. [PubMed: 10935470]
54. Gehlsen KR, Davis GE, Sriramaraio P. Integrin expression in human melanoma cells with differing invasive and metastatic properties. *Clin Exp Metastasis*. 1992; 10(2):111–120. [PubMed: 1311225]
55. Urquidi V, Sloan D, Kawai K, et al. Contrasting expression of thrombospondin-1 and osteopontin correlates with absence or presence of metastatic phenotype in an isogenic model of spontaneous human breast cancer metastasis. *Clin Cancer Res*. 2002; 8(1):61–74. [PubMed: 11801541]
56. Cailleau R, Olive M, Cruciger QV. Long-term human breast carcinoma cell lines of metastatic origin: preliminary characterization. *In Vitro*. 1978; 14(11):911–915. [PubMed: 730202]
57. Goodison S, Kawai K, Hihara J, et al. Prolonged dormancy and site-specific growth potential of cancer cells spontaneously disseminated from nonmetastatic breast tumors as revealed by labeling with green fluorescent protein. *Clin Cancer Res*. 2003; 9(10 Pt 1):3808–3814. [PubMed: 14506175]
58. Agarwal D, Goodison S, Nicholson B, Tarin D, Urquidi V. Expression of matrix metalloproteinase 8 (MMP-8) and tyrosinase-related protein-1 (TYRP-1) correlates with the absence of metastasis in an isogenic human breast cancer model. *Differentiation*. 2003; 71(2):114–125. [PubMed: 12641565]
59. Goodison S, Viars C, Urquidi V. Molecular cytogenetic analysis of a human breast metastasis model: identification of phenotype-specific chromosomal rearrangements. *Cancer Genet Cytogenet*. 2005; 156(1):37–48. [PubMed: 15588854]

60. Kreunin P, Urquidi V, Lubman DM, Goodison S. Identification of metastasis-associated proteins in a human tumor metastasis model using the mass-mapping technique. *Proteomics*. 2004; 4(9): 2754–2765. [PubMed: 15352249]
61. Kreunin P, Yoo C, Urquidi V, Lubman DM, Goodison S. Differential expression of ribosomal proteins in a human metastasis model identified by coupling 2-D liquid chromatography and mass spectrometry. *Cancer Genomics Proteomics*. 2007; 4(5):329–339. [PubMed: 17993717]
62. Kreunin P, Yoo C, Urquidi V, Lubman DM, Goodison S. Proteomic profiling identifies breast tumor metastasis-associated factors in an isogenic model. *Proteomics*. 2007; 7(2):299–312. [PubMed: 17205601]
63. Ruggero D, Pandolfi PP. Does the ribosome translate cancer? *Nat Rev Cancer*. 2003; 3(3):179–192. [PubMed: 12612653]
64. Holland EC. Regulation of translation and cancer. *Cell Cycle*. 2004; 3(4):452–455. [PubMed: 14976427]
65. Goodison S, Yuan J, Sloan D, et al. The RhoGAP protein DLC-1 functions as a metastasis suppressor in breast cancer cells. *Cancer Res*. 2005; 65(14):6042–6053. [PubMed: 16024604]
66. Kreunin P, Zhao J, Rosser C, et al. Bladder cancer associated glycoprotein signatures revealed by urinary proteomic profiling. *J Proteome Res*. 2007; 6(7):2631–2639. [PubMed: 17518487]
67. Soule HD, Maloney TM, Wolman SR, et al. Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, MCF-10. *Cancer Res*. 1990; 50(18):6075–6086. [PubMed: 1975513]
68. Basolo F, Elliott J, Tait L, et al. Transformation of human breast epithelial cells by c-Ha-ras oncogene. *Mol Carcinog*. 1991; 4(1):25–35. [PubMed: 2009132]
69. Santner SJ, Dawson PJ, Tait L, et al. Malignant MCF10CA1 cell lines derived from premalignant human breast epithelial MCF10AT cells. *Breast Cancer Res Treat*. 2001; 65(2):101–110. [PubMed: 11261825]
70. Miller FR, Soule HD, Tait L, et al. Xenograft model of progressive human proliferative breast disease. *J Natl Cancer Inst*. 1993; 85(21):1725–1732. [PubMed: 8411256]
71. Hamler RL, Zhu K, Buchanan NS, et al. A two-dimensional liquid-phase separation method coupled with mass spectrometry for proteomic studies of breast cancer and biomarker identification. *Proteomics*. 2004; 4(3):562–577. [PubMed: 14997480]
72. Chong BE, Yan F, Lubman DM, Miller FR. Chromatofocusing nonporous reversed-phase high-performance liquid chromatography/electrospray ionization time-of-flight mass spectrometry of proteins from human breast cancer whole cell lysates: a novel two-dimensional liquid chromatography/mass spectrometry method. *Rapid Commun Mass Spectrom*. 2001; 15(4):291–296. [PubMed: 11223961]
73. Buchanan NS, Zhao J, Zhu K, et al. Differential expression of acidic proteins with progression in the MCF10 model of human breast disease. *Int J Oncol*. 2007; 31(4):941–949. [PubMed: 17786328]
74. O’Neil KA, Miller FR, Barder TJ, Lubman DM. Profiling the progression of cancer: separation of microsomal proteins in MCF10 breast epithelial cell lines using nonporous chromatophoresis. *Proteomics*. 2003; 3(7):1256–1269. [PubMed: 12872226]
75. Tan K, Tegner J, Ravasi T. Integrated approaches to uncovering transcription regulatory networks in mammalian cells. *Genomics*. 2008; 91(3):219–231. [PubMed: 18191937]
76. Lin J, Qian J. Systems biology approach to integrative comparative genomics. *Expert Rev Proteomics*. 2007; 4(1):107–119. [PubMed: 17288519]

Key issues

- Highly sensitive and specific biomarkers for breast cancer progression are urgently needed. It is expected that compound molecular signatures, rather than single biomarkers, will achieve the desired accuracy.
- The combination of data from both tissue-based studies and from the analysis of appropriate models will identify biomarkers with potential clinical utility, and will provide the information to better understand the biology of breast cancer progression.
- The continued evolution of proteomic methods (and the concomitant reduction in costs) will enable the accumulation of large amounts of data associated with human disease.
- Mass spectrometry is at the core of proteomics, but coupling with multiple separation and downstream analytical techniques is required for the comprehensive profiling of the proteome.
- The availability of proteomic data in managed, standardized databases is essential.
- The integration of data from proteomics with multiple sources of global profile and target-specific information will provide better information for the modeling of disease progression than any single discipline alone.

Figure 1. Experimental overview of the liquid-phase 2D separation and downstream techniques that we have combined for the proteomic analysis of models of metastasis

HPLC: High-performance liquid chromatography; MS: Mass spectrometry; MS/MS: Tandem MS; NPS: Nonporous silica; PMF: Peptide mass fingerprinting; RP: Reversed phase.

Figure 2. A representative mass map of the secreted proteome from the metastatic M4A4 tumor cell line

The relative intensities of the band are quantitatively proportional to the amount of corresponding protein detected by UV absorption.

Table 1

Breast cancer metastasis-associated proteins identified in multiple proteomic studies of human tissues and/or xenograft models.

Protein	MCF10 series	MDA-MB-435 series	Human tissues	Ref.
Thioredoxin domain-containing protein-5	+		+	[12,73]
Pyruvate kinase, M1/M2 isozymes	+	+		[62,71,73]
Annexin II	+	+		[62,71]
Heat-shock protein-70 family	+	+		[62,71,73]
Ubiquitin	+		+	[24,71]
Ubiquitin-conjugating enzyme E2	+	+		[62,73]
Galectin-1	+	+		[62,71,73]
Peroxiredoxin-2	+	+		[62,71]
Heat-shock protein-27 family	+	+		[62,71]
α -enolase	+	+		[60,71]
Vimentin	+		+	[12,71]