Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Expert Rev Proteomics. Author manuscript; available in PMC 2019 February 01.

Published in final edited form as:

Expert Rev Proteomics. 2018 February ; 15(2): 131–152. doi:10.1080/14789450.2018.1421071.

Protein biomarkers for subtyping breast cancer and implications for future research

Claudius Mueller, **Amanda Haymond**, **Justin B. Davis**, **Alexa Williams**, and **Virginia Espina**

Center for Applied Proteomics and Molecular Medicine, George Mason University, Manassas, Virginia, USA

Abstract

Introduction: Breast cancer subtypes are currently defined by a combination of morphologic, genomic, and proteomic characteristics. These subtypes provide a molecular portrait of the tumor, that aids diagnosis, prognosis, and treatment escalation/de-escalation options. Gene expression signatures describing intrinsic breast cancer subtypes for predicting risk of recurrence have been rapidly adopted in the clinic. Despite the use of subtype classifications, many patients develop drug resistance, breast cancer recurrence, or therapy failure.

Areas covered: This review provides a summary of immunohistochemistry, reverse phase protein array, mass spectrometry, and integrative studies that are revealing differences in biological functions within and between breast cancer subtypes. We conclude with a discussion of rigor and reproducibility for proteomic-based biomarker discovery.

Expert commentary: Innovations in proteomics, including implementation of assay guidelines and standards, are facilitating refinement of breast cancer subtypes. Proteomic and phosphoproteomic information distinguish biologically functional subtypes, are predictive of recurrence, and indicate likelihood of drug resistance. Actionable, activated signal transduction pathways can now be quantified and characterized. Proteomic biomarker validation in large, well-designed studies should become a public health priority to capitalize on the wealth of information gleaned from the proteome.

Keywords

basal-like; biomarker; breast cancer; Estrogen Receptor; HER2; mass spectrometry; Progesterone Receptor; reverse phase protein array; signal transduction; triple negative breast cancer

Correspondence: Virginia Espina, George Mason University, Center for Applied Proteomics and Molecular Medicine, 10920 George, Mason Circle, MS1A9, Manassas, VA 20110, vespina@gmu.edu, Phone: 703-993-8062, Fax: 703-993-8606, Website: http://capmm.gmu.edu/.

Declaration of interest

C. Mueller, V. Espina and A. Haymond all receive salary support from the National Institutes of Health grant 1R33CA20693701. V Espina is also an inventor of protein painting technology and, as a university employee, may receive patent royalties per university policies. The authors have no other 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 apart from those disclosed. Peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

1. Introduction

Breast cancer is classified into subtypes to aid in diagnosis, prognosis, and treatment escalation/de-escalation options. Breast cancer subtype designations are based on clinical data, proteomic and genomic characteristics, and histomorphology. Subtype designations are clinically useful because breast cancer exhibits intra- and inter-patient tumor heterogeneity. Heterogeneity manifests itself in several biologically important forms: as variation in the proportion of cellular components within the tumor microenvironment, as spatial and temporal differences in biomarker expression, as tumor clonal populations, and as patient clinical variables (age, race, lymph node and menopausal status). Heterogeneity is the underlying reason that breast cancers possess different clinical behaviors and biological functions [1–5]. Based on histomorphology and growth patterns alone, 21 histological types of breast cancer have been defined by the World Health Organization [2]. Two broad categories of breast cancer are *in situ* carcinoma and invasive carcinoma. Ductal carcinoma in situ (DCIS) and lobular carcinoma in situ (LCIS) are differentiated by growth patterns and cytological features, and DCIS is further characterized by tumor architecture [3]. Invasive carcinoma histological subtypes are designated by their architecture, secretion (mucinous/colloid), or structural form (medullary, tubular, papillary) [2,3]. Infiltrating ductal carcinoma (IDC) is classified into tumor grades (well, moderately or poorly differentiated) based on mitotic index, tubule formation, and nuclear polymorphisms, further aiding prognosis [3]. Infiltrating ductal carcinoma accounts for 70–80% of female invasive breast tumors and represents the majority of breast cancer cases in The Cancer Genome Atlas (TCGA) [6–8] and the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) cohorts [9,10]. Other rare subtypes exist which are reviewed elsewhere [2,11].

1.1 Historical classification of breast tumors and breast cancer subtypes

Complex breast biology underscores the need for biomarkers that can differentiate indolent from aggressive growth and foretell treatment response. Breast tumor biology has historically been classified based on immunohistochemical (IHC) staining of proliferation proteins (Ki-67), hormone receptor status (estrogen receptor alpha (ER), progesterone receptor (PR) and/or androgen receptor (AR)), and the presence/absence of specific cytokeratins (CK) [11–17]. Ki- 67 expression is inversely correlated with outcome: high Ki-67 proliferation index correlates with poor outcome [18]. The clinical use of Ki-67 is controversial due to reported poor inter-laboratory reproducibility of Ki-67 assays, differences in thresholds for low and high proliferation indices, and differences in assay methods [18]. Despite these limitations, Ki-67 expression has demonstrated clinical prognostic value at the low and high thresholds.

ER and PR status predict endocrine therapy sensitivity. Two types of estrogen receptor exist, ERα and ERβ. An underappreciated fact is that both ERα and ERβ are biologically functional. The complex biology begins with ER cross-talk between epidermal growth factor receptor (EGFR or HER1) and HER2 [19]. ERα functions as a ligand (estrogen) dependent receptor for promoting cell proliferation, while ERβ can antagonize ERα [20]. Phosphorylation of nuclear ERα on Ser305 causes cyclin D1 transcription and phosphorylation on Ser118 and Ser167 results in increased transcriptional activity [19].

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Selective estrogen receptor modulators (tamoxifen and raloxifene) exhibit both ER agonist and antagonist activity [19]. A limitation of current estrogen receptor prognostic clinical evaluation is that only ERα has been validated for clinical value. ERβ is not measured despite availability of validated antibodies [21]. Therefore, except where noted in this review, ER refers only to ERα.

ER and PR are transcriptional regulators belonging to the nuclear receptor superfamily that includes receptors for steroid and thyroid hormones, vitamin D, and peroxisome proliferator-activated receptors [22]. ER and PR positive tumors are associated with a favorable prognosis [18] compared to a poor prognosis for patients with ER negative breast tumors [23]. ER/PR IHC scores consider both the percentage of positive cells and the staining intensity (Allred score) [12,15,24]. Currently the ER positivity threshold is 1% tumor cells [25]. ER positive tumors indicate a greater likelihood of response to endocrine therapy (estrogen receptor modulators, aromatase inhibitors, or estrogen receptor inhibitors), however treatment response is not uniform and treatment resistance develops in a subset of ER+ patients [16,18].Histomorphology and ER and PR status represent the main clinicopathological classification scheme for guiding endocrine therapy (Figure 1) [26].

Growth factor receptors and their ligands regulate cell proliferation, which is a key regulator of oncogenesis [27]. Growth factor receptors include EGFR, HER2, HER3 and HER4, platelet derived growth factor (PDGFβ), and insulin like growth factor 1 receptor (IGF1-R) [28]. These growth factor receptors regulate cellular proliferation via tyrosine kinase activity of plasma membrane proteins, with extensive dimerization and receptor cross-talk [28]. The discovery of the HER2 gene [27,29], its presence in a breast cancer cell line [30], and recognition of the role of HER2 in driving human breast tumor growth [31], transformed the subtype classification of breast cancer. HER2 overexpression is associated with aggressive tumors, poor prognosis, and response to chemotherapy [4,28,32–36]. HER2 testing is clinically validated to predict response to anti-HER2 therapies in neoadjuvant and adjuvant settings [18,36,37] (Figure 1).

1.2 Intrinsic gene expression signatures

Histopathologic and immunohistochemical classification of breast tumors undoubtedly provides diagnostic and prognostic value [38,39]. However, a discussion of breast cancer subtypes would be incomplete without considering the influence of gene expression signatures as subtype classifiers. More than 15 years ago, transcriptional profiling of breast tumors revealed a gene set whose expression varied significantly between tumors and this variation was not due to tissue sampling bias [26,40,41]. Hierarchical cluster analysis of this intrinsic 500 gene set revealed 5 gene expression profiles, which were labeled as luminal A, luminal B, basal-like, HER2+, and normal breast like (Figure 1) [26,40]. The luminal, basal, and normal subtype labels were derived from breast histomorphology. Luminal subtypes are associated with epithelial cells surrounding the duct lumen. Basal-like subtypes are associated with myoepithelial cells lining the basement membrane (basal layer of the duct) [42]. Normal-like signifies normal breast tissue, likely due to an abundance of adjacent normal breast tissue in heterogeneous tumor specimens. The HER2+ subtype represents amplification of the HER2neu gene and overexpression of the corresponding HER2 receptor.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

The intrinsic subtypes were shown to correlate with overall survival, prognosis, and therapy response [26,43]. In 2008, Parker et al. proposed that recurrence risk could be further defined using a 50 gene subset of the original intrinsic genes plus a classification algorithm, the Prediction Analysis for Microarrays (PAM50) [44]. PAM50 was rapidly adopted in the clinical setting for predicting risk of recurrence in ER+/PR+/HER2neg, lymph node-negative (stage I or II) or node-positive (stage II) patients [37]. Due to the widespread adoption of the 'intrinsic subtype' classification, many subsequent genomic and proteomic studies refer to this nomenclature.

New algorithms and subtype classification systems continue to be developed and refined. Using ensemble learning rather than a single sample predictor model, Milioli et al. developed a novel gene expression score to rank features between two different classes using a supervised univariate method [10]. Using this scoring system, they identified seven under-explored genes that could potentially aid prognosis or predictive value [10]. The St. Gallen 5 tier classification scheme removed the normal-like subtype and stratified luminal B and HER2 subtypes into luminal B HER2+, luminal B HER2 negative, and HER2 non-luminal [4]. Furthermore, enhancements to intrinsic subtype classifications have been proposed because factors other than tumor biology impact prognosis and risk of recurrence [45]. Age [1,45–47], race [46,48,49], metastatic site [46,50–53] and immune cell interactions [54–57] were shown to impact risk of recurrence, survival, or treatment options.

1.3 Male breast cancer

Molecular heterogeneity also exists between male and female breast cancers. Male breast cancer represents a mere 1% of the diagnosed breast cancer cases [58]. Male breast cancers are more likely to carry *BRCA2* mutations and less likely to carry *PIK3CA* and *TP53* mutations compared to females with the same ER+/HER2- phenotype [58,59]. Male breast cancers are predominately ER+ and AR+, and are rarely triple negative (ER, PR, and HER2 negative) [58]. Male breast cancers exhibited frequent expression of apoptotic (Bcl-2) and growth factor proteins (IGF1-R) compared to female breast cancers [58,60]. Using tissue microarrays and immunohistochemical stains, male breast cancers were found to cluster into four groups: (1) hormone receptor negative, (2) ER+ high grade, (3) ER+ intermediate grade, and (4) ER+ low grade [60]. The ER+ high and intermediate grade groups could be further stratified by BRST2 (gross cystic disease fluid protein 15 (GCDFP-15)) expression, a marker of apocrine differentiation. Classification into 4 subtypes based on ER expression clearly shows that the distribution of subtypes in male breast cancer is different compared to female breast cancers [61]. Due to the low incidence and limited literature regarding proteomic subtypes of male breast cancer, the remainder of this review focuses solely on female breast cancer subtypes.

1.4 Clonal heterogeneity

Breast cancer subtyping on the molecular level is now possible due to the convergence of technology, bioinformatics, and drug discovery. Categorizing biological features of breast tumors into subtypes provides clinically useful information to guide diagnosis, prognosis, therapy, and estimate recurrence risk [4,26,33,37,62]. Molecular classification of breast tumors currently relies on a suite of genomic and proteomic analysis, including customary

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

histomorphology, IHC, and clinical data. However, even the newly proposed integrative gene cluster taxonomy fails to consider the complex positive and negative feedback loops in signal transduction pathways that reflect the underlying breast tumor clonal heterogeneity [63]. Intra tumor clonal heterogeneity is due to active maintenance of more than one tumor clone population [64,65]. Clonal heterogeneity is proposed to ensure survival of the entire tumor, even if one clone is killed during treatment or if it moves to a metastatic niche [26,66] (Figure 2). The spatial features of intra-tumoral heterogeneity can be appreciated when one considers the simple fact that Ki-67 is often highly expressed at the tumor periphery [67]. Temporal accumulation of common oncogenic mutations in *TP53, PIK3CA, PTEN, HER2,* F*G*F*R1,* and *MYC* genes in triple negative breast cancer have been shown to occur early in breast cancer development, whereas later developing clones may possess different mutations [67].

To improve and refine breast cancer subtyping for precision medicine, we must elucidate the 'omics' of the tumor <u>and</u> its microenvironment [68]. Insights into the genome, epigenome, proteome, plus metabolome, will transform our understanding of individual tumor clones and overall tumor behavior [69]. In this review, we emphasize the contribution of protein biomarkers for classifying breast cancer subtypes based on functional phenotypes (Table 1). The plethora of literature on breast cancer subtypes limits our ability to summarize all the current papers. We have selected recent, representative studies, some which integrate multi-omic methods, and some that include *de novo* protein biomarker identification, to highlight progress within the proteomic biomarker realm.

2. Protein biomarkers associated with breast cancer subtypes

Proteomic assays have rapidly matured beyond total numbers of proteins identified and lists of proteins and peptides identified in a specimen. We can now identify and quantify low abundance proteins, characterize and quantify their post-translational modifications, validate protein identity with antibody and antibody-free methods, and pinpoint potential therapeutic protein targets. Protein biomarkers emanating from the breast microenvironment may be detected within breast tissue itself, within hematogenous circulation, or within lymphatic circulation [70]. Nonetheless, it is important to know from which cell population these biomarkers were derived. Breast cancer protein biomarkers may originate from breast tumor cells, stromal cells, or tumor/stromal infiltrating immune cells. Due to varying levels of discordance between RNA and protein expression, functional biological features are not fully represented in intrinsic gene expression signatures [8]. Functional proteomic analysis provides complementary information that can be integrated with the genomic and transcriptomic data to enhance biomarker discovery [71]. Signaling pathways, potential biomarkers, actionable drug targets, and overall tissue biology can be discerned both within and across tumors and breast cancer subtypes by integrating genomic data with immunohistochemistry (IHC), reverse phase protein arrays (RPPA), and/or liquid chromatography mass spectrometry (LC-MS/MS) proteomic data.

Despite technological improvements, the number of clinically validated biomarkers is dismally small [68]. The perceived failure of proteomics has been attributed to biased

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

sample preparation, flawed experimental design, and inappropriate statistical data analysis [68,72].

Cell line models are appealing to use for pre-clinical biomarker studies. Cell lines have been classified into the intrinsic subtype categories as well as groups based on their protein receptor expression [73]. The major caveats to using cell lines for breast cancer subtyping studies are (a) the difference in genetic mutations in transformed cell lines compared to patient samples, i.e. the *K-RAS* mutation in the MDA-MD-453 cell line that is used as a model of apocrine breast cancer has not been identified in patient specimens [74], and b) a lack of mixed cell- type cultures that recapitulate the tumor microenvironment, and c) lack of protein expression concordance between cell lines and TCGA breast proteomic data for proteins other than ER, PR and HER2 [75]. Acknowledging and understanding these criticisms paves the way for extracting valuable information from existing breast cancer studies, as well as improving prospective proteomic analysis. Experimental factors and study design need to be critically evaluated for retrospective and prospective studies [76,77]. Pre-analytical variability must be minimized by establishing consistent tissue collection, fixation, and processing procedures. Study participants must represent individuals with and without the disease being studied. Analytical methods must consider and reduce variability introduced into the data due to the heterogeneous nature of breast tissue. Potential biomarkers must be verified and validated in large numbers of people to improve predictive power. Designing studies that allow serial sampling of the tumor during neoadjuvant therapy may facilitate identifying proteomic biomarkers of treatment response and resistance [78,79].

With these experimental caveats in mind, we provide an overview of immunohistochemical, RPPA, and mass spectrometry studies that are addressing biomarker classification and biological heterogeneity within and between breast cancer subtypes. These proteomic studies are helping to clarify clinically relevant issues including better disease free and overall prognosis and predicting therapy response with the goal of limiting unnecessary toxicity and reducing development of drug resistance.

2.1 IHC4 score for rapid, inexpensive breast cancer prognosis

Rapid, inexpensive, widely available IHC assays can reduce overall health care costs and are amenable to standardized, reproducible methods. As individual biomarkers, ER, PR, Ki-67, HER2 have been used clinically as prognostic indicators of disease recurrence/clinical outcomes (Figure 3). Combining all 4 markers could have added prognostic value [80]. With this philosophy in mind, Cuzick et al. developed a 4 protein IHC panel as a prognostic tool to estimate residual recurrence risk for patients previously treated with hormone therapy [80]. IHC markers for ER, PR, Ki-67, HER2 were combined with nodal status, tumor size, grade and age, and type of drug treatment (anastrozole versus tamoxifen) to generate the IHC4 algorithm. Compared to the *Oncotype* Dx genomic recurrence scores, IHC4 was slightly more prognostic for distant recurrences, whereas *Oncotype* Dx performed better for all recurrences [80]. This study was performed with two clinical cohorts, using Ki-67 antibody clone SP6 or MIB-1, and an ER positivity cut-off of H>10. The IHC4 score has since been evaluated in a larger clinical cohort using different IHC detection methods [81],

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and in a setting with a multidisciplinary clinical decision team for recommending adjuvant chemotherapy [82]. Both studies reported improved prognostic value of the IHC4 score.

Biomarkers for predicting response or resistance to radiotherapy are urgently needed. Lakhanpal et al. evaluated whether IHC4 algorithm, which includes nodal status, tumor size, grade and age, and type of drug treatment (IHC4+CTS score), could accurately predict the risk of locoregional recurrence in women who had breast conservation surgery (BCS) for early invasive breast cancer [83]. The study consisted of 3 groups: 1) the study group who had BSC but no adjuvant radiotherapy (n=81), 2) the control group had both BCS and adjuvant radiotherapy (n=1406), and 3) the validation group included women from the control group who were post-menopausal and did not have adjuvant chemotherapy [83]. The IHC4+CTS score was divided into tertiles of low risk, intermediate risk, or high risk. The actual locoregional recurrence rate predicted by IHC4 + CTS scoring was 2.7% (1/37), 22.2% (4/18) and 23.7% (6/26) for the low, intermediate, and high risk groups, respectively. At 5 year follow-up, 97% of the patients in the low-risk group, 79% of the patients in the intermediate-risk group, and 75.2% of the patients in the high risk group were recurrence free. The results suggest that the IHC4+CTS score shows promise for predicting patients with a low local recurrence risk, providing cost and quality of life benefits for those who can avoid adjuvant radiotherapy. However, others have not endorsed the IHC4 score for ER/PR positive/HER2 negative patients [37], highlighting the urgent need for large scale validation proteomic biomarker panel.

2.2 Immunocytochemical analysis of triple negative breast cancer

Triple negative breast cancers (TNBCs) are ER/PR/HER2 negative, with a high mitotic index, prominent lymphocytic infiltrate, and a high probability of recurrence and metastasis. TNBC has the poorest prognosis within breast cancer subtypes. The absence of biomarkers for predicting outcome or tailoring therapy, leaves surgery and adjuvant chemotherapy as the current treatment options for TNBC.

Four subtypes of TNBC have recently been defined by combining a histological assessment of lymphocyte invasion and stromal gene expression with transcriptomic data [84]. The 4 subtypes reflect more intricate histomorphology designations compared to the previously published six TNBC subtype designations developed by the same group. The revised subtypes are: basal like 1, basal like 2, mesenchymal, and luminal androgen receptor (AR+). This retrospective study of publicly available data sets, including TCGA cases, revealed differences in clinical factors and distant disease progression based on the revised 4 subtype classification of TNBC [84]. In the same study, using a small cohort of specimens (n=10), transcripts from tumor and stroma were confirmed to be discordant, highlighting differences in homogenized tissue specimens versus enriched cell populations and the effect of using homogenized tissue for defining subtypes [84].

Triple negative breast cancer falls within the basal-like intrinsic subtype [41,85]. Despite the basal-like classification, not all TNBC is basal-like and not all basal-like tumors are triple negative. Basal-like and non-basal-like TNBC can be distinguished on a protein level by EGFR and CK5/6 expression. Basal-like TNBC expressing CK5/6, CK14, CK17 and/or EGFR have a poor prognosis [85,86]. A triple immunofluorescence test for CK5/6, AR and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

p53 found that TNBC cells do not simultaneously express all three proteins, although each marker alone was expressed in specific cell populations as noted by IHC on formalin fixed paraffin embedded (FFPE) tissue sections (n=52) [86].

Immune cell activation has been documented in TNBC breast tumors. A subset of TNBC tumors was shown by IHC to express both programed cell death 1 (PD-1) protein and lymphocyte activation gene 3 (LAG-3) protein [54]. PD-1/LAG-3 expression could potentially be used to identify TNBC patients that could benefit from immunomodulatory therapy.

To further complicate the molecular portrait of TNBC, a subset of TNBC expresses AR [87]. AR expression has been correlated with EGFR and PDGFRβ phosphorylation indicating activation of receptor tyrosine kinase (RTK) mediated growth pathways [88]. Androgen receptor expression in TNBC has also been correlated with locoregional recurrence following radiation therapy and radioresistance [89]. Inhibiting AR with enzalutamide induced radiation sensitivity in MDA-MB-453 and ACC-422 TNBC cell line models which highly express AR. Enzalutamide induced radiosensitivity in a dose-dependent manner, similar to cisplatin. Radiosensitization was not observed in AR negative cell lines (T47D, MDA-MB-231, MDA-MB- 468) [89]. MDA-MB-453 mouse xenograft models treated with enzalutamide demonstrated radiosensitivity documented as delayed tumor doubling time. DNA damage repair genes, regulated by AR, were identified from gene set enrichment analysis [89]. Western blotting was used to determine protein expression of DNA-PKcs and phosphoDNA-PKCs Ser2056 with and without ionizing radiation in MDA-MB-453 cell lines, treated or untreated with enzalutamide. DNA-PKcs Ser2056 increased with ionizing radiation treatment but decreased over time following AR inhibition [89]. Large clinical studies are assessing the effects of breast cancer subtypes on radiosensitivity [90], but to be applicable to all TNBC subsets, the studies should include assessment of AR expression as well as ER/PR/HER2.

CD44, a transmembrane adhesion protein that binds hyaluronic acid, is often expressed in breast tumors and is associated with the cancer stem cell phenotype. CD44 acts as a coreceptor for HER2 and MET and a scaffold for heparin sulfate proteogylcans, whereby it can modify cell shape via linkages to the actin cytoskeleton. Using tissue microarray arrays from 240 stage 2 breast cancer patients representing all subtypes, CD44 expression was frequently associated (63%) with triple negative basal-like tumors that also expressed CK5/14 and EGFR, and with *BRAC1* hereditary tumors [91]. HER2+ breast tumors exhibited the lowest CD44 expression. Alternatively spliced isoforms of CD44 have variable numbers of exons and specific isoforms have been associated with breast cancer subtypes [92]. Using tissue microarrays, western blotting, and qRT-PCR, isoform CD44v8-v10, containing three variable exons, and CD44S, without any variable regions, were associated with basal-like tumors [92].

3. Reverse phase protein arrays, antibody arrays, and western blotting

RPPA are ideally suited for quantifying and characterizing proteins and their post-translationally modified forms [93]. RPPA are utilized to identify and quantify changes in

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

levels of expressed proteins and to map the functional protein signal transduction pathways that represent actionable drug targets [94]. RPPA are not used for *de novo* protein identification, as in mass spectrometry, but instead provide high-throughput protein analysis using antibody based detection systems.

3.1 Signal kinase pathway regulation in triple negative breast cancer

In a small set of previously treated, metastatic TNBC tumors (n=14), a variety of genetic alterations were identified by whole genome and transcriptome sequencing [95]. Although it was a small study, the goal was to identify potentially actionable molecules. A large variety of inter-patient genomic alterations were discovered, but some commonalities emerged including alterations in genes regulating DNA repair and in genes driving the RAS/RAF/MEK/ERK and PI3K/AKT/mTOR signaling cascades [95]. Concurrently, another group was also investigating actionable proteins in TNBC tumors by profiling receptor tyrosine kinases using antibody arrays (human phospho-RTK array kit (R&D Systems) and the PathScan RTK Signaling Antibody Array Kit (Cell Signaling) [96]. Co-activation of mTOR (mammalian target of rapamycin) and ERK1/2 (extra-cellular signal related protein kinases 1 and 2) in patient derived TNBC tumors and TNBC cell lines was shown by antibody arrays and western blotting to be modulated by BEZ235, a dual PI3K/mTOR inhibitor [96]. These two studies accentuate the inter-patient heterogeneity in TNBC and potential actionable drug targets within the protein translation and DNA repair pathways. Recently, two types of cluster analysis were applied to RPPA data generated from 80 TNBC tumors to decipher proteomic differences within the TNBC subtype [97]. The tumors consisted of invasive lobular or ductal breast cancer. 152 proteins (n=46 phosphorylated proteins) were quantified by RPPA. Clustering analysis showed 42 differentially expressed proteins in the two clusters. The clusters were dominated by inflammatory, hormone receptor, and MAPK pathway signaling, or by DNA damage, GADD45, and p53 pathway signaling [97]. The functional phenotypes highlight the heterogeneity within TNBC and identified two clusters representing patients that could benefit from different types of treatment

3.2 Proteomic heterogeneity within HER2 positive and negative breast tumors

Dynamic receptor switching, receptor cross-talk, and receptor activation without receptor overexpression complicate efforts to classify breast cancer subtypes [51,98,99]. Truncation of the HER2 extracellular domain results in a hyperactive receptor, p95HER2, due to the ability of the C-terminus fragment, which contains tyrosine kinase activity, to form stable homodimers. p95HER2 has been associated with trastuzumab resistance [51]. HER2 and HER3 readily form heterodimers that activate the AKT-mTOR protein translation and growth pathways. In Her2+ FFPE tumor specimens, high levels of p95 and/or HER3 were associated with poor response to trastuzumab [51]. Furthermore, plasma proteome analysis of patients with HER2 negative breast cancer revealed elevated plasma and tumor tissue levels of calpain-10 [100]. Calpains cleave the extracellular domain of HER2, thus forming a hyperactive, but phenotypically HER2 negative cell [100].

Wulfkuhle et al. reported HER2 signal pathway activation, via phosphorylated EGFR and HER3, in frozen and FFPE breast cancer tissues that were defined as HER2 negative by

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

fluorescent in situ hybridization (FISH) and IHC (HER2 <2+, FISH negative) [99]. Phosphorylated proteins were quantified using RPPA in microdissected breast tumors from patients enrolled in the ISPY-1 TRIAL (CALGB 150007/150012, ACRIN 6657). Assessment of phosphorylated HER2 (HER2 Tyr1248) levels in the FFPE surgical specimens revealed 46% of the HER2 cases that were IHC/FISH+ expressed HER2 Tyr1248 compared to 26.5% of the IHC/FISH negative HER2 samples [99]. This pattern of HER2 pathway activation was also present in the frozen specimens from the IHC negative/FISH negative/phosphoHER2 positive subset. Phosphorylation levels of HER2 Tyr1248, EGFR Tyr1173, and EGFR Tyr1148 were elevated (p=0.00002, 0.01, and 0.07, respectively) in the IHC negative/FISH negative /pHER2 positive population compared with the IHC/FISH/pHER2 negative population. In addition, the HER2+ cluster of breast specimens in the TCGA showed over-expression of EGFR, EGFR Tyr992, HER2, and HER2 Tyr1248 [101]. The clinical implications regarding activated HER2 receptor, via phosphorylation rather than receptor amplification, could alter therapy recommendations for the subset of patients who are HER2 IHC/FISH negative but phosphoHER2 positive. Patients whose tumors are HER2 negative based on IHC/FISH scores are currently not considered candidates for anti-HER2 therapy. However, activation of HER2 family proteins and related downstream RTKs indicate that a subset of patients could potentially derive clinical benefit from molecular targeted inhibitors [99]. Furthermore, these HER2 receptor studies indicate that the internal domain drives the downstream signaling, either via hyperactivation of a truncated receptor or through phosphorylation via heterodimerization with HER3. Molecular profiling of HER2 needs to include phosphoHER2 and p95HER2 to provide a complete snapshot of receptor status.

3.3 Heterogeneity within the iuminai A subtype

The luminal A subtype is considered highly heterogeneous but also has an overall good prognosis. However, despite hormone therapy a subset of patients relapse and other patients may receive unnecessary hormone therapy. Using integrative clustering of genomic and proteomic data for frozen primary breast tumors, a refinement of subtypes was undertaken to identify potential differences that could be exploited for therapeutic options [102]. RPPA data for 148 proteins from 173 specimens resulted in 5 groups: luminal, HER2+, basal, reactive 1 and reactive 2 [102]. The reactive groups exhibited proteins related to fibroblasts and the microenvironment. Based on RPPA functional protein classification, subtype differences were noted within the luminal A group. 6 proteins were statistically different between the RPPA- defined luminal A subgroups: Cleaved Caspase 9, 53BP1, AMPKa, GATA3, Rad51 and p90RSK Thre359/Ser363 [102]. A caveat of this study is the wide range of tumor content in the specimens (0–90% tumor, average 53%) that could potentially explain overlap between the reactive and luminal A classifications.

Another method for identifying disease recurrence in hormone positive breast cancer subtypes was designed using quantitative RPPA data and gene expression data [103]. Three classification methods were combined: support vector machines, random forests, and prediction analysis for microarrays. 3 proteins, NDKA, RPS6 and Ki-67 were highly expressed in high risk tumors and caveolin-1 was downregulated [103]. The data was confirmed in an independent set of specimens with good sensitivity and specificity.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Although race has been associated with clinicopathological breast cancer risk factors, functional phenotype differences within breast cancer subtypes based on self-reported race/ethnicity were not found in a small study of hormone positive, HER2 positive, and triple negative invasive breast tumors [104]. Fine needle aspirates were lysed for RPPA analysis. 177 proteins were not considered differentially expressed (false discovery rate (FDR) <5%) among different races/ethnicities (n=255, 57.6% White, 18.4% Hispanic, 18.1% Black, 5.9% other) by breast cancer subtypes. The caveats to this study are; (a) the small sample sizes in each breast cancer subtype, (b) the potential paucity of tumor cells and/or overabundance of stromal cells in the fine needle aspirates, and (c) the lack of ancestry-confirmed race/ethnicity [104].

3.4 Invasive lobular breast cancer subtypes

Invasive lobular cancer (ILC) represents about 10% of all breast cancers but ILC specimens are not well represented in most study cohorts or clinical trials studying prognosis and treatment response [105]. To address this lack of knowledge regarding ILC, a multi-omic study of 144 untreated ILC tumors was undertaken to improve the molecular characterization of ILC. As part of the Rational Therapy for Breast Cancer (RATHER) consortium, DNA sequencing of somatic variants, copy number variation, gene expression, and RPPA protein data were integrated to provide molecular portraits of ILC and identify potential subtypes [105]. Gene expression data elucidated two subtypes identified by pathway analysis as immune related and hormone related. The immune related group exhibited mRNA upregulation of lymphoid signaling/cytokine receptor transcripts [105]. Transcripts upregulated in the hormone related group included ER, PR, GATA3, and cell cycle. Somatic mutations were identified in both groups of ILC tumors including *CDH1 (E-Cadherin), PIK3CA, GATA3, MAPK, HER2, NF1,* and *TP53* [105]. Using comparative analysis between the TCGA and METABRIC breast subtype classifications, the immune related and hormone related subtypes of ILC were found to be unique classifiers [105]. RPPA analysis confirmed higher expression of HER2, ER, ER Ser118, PR, fibronectin, and GATA3 in the hormone related group of ILC [105].

Using a data integration tool, mutations, copy number variations, and RPPA data between the immune and hormone related groups revealed correlations only in the hormone group. Upregulation of PR, GATA3, fibronectin and down regulation of yes associated protein-1 (YAP1) were verified at the transcript and protein level [105]. YAP1, a Hippo pathway protein, regulates growth and repair. Downregulation of YAP1 has been shown with *in vitro* models to lead to ER/PR overexpression [105]. The lack of correlations in the immune related group may be due to under-representation of immune signaling antibodies available for this particular RPPA [105]. Overall survival (10 years post diagnosis) was associated positively with higher expression of H2AX, while higher expression of eIF4B was associated with poor survival [105]. Furthermore, a decision tree based on mutation rate, eIF4B protein expression, and positive lymph node count were predictive of survival [105]. This comprehensive integrative study clearly shows the contribution of RPPA for validating gene expression data, quantifying protein expression, and characterizing post-translationally modified proteins.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

4. Emerging biomarkers

Emerging biomarkers have limited validation due to evaluation in small patient cohorts. However, emerging biomarkers may be extremely useful in research settings for deciphering drug sensitivity and resistance, or developing additional clinical disease monitors and treatment monitors. Several studies have recognized the forkhead box transcription factors as potential predictive biomarkers for overcoming trastuzumab resistance or for the need for adjuvant chemotherapy [106–109]. FOXA1 has been positively correlated with good prognosis in ER/PR positive breast cancer, and negatively correlated with Ki-67 and nuclear grade [106,107]. GATA-3, transacting T -cell specific transcription factor, has also been found to be associated with luminal A and luminal B subtypes [110]. GATA-3 correlated with FOXA1 but FOXA1 remained the better independent marker for prognosis and prediction for adjuvant chemotherapy [106,107].

The Rac activator, P-REX1, was found by mass spectrometry profiling of MCF7 cell lines during a screening for phosphatidylinositol 3,4,5-triphosphate (PIP3) regulated proteins [111]. Using the RPPA data from the TCGA primary breast cancer cohort [112], P-REX1 was found to be inversely correlated with PI3K pathway inhibition and P-REX1 levels decreased with loss of PTEN. Both protein and mRNA levels of P-REX1 were positively correlated with ER [111]. P-REX1 was also shown to be elevated in ER+ luminal breast tumors. P-REX1 will need to be verified in larger cohorts of microdissected tumors to fully assess its tumor or stroma related biology.

Numerous potential biomarkers languish in the literature due to the need for rigorous validation studies prior to clinical adoption. Other emerging biomarkers relevant to breast cancer subtypes awaiting validation are mentioned here. A biomarker of interest is heterochromatin protein 1 family (HP1β), which regulates gene expression and DNA damage, as a potential prognostic and predictive biomarker for chemotherapy and PARP inhibitor treatment [113], Evidence from gene expression, IHC, and MCF7 cell lines indicate that overexpression of HP1 β was associated with poor prognosis [113]. A biomarker for radio-iodine treatment of breast cancer may potentially be found in sodium-iodine symporter protein. Sodium-iodine symporter is over-expressed in ER+ breast cancer, with staining intensity equivalent to thyroid tissue [114]. Invasive tumors and DCIS express sodium iodide symporter more frequently than the normal adjacent tissue indicating that it could be a marker for radio-iodine therapy [114]. Two studies indicate the EpCAM expression was associated with worse overall survival [115,116]. Using IHC and clinicopathological data, EpCAM expression was shown to confer a poor prognosis in basal-like and luminal B HER2+ breast tumors [115,116]. Another extra cellular matrix associated protein, collagen 10a, has been shown by gene expression and IHC to be associated with poor prognosis in ER+/HER2+ breast tumors [117]. Sparse tumor infiltrating lymphocytes also correlated with lower frequency of pathological complete response. The combination of collagen 10a expression and the amount of tumor infiltrating lymphocytes had a higher predictive value [117].

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

5. Mass spectrometry

5.1 Pioneering mass spectrometry based studies of the breast proteome

For over 16 years, there has been evidence that multi-directional cellular interactions within the breast microenvironment influence breast tumor biology [118]. The challenge then, and now, is finding specific, reproducible biomarkers within the 19,000 proteins of the human proteome. Initial studies characterizing the breast tumor interstitial fluid using 2-D gel electrophoresis identified a total of 1,147 proteins [118]. Fresh mastectomy specimens containing invasive ductal carcinoma were used to extract tumor interstitial fluid as a fluid repository representative of the tumor microenvironment. Matrix assisted laser desorption ionization time of flight (MALDI-TOF) mass spectrometry and/or western blotting confirmed the protein identification [118]. Similar proteins were discovered in all 16 specimens, including proteins related to cell proliferation, invasion, inflammation, protein synthesis, and the actin cytoskeleton. However, differences were noted between specimens including immunoglobulins, Apo A-I, and actin which reflected interpatient heterogeneity. Two additional studies by the same group investigated the correlation between histomorphology and cancer phenotype in apocrine metaplasia of the breast [119,120]. These studies have elucidated proteomic alterations in the progression from benign lesions to apocrine cancer. A third of pre-menopausal women have breast cysts which are associated with an increased breast cancer risk. Apocrine cysts, also known as blue dome cysts, are generally ERneg/PRneg/ARpos, HER2 negative and Bcl-2 negative. Apocrine metaplasia arises from ductal epithelium that is ER+/PR+ but the apocrine cells loose hormone receptor expression and begin to express 15-hydroxyprostaglandin dehydrogenase (15-PGDH) [120], supporting the concepts of dynamic receptor switching in breast cancer cells and intra-tumor heterogeneity. The apocrine cyst fluid exhibits a high potassium:sodium ratio compared to serum [120], suggesting the presence of mutations in solute transporter genes. 15-hydroxyprostaglandin dehydrogenase (15-PGDH) and 3- hydroxymethylglutaryl-CoA reductase were discovered to be present only in apocrine type I cysts, apocrine metaplastic cell in type II cysts, terminal ducts and papillary lesions, but not in normal breast epithelium or type II flat cysts adipose tissue [120]. 15-PGDH was verified in a larger study to be expressed solely in invasive apocrine breast cancer but lacked expression in other breast cancer subtypes [119]. Vitamin D binding protein, cathepsin D, and Hsp60 were also found in apocrine cysts. Vitamin D binding protein is regulated by AR, thus suggesting a metabolic link between vitamin D levels and apocrine breast cancer.

In a follow-up study, Gromov et al. discovered two additional protein biomarkers, brain fatty acid binding protein (FABP7) and hydroxymethylglutarate-CoA synthase 2 (HMGCS2), that enhanced the classification of apocrine breast [121]. Apocrine phenotype could be defined by combining HMGCS2, ACSM1 and 15-PGDH. Furthermore, adding HMGCS2 to hormone the standard IHC profile decisively delineated invasive apocrine breast cancer with a phenotype of HMGCS2+/AR+/ERneg/PRneg [121].

5.2 Mass spectrometry proteome profiling of cell lines

Surface enhanced laser desorption ionization time-of-flight mass spectrometry (SELDI-TOF) and MALDI-TOF were used to profile the proteome of 27 breast cancer cell lines to

Author Manuscript

demonstrate the utility of mass spectrometry-based proteomics for delineating proteomic complexity [122]. The cell line protein profiles were distinguishable as two groups, ductal carcinoma or mixed ductal, mesenchymal, medullary, and non-cancerous. These two clusters had previously been defined as luminal-like and basal-like, respectively. 10 proteins were differentially expressed in the basal-like cell lines (GATA3, CK19, EGFR, CD10, Met, CK5/6, Caveolin-1, Moesin, CD44, and ETS1 [122]. Ubiquitin was associated with luminal cell lines, whereas S100A9 was associated with basal-like cell lines. S100A9 was verified on a tissue microarray of 547 breast cancer tumors to be present in the basal-like tumors and was associated with poor prognosis [122]. This study confirmed that mass spectrometry-based proteomic profiling could recapitulate the intrinsic breast cancer subtype classifications.

Cell lines are often used as models of breast cancer subtypes, therefore it is useful to determine functional biological differences between the cell lines. Global protein abundances were evaluated using LC-MS/MS in 4 cell lines representative of normal, luminal, triple negative, and luminal/HER2+ breast cancer subtypes [123]. The MCF10A cell line represented normal mammary epithelial cells for comparison to MCF7 (ER+, luminal-like), SKBR3 (HER2+), and MDA-MB-231 (mesenchymal, triple negative) cell lines. Spectral counting was used to determine differentially expressed proteins. The relative protein abundance was calculated by normalizing the sum of the spectral counts for all 3 tumor cell lines to the spectral counts for MCF10A cells [123]. 68/82 differentially expressed proteins were decreased in the tumor cell lines compared to MCF10A, whereas only 13 proteins were increased in abundance [123]. Filamin C gamma was increased in mesenchymal cell line MDA-MD-231, but was decreased in MCF7 and SKBR3 cells, consistent with epithelial to mesenchymal transition [123]. Bioinformatic pathway analysis revealed loss of several proteins involved in cell development and cell morphology/adhesion, including loss of key basement membrane proteins, including collagens, neprilysin, tenascin C, and integrin alpha 11 [123]. Focal adhesion kinase (FAK) was inferred by pathway analysis to be included in the protein network of the 82 differentially expressed proteins. FAK regulates adhesion, motility, spreading, and cell survival via phosphorylation of six sites [124] FAK links the extra cellular matrix to the cytoskeleton via a complex protein scaffold with integrins, and integrins in turn induce FAK auto-phosphorylation [125]. FAK was inferred to be within the same pathway network as many of the down-regulated proteins identified via global protein abundance comparison emphasizing the role of adhesion and the extracellular matrix in tumor invasion across breast cancer cell types [123].

Another study of breast cancer subtypes using cell line models was conducted through the Chromosome-Centric Human Proteome Project [126]. This project integrated RNA-Seq (Illumina TruSeq) from Chromosome 17 and LC-MS/MS data to examine splice variants in HER2+ breast cancer cell lines (n=3) with different phenotypes. The models were SKBR3 (HER2+ adenocarcinoma), SUM190 (HER2+ inflammatory breast cancer), and SUM149 (low HER2 expression, inflammatory breast cancer) [126]. Chromosome 17 is of interest in breast cancer due to its high density of protein coding genes, including *TP53, HER2,* and *BRAC1.* HER2 splice variants were found in all three cell lines, five EGFR variants were found in SKBR3, 6 in SUM149, and no EGFR variants were discovered in SUM190 [126]. STAT3 variants were present in SKBR3 and SUM190. MAPK1 variants were found in

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

SUM149 and SUM190 [126]. Splice variants can alter protein function, therefore identifying cell line models with specific isoforms can facilitate drug testing to determine which isoforms are preferentially susceptible to specific therapeutic compounds.

Predicting therapy response for patients with TNBC is a high research priority for precision medicine. Three groups have used quantitative mass spectrometry for biomarker discovery in patient derived tumor samples [127,128] and/or cell lines [129]. Using MCF10A wild type and PIK3CA knock-in mutants, Young et al. observed 72 altered proteins between the two models, with secreted, extra cellular matrix interaction, cell surface receptors, intermediate filaments, mitochondrial proteins to be differentially expressed [129]. PIK3CA mutations are common in basal-like cancer [95,129] and are associated with poor prognosis and therapy response. PIK3CA mutations often occur in the helical domain (E545K, E542K) or the catalytic domain (H1047R) which activate the PI3K-PTEN-AKT signal transduction kinases, resulting in growth and proliferation [129]. Key up-regulated proteins expressed in the PIK3CA mutant cell lines included peroxidasin (PXDN), laminin y2, fibronectin, EphA2, and TGM2, with decreased expression of PTPRF, a protein tyrosine phosphatase receptor. Silencing of PTPRF with shRNA and upregulation of EphA2 cooperated to allow EGFR activation in the MCF10A mutant cell lines [129]. Based on the above data, using cetuximab to block EGFR, and decrease phosphoERK and proliferation in PIK3CA mutant cell lines appears to be a rationale treatment. However, cetuximab+cisplatin have failed to meet the primary endpoint in two clinical trials, although an improvement in overall response was noted in TNBC patients [130,131]. The failure of cetuximab could be due to receptor cross-talk between EGFR and amphiregulin, which was elevated in the MCF10A PIK3CA mutants and EGFR is frequently over-expressed in TNBC [129]. A comparison with proteomic data in the TCGA breast cancer cohort showed PIK3CA mutations were associated with elevated levels of EGFR and EGFR Tyr1068. Increased expression of peroxidasin mRNA, an extra cellular matrix protein involved in collage crosslinking, has been associated with decreased recurrence free survival [129].

Wu et al. used phospho-tyrosine targeted mass spectrometry, siRNA knockdown, and gene expression data to identify RTKs associated with aggressive breast cancer phenotypes [132]. Using 26 triple negative cell lines, seven phospho-tyrosine peptides were associated with aggressive phenotypes. Inhibition of aggressiveness was greatest with siRNA knock down of AXL and TNK2. AXL is a cell surface receptor that transmits signals from the extra cellular matrix via growth factor GS6. AXL ligand binding phosphorylates PI3K subunit, Grb2, resulting to AKT activation. AXL gene expression was increased in TNBC and was inversely correlated with patient relapse-free survival [132]. Phospho-AXL correlated with anchorage independent growth and invasiveness in the cell line models. [132]. 4/7 cell lines (MCF10A, MCF12A, HCC1187, SUM225) were categorized as non-aggressive based on phosphotyrosine profiling. 5 cell lines, MDA-MB-231, HCC-1569, HCC-1395, BT549, SUM159, exhibited hyperphosphorylation of AXL, DYRK2, TYK2, EphA2, and TNK2 [132]. AXL could potentially be a biomarker of poor prognosis in TNBC patients or as an indicator for AXL targeted therapy [133].

In a pre-clinical cell line study, proteogenomic technologies were combined by Lawrence et al. to identify actionable protein components in TNBC. Shotgun mass spectrometry using

Author Manuscript

intensity based absolute quantitation revealed significant differences in protein expression across the cell lines (n=16 TNBC, n=3 Luminal, and n=1 non-tumorigenic) [134]. The MDA- MB-453 triple negative cell line expressed HER2+ 20-fold greater than the median expression, which does not correspond to the IHC data for MDA-MB-453 [73,74]. Two possible explanations for this discrepancy are: (1) functional cross-talk between AR, HER2, and/or downstream kinases such as AKT and p70S6 since MDA-MB-453 expresses AR [74,88]; and (2) Dynamic receptor switching of HER2, altering the functional state of the receptor [51,98,100,135]. Heterogeneous protein expression was noted across all cell lines. However, two groups and four clusters of cell lines emerged using hierarchical clustering integrated with frequent mutations and copy number variation. The groups were defined by differences in *TP53* mutation status and insulin signaling, nucleotide excision repair, focal adhesion, ubiquitin-mediated proteolysis and extra-cellular matrix interactions [134]. Comparing the cell line data to TNBC tumor specimens (n=4), the most abundant proteins in the patient tumors were highly abundant serum proteins and blood cell proteins. The centroid of each cell line cluster was used to correlate the tumor sample proteomic data. As expected due to the influence of the microenvironment, the tumor samples contained more proteins involved in extracellular matrix interactions and antigen expression [134]. NF-κB p65, an isoform that lacks the IKKB regulatory region, was found in all four tumor specimens. NF-κB modulates inflammatory responses, which are known to be involved in invasion and TNBC has a high metastatic rate [134].

5.3 Mass spectrometry profiling of human breast tissue

Cha et al. developed Protein Set Enrichment Analysis (PSEA) to discern protein pathway patterns from data generated by shotgun proteomics and quantified via spectral index of protein abundances [110]. PSEA is a modification of gene set enrichment analysis. PSEA employs the spectral index abundance of differentially expressed proteins to highlight proteins with similar biological function but different expression levels [110]. Using this approach, they compared nine human normal breast samples to nine non-patient matched ER + luminal subtype tumors.

Normal breast or tumor epithelial cells were procured by laser capture microdissection to reduce interference from stromal and immune cells. 298 differentially expressed proteins were applied o PSEA to render functional protein maps of normal and luminal breast epithelium [110]. Based on PSEA, tumors exhibited decreased levels of cytoskeletal proteins, collagens, fibrinogen, laminins, hemopexin, 14–3-3σ, lumican, TGF-β, and serpin peptidase inhibitor. Enrichment of proteins related to transcription binding factor motifs, including p53, SMAD, NF- kB were also found. Using spectral index abundance for the tumor cohort, elevated levels of fibronectin and mitochondrial isoleucyl-tRNA synthetase were discovered [110]. Exploiting the wealth of IHC information in the Human Protein Atlas (HPA) [136,137], Cha el al. visually quantified IHC breast images in the HPA to verify expression levels and subcellular location for 25 of their differentially expressed proteins. 75% concordance was noted with the spectral index abundance. The lack of complete concordance was surmised to be due to the limited dynamic range of IHC immunoperoxidase staining intensity, or lack of breast cancer subtype annotation in version 5.0 of the HPA [110].

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Panis et al. expanded on this work using label-free mass spectrometric analysis of human breast tumors representing each of the four intrinsic breast cancer subtypes, plus normal human breast tissue [138]. Each cancer subtype consisted of 20 patients enrolled in the REMARK trial (The Reporting Recommendations for Tumor Marker Prognostic Studies). All four cancer subtypes exhibited changes in cytoskeletal proteins and cell adhesion [138].

Fresh frozen breast tissue from 477 sporadic and hereditary breast tumors has helped to further confirm biological differences in breast cancer subtypes on a proteomic level [139]. Homogenized tumor tissues were analyzed with 2D-DIGE and LC-MS/MS, using pooled specimens. The breast cancer subtype pools consisted of normal-like (n=4), luminal A (n=14), luminal B (n=4), HER2 (n=5), and basal-like (n=15) [139]. Pathway analysis demonstrated chromatin upregulation in luminal A and B tumors. Cytoskeletal remodeling proteins, clathrin, and vesicle transport were upregulated in luminal B tumors. Integrins were down-regulated in basal tumors, whereas PARP1, poly ADB-ribose polymerasel was upregulated. As expected, HER2+ tumor over-expressed HER2, but they also were found to have down-regulation of TRAF2, TRADD, and NEMO [139]. Using pooled specimens allows faster analysis time, but the caveat is the inability to discern individual patient differences.

The merits of top-down and bottom-up mass spectrometry were recently illustrated using two patient derived mouse xenograft models of basal and luminal B breast cancer. WHIM2, Washington University Human in Mouse, is a basal-like TNBC xenograft, and WHIM16 is a ER+/PR+/HER2neg luminal B xenograft [140]. Top-down proteomic analysis allows direct analysis of small intact proteins (less than 30kDa), whereas the bottom-up approach detects enzymatically digested peptides for database matching to intact proteins. After removing mouse proteins and human proteins homologous to mouse proteins, the bottom-up approach detected 3,519 protein groups, with elevated levels of α-endosulfine in WHIM2 [140]. The top down method quantified 982 proteoforms matching 358 proteins, but also revealed a diphsophorylation of α-endosulfine only in WHIM16 [140]. Phosphorylation significantly changes the secondary structure of α-endosulfine and its protein-protein interactions, highlighting the complementary information of each mass spectrometry method.

Stable isotope labeling with amino acids in cell culture (SILAC) was improved by including a different SILAC labeled cell line as an internal standard to quantify peptides (SuperSILAC) [141]. SuperSILAC was utilized to quantify proteins from human tumors to create functional network maps based on breast cancer cell type [142]. FFPE breast tissue was macrodissected to enrich the lysate with tumor cells from ER+, HER2+, or TNBC specimens. Integrated genomic and proteomic data was used to find molecular signatures that are differentially expressed between breast cancer subtypes. Annotation matrix analysis, a type of 1- dimensional enrichment test, was applied to the data. The distribution of mean differences of all proteins ina category was statistically analyzed for shifts compared to the global distribution of values in a breast cancer subtype [142]. Heterogeneity within the subtypes was clearly evident in more than 7000 proteins. Energy metabolism related proteins were different between the subtypes. The average protein fold change between the 3 tumor subtypes was compared to the Molecular Signatures Database for 706 genes that are significantly enriched in breast cancer subtypes [142]. The HER2 cluster revealed increased

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

interferon and immune signaling (IRF3) targets. The TNBC cluster showed increased EGFR and CD44 [142]. Functional pathway mapping demonstrated distinct difference in energy metabolism between ER+ (upregulated) and HER2+ (downregulated) subtypes. The TNBC subtype exhibited upregulation of translation, ribosome biogenesis, tRNA synthetase, cell growth, replication and DNA damage repair, with downregulation of PTEN [142]. A subsequent report from the same lab used SuperSILAC to identify biomarkers with clinical utility during breast cancer progression [143]. Proteins were quantified in 88 untreated ER+/HER2neg FFPE breast tissue specimens. The specimens represented lymph node negative tumors (n=21), tumors with lymph node metastasis (n=20), lymph nodes (n=25), and uninvolved adjacent breast ducts. Proteins were highly correlated between the primary tumor and lymph node metastasis, more so than with the adjacent breast ducts [143]. Extensive metabolic pathway modulations were detected between the cancer and adjacent normal tissue. DNA repair proteins and aminoacyl tRNA synthetases were down-regulated in tumors. However lysosomal and proteasome proteins were upregulated in the tumors compared to adjacent normal tissue. By IHC, two proteins, Acyl-CoA Thioesterase 1 and Glutamate-Ammonia Ligase were increased in tumors, but absent in normal tissue [143]. The lymph node proteome showed six downregulated proteins and 4 upregulated proteins, compared to the 563 upregulated and 406 downregulated proteins between tumor and normal adjacent tissue [143].

Recently, global proteome profiling utilizing LC-MS/MS of microdissected lymph node negative, therapy naive TNBC patients was used to develop an 11-protein prognostic signature [144]. This protein signature was verified in a multicenter (training set n=63, multicenter verification set n=63). The 11 protein signature was developed by selecting the protein model that showed 100% sensitivity in the area under the curve of the receiver operating characteristic curve (0.83) and a reversed model size. The 11 proteins represent cytoplasmic/cytoskeleton, Golgi apparatus, endoplasmic reticulum, nucleus and mitochondria subcellular compartments. Of note, catenin alpha-1 was included in the 11 protein signature and its corresponding gene, *CTNNA1,* was also associated with metastatic TNBC in the Craig et al. study [95]. The protein score for good (free of distant metastasis for 5 years post surgery) versus poor prognosis was compared to the consensus NIH and St. Gallen chemotherapy treatment criteria. Using the 11 protein score, 30% of the good-prognosis group would have received potentially unnecessary adjuvant chemotherapy, compared to 91–95% of the good prognosis groups based on the consensus criteria [144].

Treatment naive, frozen TNBC tumors (n=83) were analyzed using iTRAQ-OFFGEL quantitative mass spectrometry by Campone et al. to identify protein drug targets. The tumors were classified as triple negative if <10% of the tumor cells expressed ER/PR/HER2. 10% is higher than the current clinical recommendation of 1% for hormone receptor status [18,25]. Homogenized tumor containing tissues were found to possess three protein biomarkers of aggressive breast cancer that were validated by IHC in 42 tumors from the training set. Elevated levels of tryptophanyl-tRNA synthetase (TrpRs), and decreased levels of desmoplakin (DP) and thrombospondin-1 (TSP1) were predictive of five year overall survival, suggesting that these markers could evaluate tumor aggressiveness [127]. Desmoplakin forms adhesion junctions in desomosomes between epithelial cells, maintaining tissue integrity. Thrombospondin-1 is an extra cellular matrix protein that

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

exhibits anti-angiogenic and angiogenic behavior depending on the cell context. The caveat of this study is all study patients were treated with FEC100 chemotherapy and radiation. Therefore their overall survival may have been associated with factors other than expression patterns of TrpRS, DP and TSP1.

FFPE tissue sections (n=26) have also been used for targeted, label free proteomic discovery, with the goal of identifying prognostic protein biomarkers of chemotherapy response. Parallel reaction monitoring of one to four unique peptides was performed with a verification set of 114 TNBC tumors containing at least 50% tumor [128]. Shotgun proteomic data was used to correlate protein abundance with distant metastasis free survival. Transcriptomics data were available from 1,296 primary breast tumors for comparison with the proteomics data. From the 18 peptides originally identified that were associated with distant metastasis free survival, a panel of 5 peptides was verified to predict low and high risk of metastasis with a 70:30 ratio [128]. The proteins in the predictive panel were RAC2, RAB61, BLVRA, and IPYR [128]. RAC2 is proposed to regulate the actin cytoskeleton and is involved in invasion. BLVRA is biliverdin reductase A, and activated MAPK and IGF1-R signal transduction pathways. RAB6a regulates protein trafficking and IPYR is an inorganic pyrophosphatase [128]. The same group recently compared the proteome and mRNA of human FFPE ER+/PR+ tumors to TNBC using label free mass spectrometry, SRM, and a custom TaqMan Array MicroRNA card [145]. The homogenized tumors contained at least 50% tumor. 224 proteins were differentially expressed, but hierarchical clustering split the specimens into two groups, with 26% of the ER+/PR+ tumors clustered with the TNBC tumors [145]. To decipher the differences in the two different clusters of ER+ tumors, significance analysis of microarrays was performed, excluding the TNBC specimens. 44 proteins were differentially expressed between the true ER+ cluster and the triple negative-like cluster in the discovery set (FDR<5%). The ER+ cluster contained 7 different small leucine-rich proteoglycans, cathepsin G, chymase, and mast cell carboxypeptidase A. The triple negative like ER+ specimens contained cell adhesion and chaperone proteins [145]. Using SRM of a verification cohort of 46 specimens, 14 proteins were identified that could discriminate ER+ tumors from ER+ triple negative-like tumors. However, the identity of the 14 proteins was not provided [145]. Homogenized tissue was used in this study, which could either dilute or enrich the lysate with stromal and immune (mast cells). Thus, the ER+ triple negative like designation would need to be verified using microdissected tumor cells. Nonetheless, a common theme from the Young, Campone and Gamez-Pozo studies is the role of the extra cellular matrix in TNBC tumor proliferation and recurrence free survival [127–129,145].

Pre-clinical studies provide the basic research that is absolutely essential for discovering novel biomarkers. The mass spectrometry studies reviewed here are providing insights into biological processes that distinguish breast cancer subtypes and contribute to biomarker heterogeneity. As can be seen from the variation in specimen fixation and processing (homogenized tissue, macrodissection, or microdissection), the impact of biomarker study design is generally under-appreciated and under-implemented, although the concepts are well known [146,147].

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

6. Rigor and reproducibility of biomarker studies

The first drafts of the Human Proteome [148,149] generated discourse within the proteomic community regarding the definition of quality peptide identifications [150]. The positive outcome of this debate has strengthened the quality of mass spectrometry biomarker identification and proteomics databases. Consensus guidelines and standards are available for mass spectrometry [151,152]. All major human mass spectrometry data sets available through ProteomeXchange are now reanalyzed via PeptideAtlas and global proteome machine and database (GPMDB) using rigorous quality control parameters with standardized procedures [153,154]. neXtProt [155], a compendium of the human proteome, curates and integrates mass spectrometry data, including post-translational modifications, siRNA data, population variants, and protein-protein interactions, to resolve the function of uncharacterized human proteins [154]. In addition, the proteomics community, including funding agencies and journals, are facilitating the rigor and reproducibility of proteomic studies by (a) establishing data deposition guidelines, (b) creating minimum information standards for proteomic studies (Minimum Information about a Proteomics Experiment (MIAPE)) [156], (c) developing reagents and protocols for quality control of mass spectrometry instruments [157,158], and (d) creating metrics for assessing the validity of previously unknown proteins [69].

However, new biomarkers and therapies will never be validated, verified, or advance to actual patient use without adequate rigor and reproducibility in both human and animal model pre-clinical studies [68,147]. For the biomedical community to actually succeed in delivering precision medicine, pre-clinical and clinical studies must include processes to reduce variability, ensure inclusion of appropriate control subjects with adequate statistical power for discovery, authentication of reagents and chemical, followed by verification and validation of the biomarkers. The promises of precision medicine are not simply performing molecular assays for each patient and describing therapy options. The true deliverables for precision medicine are the ability to analyze an individual patient's diseased specimen, establish an optimal therapy regimen, and treat the patient with that optimal therapy, resulting in improved patient outcomes [68]. Fortunately, the importance of reproducibility, adequate statistical power, reagent verification, limiting pre-analytical and analytical variability, and including quality control processes in pre-clinical studies have renewed prominence in this era of precision medicine. European and North American agencies, such as the European Infrastructure for Translational Medicine, National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs) and the National Institutes of Health, are addressing rigor and reproducibility through guidelines, granting agency requirements for funding, access to technology, and freely available software tools to assist with statistical and experimental design [146]. The NC3Rs recently published the ARRIVE guidelines (Animals in Research: Reporting of In Vivo Experiments) describing how to design and report pre-clinical animal studies to improve the model fitness [146].

The American Society for Clinical Oncologists and the College of American Pathologists continues to update guidelines for ER, PR, and HER2 IHC-based tissue testing [25,36,159]. Methods to circumvent issues with specimen processing [33,160–166] and analysis [161] are well-documented, with general guidelines including cold ischemia time less than 60

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

minutes, 672 hours fixation in 10% neutral buffered formalin, and storing cut tissue section slides for less than 6 weeks [25,36]. An important note regarding testing guidelines concerns differences/updates in positivity thresholds. For example, the current HER2 positivity threshold is 10% strong membranous staining for cells with an IHC score 3+, but the previous threshold was 30% [167]. Differences in positivity thresholds can potentially bias meta-analysis and prospective studies in comparison to previously published data.

6.1 Protein carry-over in LC-MS/MS

Individual laboratories and scientists are the first-line resource to identify sources of variability and technical issues that impact biomarker discovery. Empirical evidence in our laboratory has exposed sample carry-over issues with nanoflow HPLC columns used with LC-MS/MS. Consensus guidelines state that carry-over must be assessed and rectified in any mass spectrometry method development [151,152,168]. Several salient points should be noted that are applicable to all LC-MS/MS studies. Firstly, significant carryover of proteins can occur in simple samples (consisting of 1–3 pure proteins). Secondly, carryover can be independent of molecular weight, pI, and hydrophobicity [169]. Thirdly, changes in flow rate, both to 100nL/min (lowest range of the instrument) and the highest flow rate (500nL/min) may not be sufficient to eliminate carry-over. Finally, protein concentration and physicochemical properties of specific proteins may enhance carry-over. For robust, reproducible clinical applications sample carry over issues must be systematically addressed (Figure 4).

7. Conclusion

Proteomic technologies contribute invaluable data for deciphering functional phenotypes and subsets within breast cancer subtypes. Widely available IHC markers (ER, PR, Ki-67, and CK) are relatively inexpensive, rapid assays that provide key diagnostic and prognostic information. Standardized protocols and interpretation guidelines implemented in 2011 allow seamless comparison of HER2 and hormone receptor status between institutions/across study sets. Technological refinements in mass spectrometry are enhancing the accuracy and precision of biomarker discovery and identification. RPPA are being used in clinical research and clinical trials to quantify biomarkers, and their post-translational modified forms, within functional signal transduction networks. Integration of genomic risk recurrence scores with the functional proteomic landscapes of breast cancer subtypes has the potential to significantly augment clinical management of breast cancer patients.

8. Expert commentary

Proteomic and genomic tissue and cellular analyses indisputably contribute to biomarker identification, drug discovery, and clinical management. The plethora of literature available regarding breast cancer subtypes attests to the rapid adoption of these technology platforms.

Integrative algorithms are being developed to further categorize this vast array of data into clinical management tools which will be invaluable for precision medicine and rationale therapy design. Nevertheless, even with our ability to apply molecular labels and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

classification schemes to breast cancer, therapies continue to fail for patients due to either *de novo* resistance, acquired resistance, or unknown/under-explored molecular mechanisms.

Identifying and quantifying *in vivo* protein-protein interactions and signal transduction cascades are currently our best representation of the functional state of cells. Knowing which cell populations are present, the *in vivo* activity state of protein kinases, and the sub-cellular location of these proteins is only possible via proteomic analyses. However, these interactions are more complex than we are able to depict in 2-D network maps or through gene expression clustering algorithms. To illustrate this proteomic network complexity, consider a single protein, HER2, whose expression level is currently used to guide breast cancer therapy. HER2 readily forms homodimers or heterodimers due to structural similarity with other RTKs [28,35,170]. Recently, HER2 expression was shown to be a dynamic state in circulating breast cancer cells [98]. HER2 negative cells that were propagated as multicellular colonies produced HER2 positive daughter cells when the colonies reached a cell density of 5–9 cells [98]. These findings substantiate the temporal variation in TNBC mutations during clonal evolution [171], reflecting a yet unknown mechanism of dynamic mutation and receptor expression. Promiscuous dimerization and dynamic expression provide HER2 overexpressing breast cancer cells with multiple options for orchestrating downstream signal transduction cascades leading either to cell proliferation and potential therapy resistance, or alternatively, a lower proliferation rate [28,34,35]. The obvious, clinically relevant questions are: (a) How does the percentage of receptor positive and negative cells change over time?, (b) Does this percentage change during treatment?, (c) If receptor expression is temporal, can the timing of therapy be coordinated to the dynamic receptor expression?, and (d) Are the current ER, PR, and HER2 immunohistochemical expression thresholds truly adequate for guiding therapy decisions? These questions can only be addressed in large, well-designed proteomic-based studies that include resampling the residual tumor following neoadjuvant therapy [172].

Proteomics also provides functional information regarding the state of post-translationally modified proteins and the subcellular location. Many proteins are able to translocate between sub-cellular compartments, with often disparate functions depending on the protein location [108,173]. ER, AR, and EGFR are just three examples of RTKs that translocate from the cytoplasm to the nucleus, where they exhibit enhanced transcriptional activity, ultimately enhancing cell proliferation [173–177]. The FOXO1 transcription factors functions as tumor suppressor in the nucleus [178]. However, phosphorylation causes FOXO to be expelled from the nucleus, thus inactivating its tumor suppressor function [132]. Low levels of phosphorylated FOXO1/O3A have been associated with pathologic complete response in HER2+ breast cancer patients treated with anti-HER2 therapy (trastuzumab/lapatinib) [108]. From these examples, one can easily appreciate the contribution that IHC, mass spectrometry, and RPPA have in defining the functional state of cells. This suite of proteomic technologies illuminates the subcellular location, identifies the proteins, and provides quantitative expression of complex downstream signaling events, none of which are knowable from gene expression levels.

Despite our progress in deciphering breast tumor heterogeneity, critical analysis of pre-clinical and clinical trial results, with special emphasis on the study design and

methodologies, are required to interpret the findings within the limitations of the study. To realize the promise of precision medicine, further research funding should be allocated to understand the molecular characteristics for the minority of patients that either are therapy responders or in whom therapy fails. Critically important biological information is waiting to be discovered in these 'outlier' patient populations. The key questions we should be asking are: What are the biological differences in the tumor and the tumor microenvironment between the responders and non-responders? and How do we harness this information to optimize treatment for all patients, at diagnosis, recurrence, and if metastasis occurs?

Discerning biological underpinnings of breast cancer requires the scientific and clinical communities to be aware of numerous factors that can potentially introduce bias in biomarker studies. Firstly, the widespread availability of large genomic and proteomic data sets must be evaluated not only on the data content, but also on specimen collection and pre-analytical variables that directly impact data quality [8,160,165,179]. Large data sets such as TCGA [180] and METABRIC [9,10], provide well-annotated breast cancer data sets and provide opportunities to reanalyze the data using different methods and hypotheses. However, even in large, international specimen cohorts, pre-analytical variables are under-appreciated regarding their effect on protein and phosphoprotein levels. Fixation conditions (formalin or freezing), cold ischemia time, and size of the tissue contribute to pre-analytical variability [160,163,166,181]. Furthermore, interpreting any breast cancer study requires critical evaluation regarding the number of specimens from each histomorphology category (Figure 1), race, and age. Each of these factors contributes to biological differences in breast cancer subtypes.

Molecular characterization of tumors, and specifically of tumor subtypes, requires enrichment of tumor cell populations [182,183]. Enrichment techniques drastically reduce the number of non-tumor cells, which can mask or confound the underlying proteomic biology [179,184–186]. Macrodissection may be performed manually using needles or scalpels. Laser capture microdissection, with dedicated instruments (available from ThermoFisher, Carl Zeiss, Molecular Machines & Industries, or Leica), enables precise selection of specific cell populations, including tumor, stroma, normal adjacent tissue, and immune infiltrates [182,183]. Enrichment of specific cell populations also allows direct comparison of the tumor to other tissue types. Numerous studies have demonstrated the lack of concordance in gene expression and proteomic signal transduction kinases between enriched and non-enriched cell populations [179,184,185,187]. Thus, even publicly available, well-annotated data sets should be assessed with the knowledge that homogenized tissue, without *a priori* cell population enrichment, may lack low abundance biomarkers or have notable molecular differences compared to tumor enriched specimens [8,187]. As additional proof, microdissected tumors have been shown to be highly concordant with the underlying genomic alterations [184,187]. Molecular characterization of specimens that have been macro or microdissected to enrich the tumor, stroma, and immune cell populations will undoubtedly yield more informative biological insights for predictive and therapeutic patient stratification [8].

Secondly, analytical methodologies must be accurately communicated with adequate details for interpreting the data in the context of the patient population, sample handling and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

preparation, and analysis [69,188,189]. Meta-analysis and systematic reviews of numerous, large published datasets offer unprecedented opportunities for comparing and contrasting studies, identifying new biological information, and deciphering patterns or discordant results within the study data. Nonetheless, issues can arise with meta-analysis and systematic reviews, potentially leading to misinterpretation of data. Examples of issues that may be encountered include failure to account for pre-analytical variables such as using homogenized tissue and stroma [8], variations in cold ischemia times [190], over-interpretation of data [191], bias in interpretation due to mistakes in the meta-analysis [192], or differences in analytical procedures prior to the publication of standard operating procedures [16], for example, non- standardized staining protocols for ER and PR prior to the 2010 publication of ASCO/CAP guidelines for IHC [25]. The same lack of standardized guidelines applies to HER2 testing prior to 2007 [159]. Updated HER2 testing guidelines were implemented in 2014 [36].

Thirdly, funding is urgently needed to systematically validate promising new biomarkers in large, well-annotated specimen cohorts which were collected and analyzed using appropriate standardized protocols. Biomarkers need to be evaluated in cohorts of patients that are both biomarker-positive and biomarker-negative. Inclusion of both populations, and control subjects with co-morbidities, in addition to healthy controls, can potentially improve approval rates by regulatory agencies such as the United States Food and Drug Administration (FDA) for “treatment-by-biomarker” compounds and companion diagnostic assays [193].

Two pressing unmet needs currently limit the widespread adoption of proteomic analysis in routine clinical patient management: (1) the lack of proteomic population data for normal and/or breast tumor cells, particularly for low abundance proteins and post-translationally modified proteins [194], and (2) computational tools for comparing individual patient’s proteomic data with population data. Recently, a genomic analysis tool (iCAGES) has been devised for ranking driver genes for individual patients [195]. To optimize delivery of precision medicine, physicians and scientists urgently need similar tools for ranking and prioritizing the functional status of proteins and protein pathways in individual patients.

9. Five-year view

Complete -omic integration, the true embodiment of precision medicine, looms on the horizon. Proteomic technologies are poised to be adopted into routine clinical practice, as clinical laboratory assays for patient management. Digital pathology and image recognition software enhancements reduce variability in inter-pathologist scoring of immunohistochemical stains. Despite improvements in proteomic assay sensitivity and reliability, clinical diagnostics is not keeping pace with biomarker research and discovery. For example, only ERα, and not ERβ, is routinely measured for characterizing the breast cancer subtype. Estrogen receptor isoforms ERα and ERβ exhibit different physiological outcomes and signaling interactions. ERβ protein expression potentially holds diagnostic and prognostic information that could help identify hormone resistant subclones to further tailor hormonal therapy options. Another understudied area is immune cell signaling in breast cancer subtypes. We often see large zones of stromal infiltrating immune cells

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

surrounding areas of DCIS. How do these immune cells influence the emergence or repression of invasive breast cancer? Does the immune cell composition reflect the breast cancer subtype? Hopefully, as multiplex phenotyping technologies evolve, the predictive and prognostic contribution of hormone receptor isoforms and immune cell phenotype will be realized [117,196].

In the near future, further integrative analyses including glycomic [197–202], metabolomic [203–205] and proteomic biomarkers, could yield wholly new types and combinations of biomarkers which reflect the biological state of the tumor microenvironment. Glycomic analysis has already highlighted glycoprotein profile differences between basal and luminal breast cancer cell lines [202] and the presence of high-mannose content glycans in mouse xenograft specimens [200].

Within five years, mass spectrometry-based proteomic assays will become clinically accepted companion diagnostic tests aiding breast cancer diagnosis, prognosis, and therapy decisions similar to IHC. Mass spectrometry biomarker validation, based on multiple reaction monitoring (MRM), provides antibody-free confirmation of protein/peptide identity. MRM assays are becoming more common and are readily adoptable in clinical laboratories.

Additional xenograft models, in a variety of species, such as canines, will aid pre-clinical investigation for studying breast cancer subtypes [206,207]. Shinoda et al. have shown prognostic significance of ER, HER2 and caveolin-1 IHC in canine mammary gland tumors, which mirrors prognostic data for humans [208]. An example of improved animal modelling is the currently available xenograft model of ER+ breast adenocarcinoma that maintains an intact microenvironment [209]. This model recapitulates breast architecture which is critical for decoding spatial and temporal tumor-stromal interactions. Patient derived xenograft models can also be used to investigate clonal tumor behavior, both as individual clonal populations or as heterogeneous tumor, thus elucidating optimal treatment schedules to eliminate all tumor subclones (Figure 1).

A new niche for proteomic assays will be in the area of biosimilar therapies. A biosimilar product is interchangeable with a FDA approved biological product if it can be shown that it is highly similar to an approved product [210]. However, ‘highly similar’ is not equivalent to identical and not all biosimilar products are interchangeable with their corresponding FDA-approved product. Biosimilar products may possess variant protein substitutions, such as glycosylation, which could potentially modulate their intended effect. These functional effects cannot be discerned via genomics or structural analysis of the compound. Proteomic-based assays will be essential for confirming the on and off target effects of biosimilar breast cancer therapies. breast ca

Acknowledgments

Funding

This work was supported in part by George Mason University, and the National Institutes of Health Innovative Molecular Analysis Technologies (IMAT) program grant 1R33CA20693701, “Protein painting identifies therapeutic targets at protein-protein interfaces” to Lance Liotta. The funding sources did not have any role in the study design; the collection, analysis an interpretation of data; manuscript preparation; or the decision to submit the paper for publication. We thank Paul Russo, George Mason University, for mass spectrometry analysis, and Lance

Liotta for editorial advice. Conflict of interest statement: CM, VE and AH receive salary support from the National Institutes of Health grant 1R33CA20693701. VE is an inventor of the protein painting technology and, as a university employee, may receive patent royalties per university policies.

References

Papers of special note have been highlighted as:

* of interest

** of considerable interest

1. Anderson WF, Rosenberg PS, Prat A, et al. How many etiological subtypes of breast cancer. Journal of the National Cancer Institute. 2014;106 [PubMed: 25174031]
2. Dieci MV, Orvieto E, Dominici M, et al. Rare breast cancer subtypes. The Oncologist. 2014;19:805–13. [PubMed: 24969162]
3. Malhotra GK, Zhao X, Band H, et al. Histological, molecular and functional subtypes of breast cancers. Cancer Biol Ther. 2010;10(10):955–60. [PubMed: 21057215] *Provides very good descriptions of breast histomorphology.
4. Vasconcelos I, Hussainzada A, Berger S, et al. The St. Gallen surrogate classification for breast cancer subtypes successfully predicts tumor presenting features, nodal involvement, recurrence patterns and disease free survival. Breast. 2016;29:181–5. [PubMed: 27544822]
5. Weigelt B, Baehner FL, Reis-Filho JS. The contribution of gene expression profiling to breast cancer classification, prognostication and prediction. The Journal of Pathology. 2010;220:263–80. [PubMed: 19927298]
6. Ciriello G, Gatza ML, Beck AH, et al. Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. Cell. 2015;163:506–19. [PubMed: 26451490]
7. Network CGA. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490(7 8):61–70. [PubMed: 23000897]
8. Network CGAR, Weinstein JN,Collisson EA, et al. The Cancer Genome Atlas Pan Cancer analysis project. Nature Genetics. 2013;45:1113–20. [PubMed: 24071849]
9. Curtis C, Shah SP, Chin SF, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012;486(7403):346–52. [PubMed: 22522925]
10. Milioli HH, Vimieiro R, Riveros C, et al. The Discovery of Novel Biomarkers Improves Breast Cancer Intrinsic Subtype Prediction and Reconciles the Labels in the METABRIC Data Set. PloS One. 2015;10:e0129711. [PubMed: 26132585]
11. Weigelt B, Geyer FC, Reis-Filho JS. Histological types of breast cancer. Molecular Oncology. 2010;4:192–208. [PubMed: 20452298]
12. Allred DC, Harvey JM, Berardo M, et al. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. Mod Pathol. 1998;11(2):155–68. [PubMed: 9504686]
13. Brotherick I, Robson CN, Browell DA, et al. Cytokeratin expression in breast cancer: phenotypic changes associated with disease progression. Cytometry. 1998;32(4):301–8. [PubMed: 9701399]
14. Cui X, Schiff R, Arpino G, et al. Biology of Progesterone Receptor Loss in Breast Cancer and Its Implications for Endocrine Therapy. Journal of Clinical Oncolgy. 2005;23:7721–35.
15. Harvey JM, Clark GM, Osborne CK, et al. Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. J Clin Oncol. 1999;17(5):1474–81. [PubMed: 10334533]
16. Hill DA, Barry M, Wiggins C, et al. Estrogen receptor quantitative measures and breast cancer survival. Breast Cancer Res Treat. 2017.
17. Korsching E, Packeisen J, Agelopoulos K, et al. Cytogenetic alterations and cytokeratin expression patterns in breast cancer: integrating a new model of breast differentiation into cytogenetic pathways of breast carcinogenesis. Lab Invest. 2002;82(11):1525–33. [PubMed: 12429812]
18. Duffy MJ, Harbeck N, Nap M, et al. Clinical use of biomarkers in breast cancer: Updated guidelines from the European Group on Tumor Markers (EGTM). Eur J Cancer. 2017;75:284– 98. [PubMed: 28259011]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

19. Arpino G, Wiechmann L, Osborne CK, et al. Crosstalk between the Estrogen Receptor and the HER Tyrosine Kinase Receptor Family. Endocrine Reviews. 2008;29:217–33. [PubMed: 18216219] ** Documents receptor cross-talk and implications for prognosis and treatment.
20. Powell E, Shanle E, Brinkman A, et al. Identification of estrogen receptor dimer selective ligands reveals growth-inhibitory effects on cells that co-express ERalpha and ERbeta. PLoS One. 2012;7(2):e30993. [PubMed: 22347418]
21. Nelson AW, Groen AJ, Miller JL, et al. Comprehensive assessment of estrogen receptor beta antibodies in cancer cell line models and tissue reveals critical limitations in reagent specificity. Mol Cell Endocrinol. 2017;440:138–50. [PubMed: 27889472] *Provides an essential comparison of ERβ antibodies for accurate IHC results.
22. Mangelsdorf DJ, Thummel C, Beato M, et al. The nuclear receptor superfamily: the second decade. Cell. 1995;83(6):835–9. [PubMed: 8521507]
23. Hahnel R, Woodings T, Vivian AB. Prognostic value of estrogen receptors in primary breast cancer. Cancer. 1979;44(2):671–5. [PubMed: 476576]
24. Phillips T, Murray G, Wakamiya K, et al. Development of standard estrogen and progesterone receptor immunohistochemical assays for selection of patients for antihormonal therapy. Appl Immunohistochem Mol Morphol. 2007;15(3):325–31. [PubMed: 17721279]
25. Hammond ME, Hayes DF, Dowsett M, et al. American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. J Clin Oncol. 2010;28(16):2784–95. [PubMed: 20404251] **Provides essential guidelines for ER/PR immunohistochemistry testing and interpretation.
26. Norum JH, Andersen K, Sorlie T. Lessons learned from the intrinsic subtypes of breast cancer in the quest for precision therapy. Br J Surg. 2014;101(8):925–38. [PubMed: 24849143]
27. Coussens L, Yang-Feng TL, Liao YC, et al. Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with neu oncogene. Science (New York, NY). 1985;230:1132–9.
28. Moasser MM. The oncogene HER2: its signaling and transforming functions and its role in human cancer pathogenesis. Oncogene. 2007;26(45):6469–87. [PubMed: 17471238]
29. Schechter AL, Hung M- C, Vaidyanathan L, et al. The neu Gene. Science. 1985;229:976–8. [PubMed: 2992090]
30. King CR, Kraus MH, Aaronson SA. Amplification of a novel v-erbB-related gene in a human mammary carcinoma. Science (New York, NY). 1985;229:974–6.
31. Slamon DJ, Clark GM, Wong SG, et al. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science. 1987;235(4785):177–82. [PubMed: 3798106]
32. Ciocca DR, Fujimura FK, Tandon AK, et al. Correlation of HER-2/neu amplification with expression and with other prognostic factors in 1103 breast cancers. J Natl Cancer Inst. 1992;84(16):1279–82. [PubMed: 1353538]
33. Duffy MJ, O'Donovan N, McDermott E, et al. Validated biomarkers: The key to precision treatment in patients with breast cancer. Breast. 2016;29:192–201. [PubMed: 27521224] *Provides an excellent perspective on biomarkers and issues with validation.
34. Ross JS, Fletcher JA, Bloom KJ, et al. Targeted therapy in breast cancer: the HER- 2/neu gene and protein. Mol Cell Proteomics. 2004;3(4):379–98. [PubMed: 14762215]
35. Rubin I, Yarden Y. The basic biology of HER2. Ann Oncol. 2001;12 Suppl 1:S3–8.
36. Wolff AC, Hammond ME, Hicks DG, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. Arch Pathol Lab Med. 2014;138(2):241–56. [PubMed: 24099077] **Updated guidelines for HER2 immunohistochemistry testing and analysis
37. Harris LN, Ismaila N, McShane LM, et al. Use of Biomarkers to Guide Decisions on Adjuvant Systemic Therapy for Women With Early-Stage Invasive Breast Cancer: American Society of Clinical Oncology Clinical Practice Guideline. J Clin Oncol. 2016;34(10):1134–50. [PubMed: 26858339]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

38. Northridge ME, Rhoads GG, Wartenberg D, et al. The importance of histologic type on breast cancer survival. J Clin Epidemiol. 1997;50(3):283–90. [PubMed: 9120527]
39. Rakha EA, Lee AH, Evans AJ, et al. Tubular carcinoma of the breast: further evidence to support its excellent prognosis. J Clin Oncol. 2010;28(1):99–104. [PubMed: 19917872]
40. Perou CM, Scrlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature. 2000;406:747–52. [PubMed: 10963602]
41. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A. 2001;98(19):10869–74. [PubMed: 11553815]
42. Nagle RB, Bocker W, Davis JR, et al. Characterization of breast carcinomas by two monoclonal antibodies distinguishing myoepithelial from luminal epithelial cells. J Histochem Cytochem. 1986;34(7):869–81. [PubMed: 2423579]
43. Sotiriou C, Neo SY, McShane LM, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proc Natl Acad Sci U S A. 2003; 100(18): 10393–8. [PubMed: 12917485]
44. Parker JS, Mullins M, Cheang MCU, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology. 2009;27:1160–7. [PubMed: 19204204] *PAM50 intrinsic breast cancer classification algorithm
45. Kuijer A, King TA. Age, molecular subtypes and local therapy decision-making. The Breast. 2017;34:S70–S7. [PubMed: 28666919]
46. Anders CK, Deal AM, Miller CR, et al. The prognostic contribution of clinical breast cancer subtype, age, and race among patients with breast cancer brain metastases. Cancer. 2011;117(8): 1602–11. [PubMed: 21472708]
47. Morrison DH, Rahardja D, King E, et al. Tumour biomarker expression relative to age and molecular subtypes of invasive breast cancer. British Journal of Cancer. 2012;107:382–7. [PubMed: 22713661]
48. Carey LA, Perou CM, Livasy CA, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA. 2006;295:2492–502. [PubMed: 16757721]
49. Kaur JS, Vierkant RA, Hobday T, et al. Regional differences in breast cancer biomarkers in american Indian and Alaska native women. Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology. 2014;23:409–15.
50. Kast K, Link T, Friedrich K, et al. Impact of breast cancer subtypes and patterns of metastasis on outcome. Breast Cancer Research and Treatment. 2015;150:621–9. [PubMed: 25783184]
51. Lipton A, Goodman L, Leitzel K, et al. HER3, p95HER2, and HER2 protein expression levels define multiple subtypes of HER2-positive metastatic breast cancer. Breast Cancer Research and Treatment. 2013;141:43–53. [PubMed: 23959396] **Highlights the issue of hyperactive HER2 signaling with truncated HER2 receptors.
52. Savci-Heijink CD, Halfwerk H, Hooijer GKJ, et al. Retrospective analysis of metastatic behaviour of breast cancer subtypes. Breast Cancer Research and Treatment. 2015;150:547– 57. [PubMed: 25820592]
53. Smid M, Wang Y, Zhang Y, et al. Subtypes of breast cancer show preferential site of relapse. Cancer Research. 2008;68:3108–14. [PubMed: 18451135]
54. Bottai G, Raschioni C, Losurdo A, et al. An immune stratification reveals a subset of PD-1/LAG-3 double-positive triple-negative breast cancers. Breast cancer research: BCR. 2016;18:121. [PubMed: 27912781]
55. Burugu S, Asleh-Aburaya K, Nielsen TO. Immune infiltrates in the breast cancer microenvironment. Breast Cancer (Tokyo, Japan). 2017;24:3–15.
56. Luen S, Virassamy B, Savas P, et al. The genomic landscape of breast cancer and its interaction with host immunity. Breast. 2016;29:241–50. [PubMed: 27481651]
57. Nawaz S, Heindl A, Koelble K, et al. Beyond immune density: critical role of spatial heterogeneity in estrogen receptor-negative breast cancer. Mod Pathol. 2015;28(6):766–77. [PubMed: 25720324]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

58. Zografos E, Gazouli M, Tsangaris G, et al. The Significance of Proteomic Biomarkers in Male Breast Cancer. Cancer Genomics Proteomics. 2016;13(3):183–90. [PubMed: 27107060]
59. Piscuoglio S, Ng CKY, Murray MP, et al. The Genomic Landscape of Male Breast Cancers. Clinical Cancer Research: An Official Journal of the American Association for Cancer Research. 2016;22:4045–56. [PubMed: 26960396]
60. Kornegoor R, Verschuur-Maes AH, Buerger H, et al. Immunophenotyping of male breast cancer. Histopathology. 2012;61(6):1145–55. [PubMed: 22958056]
61. Kornegoor R, Verschuur-Maes AHJ, Buerger H, et al. Molecular subtyping of male breast cancer by immunohistochemistry. Modern Pathology: An Official Journal of the United States and Canadian Academy of Pathology, Inc. 2012;25:398–404.
62. Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer Mol Oncol.2011;5(1):5–23 [PubMed: 21147047]
63. Russnes HG, Lingjærde OC, Børresen-Dale A-L, et al. Breast Cancer Stratification. The American Journal of Pathology. 2017.
64. Place AE, Jin Huh S, Polyak K. The microenvironment in breast cancer progression: biology and implications for treatment. Breast Cancer Res. 2011;13(6):227. [PubMed: 22078026]
65. Polyak K Heterogeneity in breast cancer. J Clin Invest. 2011;121(10):3786–8. [PubMed: 21965334]
66. McFarland CD, Mirny LA, Korolev KS. Tug-of-war between driver and passenger mutations in cancer and other adaptive processes. Proc Natl Acad Sci U S A. 2014;111(42):15138–43. [PubMed: 25277973]
67. Yates LR. Intratumoral heterogeneity and subclonal diversification of early breast cancer. Breast. 2017;34 Suppl 1:S36–S42.
68. Letai A Functional precision cancer medicine-moving beyond pure genomics. Nat Med. 2017;23(9):1028–35. [PubMed: 28886003] *Timely and accurate assessment of the state of precision medicine and how proteomics can contribute essential biological information.
69. Baker MS, Ahn SB, Mohamedali A, et al. Accelerating the search for the missing proteins in the human proteome. Nat Commun. 2017;8:14271. [PubMed: 28117396]
70. Wong SY, Hynes RO. Lymphatic or hematogenous dissemination: how does a metastatic tumor cell decide? Cell Cycle. 2006;5(8):812–7. [PubMed: 16627996]
71. Akbani R, Ng PKS, Werner HMJ, et al. A pan-cancer proteomic perspective on The Cancer Genome Atlas. Nature Communications. 2014;5:3887.
72. Orton DJ, Doucette AA. Proteomic Workflows for Biomarker Identification Using Mass Spectrometry - Technical and Statistical Considerations during Initial Discovery. Proteomes. 2013;1(2):109–27. [PubMed: 28250400]
73. Neve RM, Chin K, Fridlyand J, et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell. 2006;10(6):515–27. [PubMed: 17157791] *Provides western blot results for a vareity of proteins and breast cancer cell lines.
74. Vranic S, Gatalica Z, Wang ZY. Update on the molecular profile of the MDA-MB-453 cell line as a model for apocrine breast carcinoma studies. Oncol Lett. 2011 ;2(6):1131–7. [PubMed: 22121396] *This paper highlights important caveats of using cell lines for biomarker studies.
75. Jiang G, Zhang S, Yazdanparast A, et al. Comprehensive comparison of molecular portraits between cell lines and tumors in breast cancer. BMC Genomics. 2016; 17. [PubMed: 26725242]
76. Chan D, Lee S-C. Serial Changes in Expression of Proteins in Response to Neoadjuvant Chemotherapy in Breast Cancer. 2013 9 26, 2017 In: Oncogene and Cancer – From Bench to Clinic [Internet]. In Tech . Availabel from: https://www.intechopen.com/books/oncoqene-and-cancer-from-bench-to-clinic/serial-chanqes-in-expression-of-proteins-in-response-to-neoadiuvant-chemotherapy-in-breast-cancer.
77. Karp NA, Lilley KS. Design and analysis issues in quantitative proteomics studies. Proteomics. 2007;7 Suppl 1:42–50. [PubMed: 17893850]
78. Ellis MJ. Lessons in precision oncology from neoadjuvant endocrine therapy trials in ER+ breast cancer. The Breast. 2017;34:S104–S7. [PubMed: 28669712]
79. Ellis MJ, Suman VJ, Hoog J, et al. Ki67 Proliferation Index as a Tool for Chemotherapy Decisions During and After Neoadjuvant Aromatase Inhibitor Treatment of Breast Cancer: Results From the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

American College of Surgeons Oncology Group Z1031 Trial (Alliance). J Clin Oncol. 2017;35(10): 1061–9. [PubMed: 28045625]

80. Cuzick J, Dowsett M, Pineda S, et al. Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the Genomic Health recurrence score in early breast cancer. J Clin Oncol. 2011;29(32):4273–8. [PubMed: 21990413] **Describes the IHC4 score.
81. Bartlett JM, Christiansen J, Gustavson M, et al. Validation of the IHC4 Breast Cancer Prognostic Algorithm Using Multiple Approaches on the Multinational TEAM Clinical Trial. Arch Pathol Lab Med. 2016;140(1):66–74. [PubMed: 26717057]
82. Yeo B, Zabaglo L, Hills M, et al. Clinical utility of the IHC4+C score in oestrogen receptor-positive early breast cancer: a prospective decision impact study. Br J Cancer. 2015;113(3):390–5. [PubMed: 26180920]
83. Lakhanpal R, Sestak I, Shadbolt B, et al. IHC4 score plus clinical treatment score predicts locoregional recurrence in early breast cancer. Breast. 2016;29:147–52. [PubMed: 27498128]
84. Lehmann BD, Jovanovi B, Chen X, et al. Refinement of Triple-Negative Breast Cancer Molecular Subtypes. PloS One. 2016;11:e0157368. [PubMed: 27310713]
85. Reis-Filho JS, Tutt AN. Triple negative tumours: a critical review. Histopathology. 2008;52(1): 108–18. [PubMed: 18171422] **An excellent guide to triple negative breast tumors.
86. Maeda T, Nakanishi Y, Hirotani Y, et al. Immunohistochemical co-expression status of cytokeratin 5/6, androgen receptor, and p53 as prognostic factors of adjuvant chemotherapy for triple negative breast cancer. Med Mol Morphol. 2016;49(1):11–21. [PubMed: 26009308]
87. Geyer FC, Pareja F, Weigelt B, et al. The Spectrum of Triple-Negative Breast Diseasse.The American Journal of Pathology. 2017
88. Cuenca-Lopez MD, Montero JC, Morales JC, et al. Phospho-kinase profile of triple negative breast cancer and androgen receptor signaling. BMC Cancer. 2014;14:302. [PubMed: 24779793]
89. Speers C, Zhao SG, Chandler B, et al. Androgen receptor as a mediator and biomarker of radioresistance in triple-negative breast cancer. NPJ Breast Cancer. 2017;3:29. [PubMed: 28840192]
90. Orecchia R Tailoring radiotherapy according to cancer subtypes. The Breast. 2017;34:S91–S4. [PubMed: 28673538]
91. Honeth G, Bendahl PO, Ringner M, et al. The CD44+/CD24- phenotype is enriched in basal-like breast tumors. Breast Cancer Res. 2008;10(3):R53. [PubMed: 18559090]
92. Olsson E, Honeth G, Bendahl P-O, et al. CD44 isoforms are heterogeneously expressed in breast cancer and correlate with tumor subtypes and cancer stem cell markers. BMC cancer. 2011;11:418. [PubMed: 21957977]
93. Paweletz CP, Charboneau L, Bichsel VE, et al. Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. Oncogene. 2001;20(16):1981–9. [PubMed: 11360182]
94. Boyd ZS, Wu QJ, O'Brien C, et al. Proteomic analysis of breast cancer molecular subtypes and biomarkers of response to targeted kinase inhibitors using reverse-phase protein microarrays. Molecular Cancer Therapeutics. 2008;7:3695–706. [PubMed: 19056674]
95. Craig DW, O'Shaughnessy JA, Kiefer JA, et al. Genome and transcriptome sequencing in prospective metastatic triple-negative breast cancer uncovers therapeutic vulnerabilities. Mol Cancer Ther. 2013;12(1):104–16. [PubMed: 23171949]
96. Montero JC, Esparis-Ogando A, Re-Louhau MF, et al. Active kinase profiling, genetic and pharmacological data define mTOR as an important common target in triple-negative breast cancer. Oncogene. 2014;33(2):148–56. [PubMed: 23246963]
97. Masuda H, Qi Y, Liu S, et al. Reverse phase protein array identification of triple-negative breast cancer subtypes and comparison with mRNA molecular subtypes. Oncotarget. 2017;8:70481–95. [PubMed: 29050296]
98. Jordan NV, Bardia A, Wittner BS, et al. HER2 expression identifies dynamic functional states within circulating breast cancer cells. Nature. 2016;537(7618):102–6. [PubMed: 27556950]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

99. Wulfkuhle JD, Berg D, Wolff C, et al. Molecular analysis of HER2 signaling in human breast cancer by functional protein pathway activation mapping. Clinical Cancer Research: An Official Journal of the American Association for Cancer Research. 2012;18:6426–35. [PubMed: 23045247]
100. Panis C, Pizzatti L, Correa S, et al. The positive is inside the negative. Cancer Letters. 2015;357:186–95. [PubMed: 25434795]
101. Network CGA. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490:61–70. [PubMed: 23000897]
102. Aure MR, Vitelli V, Jernstrom S, et al. Integrative clustering reveals a novel split in the luminal A subtype of breast cancer with impact on outcome. Breast cancer research: BCR. 2017;19:44. [PubMed: 28356166]
103. Sonntag J, Bender C, Soons Z, et al. Reverse phase protein array based tumor profiling identifies a biomarker signature for risk classification of hormone receptro-positive breast cancer. Translational Proteomics. 2014;2:52–9.
104. Chavez-Macgregor M, Liu S, De Melo-Gagliato D, et al. Differences in gene and protein expression and the effects of race/ethnicity on breast cancer subtypes. Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology. 2014;23:316–23.
105. Michaut M, Chin S-F, Majewski I, et al. Integration of genomic, transcriptomic and proteomic data identifies two biologically distinct subtypes of invasive lobular breast cancer. Scientific Reports. 2016;6:18517. [PubMed: 26729235] **Excellent example of precision medicine applying multiple methods and data analysis tools t identify subtypes within lobular breast cancer.
106. Hisamatsu Y, Tokunaga E, Yamashita N, et al. Impact of FOXA1 expression on the prognosis of patients with hormone receptor-positive breast cancer. Annals of Surgical Oncology. 2012;19:1145–52. [PubMed: 21984487]
107. Hisamatsu Y, Tokunaga E, Yamashita N, et al. Impact of GATA-3 and FOXA1 expression in patients with hormone receptor-positive/HER2-negative breast cancer. Breast Cancer (Tokyo, Japan). 2015;22:520–8.
108. Holmes FA, Espina V, Liotta LA, et al. Pathologic complete response after preoperative anti-HER2 therapy correlates with alterations in PTEN, FOXO, phosphorylated Stat5, and autophagy protein signaling. BMC Res Notes. 2013;6:507. [PubMed: 24304724]
109. Wu Y, Shang X, Sarkissyan M, et al. FOXO1A is a target for HER2-overexpressing breast tumors. Cancer Res. 2010;70(13):5475–85. [PubMed: 20551062] **This paper describes the mechansim of FOXO nuclear-cytoplasmic translocation and its biological implications
110. Cha S, Imielinski MB, Rejtar T, et al. In situ proteomic analysis of human breast cancer epithelial cells using laser capture microdissection. Molecular & cellular proteomics: MCP. 2010;9:2529–44. [PubMed: 20739354] **Excellent example of proteomic study design using microdissected tissue and work-flow integrating genomics and proteomics
111. Dillon LM, Bean JR, Yang W, et al. P-REX1 creates a positive feedback loop to activate growth factor receptor, PI3K/AKT and MEK/ERK signaling in breast cancer. Oncogene. 2015;34:3968–76. [PubMed: 25284585]
112. Li J, Lu Y, Akbani R, et al. TCPA: a resource for cancer functional proteomics data. Nat Methods. 2013;10(11):1046–7.
113. Lee YH, Liu X, Qiu F, et al. HP1beta is a biomarker for breast cancer prognosis and PARP inhibitor therapy. PLoS One. 2015;10(3):e0121207. [PubMed: 25769025]
114. Chatterjee S, Malhotra R, Varghese F, et al. Quantitative immunohistochemical analysis reveals association between sodium iodide symporter and estrogen receptor expression in breast cancer. PLoS One. 2013;8(1):e54055. [PubMed: 23342072]
115. Agboola AJ, Paish EC, Rakha EA, et al. EpCAM expression is an indicator of recurrence in basal-like breast cancer. Breast Cancer Research and Treatment. 2012;133:575– 82. [PubMed: 22002520]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

116. Soysal SD, Muenst S, Barbie T, et al. EpCAM expression varies significantly and is differentially associated with prognosis in the luminal B HER2(+), basal-like, and HER2 intrinsic subtypes of breast cancer. British Journal of Cancer. 2013;108:1480–7. [PubMed: 23519058]

117. Brodsky AS, Xiong J, Yang D, et al. Identification of stromal ColXalpha1 and tumor- infiltrating lymphocytes as putative predictive markers of neoadjuvant therapy in estrogen receptor-positive/HER2-positive breast cancer. BMC Cancer. 2016;16:274. [PubMed: 27090210]

118. Celis JE, Gromov P, Cabezon T, et al. Proteomic characterization of the interstitial fluid perfusing the breast tumor microenvironment: a novel resource for biomarker and therapeutic target discovery. Mol Cell Proteomics. 2004;3(4):327–44. [PubMed: 14754989]

119. Celis JE, Gromov P, Cabezon T, et al. 15-prostaglandin dehydrogenase expression alone or in combination with ACSM1 defines a subgroup of the apocrine molecular subtype of breast carcinoma. Molecular & cellular proteomics: MCP. 2008;7:1795–809. [PubMed: 18632593]

120. Celis JE, Gromov P, Moreira JM, et al. Apocrine cysts of the breast: biomarkers, origin, enlargement, and relation with cancer phenotype. Mol Cell Proteomics. 2006;5(3):462–83. [PubMed: 16316978]

121. Gromov P, Espinoza JA, Talman ML, et al. FABP7 and HMGCS2 are novel protein markers for apocrine differentiation categorizing apocrine carcinoma of the breast. PLoS One. 2014;9(11):e112024. [PubMed: 25389781]

122. Gongalves A, Charafe-Jauffret E, Bertucci F, et al. Protein profiling of human breast tumor cells identifies novel biomarkers associated with molecular subtypes. Molecular & cellular proteomics: MCP. 2008;7:1420–33. [PubMed: 18426791]

123. Bateman NW, Sun M, Hood BL, et al. Defining central themes in breast cancer biology by differential proteomics. Journal of Proteome Research. 2010;9:5311–24. [PubMed: 20681588]

124. Westhoff MA, Serrels B, Fincham VJ, et al. SRC-mediated phosphorylation of focal adhesion kinase couples actin and adhesion dynamics to survival signaling. Mol Cell Biol. 2004;24(18):8113–33. [PubMed: 15340073]

125. Mitra SK, Hanson DA, Schlaepfer DD. Focal adhesion kinase: in command and control of cell motility. Nat Rev Mol Cell Biol. 2005;6(1):56–68. [PubMed: 15688067]

126. Menon R, Im H, Zhang EY, et al. Distinct splice variants and pathway enrichment in the cell-line models of aggressive human breast cancer subtypes. Journal of Proteome Research. 2014;13:212–27. [PubMed: 24111759] **Description of proteomic workflow for identifying splice

127. Campone M, Valo I, Jezequel P, et al. Prediction of Recurrence and Survival for Triple-Negative Breast Cancer (TNBC) by a Protein Signature in Tissue Samples. Mol Cell Proteomics. 2015;14(11):2936–46. [PubMed: 26209610]

128. Gamez-Pozo A, Trilla-Fuertes L, Prado-Vazquez G, et al. Prediction of adjuvant chemotherapy response in triple negative breast cancer with discovery and targeted proteomics. PLoS One. 2017;12(6):e0178296. [PubMed: 28594844]

129. Young CD, Zimmerman LJ, Hoshino D, et al. Activating PIK3CA Mutations Induce an Epidermal Growth Factor Receptor (EGFR)/Extracellular Signal-regulated Kinase (ERK) Paracrine Signaling Axis in Basal-like Breast Cancer. Molecular & cellular proteomics: MCP. 2015;14:1959–76. [PubMed: 25953087] **Integrative study shows relationship between relevant breast cancer mutations and signal transduction.

130. Baselga J, Gomez P, Greil R, et al. Randomized phase II study of the anti-epidermal growth factor receptor monoclonal antibody cetuximab with cisplatin versus cisplatin alone in patients with metastatic triple-negative breast cancer. J Clin Oncol. 2013;31(20):2586–92. [PubMed: 23733761]

131. Carey LA, Rugo HS, Marcom PK, et al. TBCRC 001: randomized phase II study of cetuximab in combination with carboplatin in stage IV triple-negative breast cancer. J Clin Oncol. 2012;30(21):2615–23. [PubMed: 22665533]

132. Wu X, Zahari MS, Ma B, et al. Global phosphotyrosine survey in triple-negative breast cancer reveals activation of multiple tyrosine kinase signaling pathways. Oncotarget. 2015;6:29143–60. [PubMed: 26356563]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

133. Wu X, Liu X, Koul S, et al. AXL kinase as a novel target for cancer therapy. Oncotarget. 2014;5(20):9546–63. [PubMed: 25337673]
134. Lawrence RT, Perez EM, Hernandez D, et al. The proteomic landscape of triple-negative breast cancer. Cell Reports. 2015;11:630–44. [PubMed: 25892236]
135. Keen JC. Breast cancer cell line development and authentication 2011 In: Breast Cancer - Focusing Tumor Microenvironment, Stem Cells and Metastasis [Internet]. InTech. Available from: https://www.intechopen.com/books/breast-cancer-focusing-tumor-microenvironment-stem-cells-and-metastasis/breast-cancer-cell-line-development-and-authentication.
136. Ponten F, Jirstrom K, Uhlen M. The Human Protein Atlas--a tool for pathology. J Pathol. 2008;216(4):387–93. [PubMed: 18853439]
137. Uhlen M, Bjorling E, Agaton C, et al. A human protein atlas for normal and cancer tissues based on antibody proteomics. Mol Cell Proteomics. 2005;4(12):1920–32. [PubMed: 16127175]
138. Panis C, Pizzatti L, Herrera AC, et al. Label-free proteomic analysis of breast cancer molecular subtypes. Journal of Proteome Research. 2014;13:4752–72. [PubMed: 25221861]
139. Waldemarson S, Kurbasic E, Krogh M, et al. Proteomic analysis of breast tumors confirms the mRNA intrinsic molecular subtypes using different classifiers. Breast cancer research: BCR. 2016;18:69. [PubMed: 27357824]
140. Ntai I, LeDuc RD, Fellers RT, et al. Integrated Bottom-Up and Top-Down Proteomics of Patient-Derived Breast Tumor Xenografts. Molecular & cellular proteomics: MCP. 2016;15:45– 56. [PubMed: 26503891]
141. Geiger T, Cox J, Ostasiewicz P, et al. Super-SILAC mix for quantitative proteomics of human tumor tissue. Nat Methods. 2010;7(5):383–5. [PubMed: 20364148]
142. Tyanova S, Albrechtsen R, Kronqvist P, et al. Proteomic maps of breast cancer subtypes. Nature Communications. 2016;7:10259.
143. Pozniak Y, Balint-Lahat N, Rudolph JD, et al. System-wide Clinical Proteomics of Breast Cancer Reveals Global Remodeling of Tissue Homeostasis. Cell Syst. 2016;2(3):172–84. [PubMed: 27135363] *Quantitiative analysis of proteins in tumors and lymph node metastatis
144. Liu NQ, Stingl C, Look MP, et al. Comparative proteome analysis revealing an 11- protein signature for aggressive triple-negative breast cancer. J Natl Cancer Inst. 2014;106(2):djt376.
145. Gamez-Pozo A, Trilla-Fuertes L, Berges-Soria J, et al. Functional proteomics outlines the complexity of breast cancer molecular subtypes. Sci Rep. 2017;7(1):10100. [PubMed: 28855612]
146. Aittokallio T, Scherer A, Poutanen M, et al. Matched preclinical designs for improved translatability. Sci Transl Med 2017;9(389).
147. Freedman LP, Venugopalan G, Wisman R. Reproducibility2020: Progress and priorities. F1000Res. 2017;6:604. [PubMed: 28620458]
148. Kim MS, Pinto SM, Getnet D, et al. A draft map of the human proteome. Nature. 2014;509(7502):575–81. [PubMed: 24870542]
149. Wilhelm M, Schlegl J, Hahne H, et al. Mass-spectrometry-based draft of the human proteome. Nature. 2014;509(7502):582–7. [PubMed: 24870543]
150. Ezkurdia I, Vazquez J, Valencia A, et al. Analyzing the first drafts of the human proteome. J Proteome Res. 2014;13(8):3854–5. [PubMed: 25014353]
151. Institute CLS. C62-Liquid Chromatography-Mass Spectrometry Methods; Approved Guideline, CLSI document C62-A. 1st ed. Wayne, PA: Clinical Laboratory Standards Institute; 2014 p. 88.
152. Institute CLS. Mass Spectrometry in the Clinical Laboratory: General Principles and Guidance; Approved Guideline, CLSI document C50 1st ed. Wayne, PA: Clinical Laboratory and Standards Institute (CLSI); 2017.
153. Craig R, Cortens JP, Beavis RC. Open source system for analyzing, validating, and storing protein identification data. J Proteome Res. 2004;3(6):1234–42. [PubMed: 15595733]
154. Omenn GS, Lane L, Lundberg EK, et al. Metrics for the Human Proteome Project 2016: Progress on Identifying and Characterizing the Human Proteome, Including Post-Translational Modifications. J Proteome Res. 2016;15(11):3951–60. [PubMed: 27487407]
155. Gaudet P, Michel PA, Zahn-Zabal M, et al. The neXtProt knowledgebase on human proteins: 2017 update. Nucleic Acids Res. 2017;45(D1):D177–D82. [PubMed: 27899619]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

156. Martinez-Bartolome S, Deutsch EW, Binz PA, et al. Guidelines for reporting quantitative mass spectrometry based experiments in proteomics. J Proteomics. 2013;95:84–8 [PubMed: 23500130]
157. Beri J, Rosenblatt MM, Strauss E, et al. Reagent for Evaluating Liquid Chromatography- Tandem Mass Spectrometry (LC-MS/MS) Performance in Bottom-Up Proteomic Experiments. Anal Chem. 2015;87(23):11635–40. [PubMed: 26537636]
158. Bramwell D An introduction to statistical process control in research proteomics. J Proteomics. 2013;95:3–21. [PubMed: 23791708]
159. Wolff AC, Hammond ME, Schwartz JN, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. J Clin Oncol. 2007;25(1):118–45. [PubMed: 17159189]
160. Espina V, Edmiston KH, Heiby M, et al. A portrait of tissue phosphoprotein stability in the clinical tissue procurement process. Mol Cell Proteomics. 2008;7(10):1998–2018. [PubMed: 18667411]
161. Gown AM. Current issues in ER and HER2 testing by IHC in breast cancer. Mod Pathol. 2008;21 Suppl 2:S8–S15. [PubMed: 18660796]
162. Gown AM, Goldstein LC, Barry TS, et al. High concordance between immunohistochemistry and fluorescence in situ hybridization testing for HER2 status in breast cancer requires a normalized IHC scoring system. Mod Pathol. 2008;21(10):1271–7. [PubMed: 18487992]
163. Neumeister VM, Parisi F, England AM, et al. A tissue quality index: an intrinsic control for measurement of effects of preanalytical variables on FFPE tissue. Lab Invest. 2014;94(4):467–74. [PubMed: 24535259]
164. Pinder SE, Provenzano E, Earl H, et al. Laboratory handling and histology reporting of breast specimens from patients who have received neoadjuvant chemotherapy. Histopathology. 2007;50:409–17. [PubMed: 17448015]
165. Unger FT, Lange N, Kruger J, et al. Nanoproteomic analysis of ischemia-dependent changes in signaling protein phosphorylation in colorectal normal and cancer tissue. J Transl Med. 2016;14:6. [PubMed: 26742633]
166. Vassilakopoulou M, Parisi F, Siddiqui S, et al. Preanalytical variables and phosphoepitope expression in FFPE tissue: quantitative epitope assessment after variable cold ischemic time. Lab Invest. 2015;95(3):334–41. [PubMed: 25418580]
167. Varga Z, Noske A. Impact of Modified 2013 ASCO/CAP Guidelines on HER2 Testing in Breast Cancer. One Year Experience. PLoS One. 2015;10(10):e0140652. [PubMed: 26473483]
168. Lynch KL. CLSI C62-A: A New Standard for Clinical Mass Spectrometry. Clin Chem. 2016;62(1):24–9. [PubMed: 26430075]
169. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol. 1982;157(1):105–32. [PubMed: 7108955]
170. Nahta R, Yuan LX, Zhang B, et al. Insulin-like growth factor-I receptor/human epidermal growth factor receptor 2 heterodimerization contributes to trastuzumab resistance of breast cancer cells. Cancer Res. 2005;65(23):11118–28. [PubMed: 16322262]
171. Shah SP, Roth A, Goya R, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature. 2012;486(7403):395–9. [PubMed: 22495314]
172. van de Ven S, Smit VT, Dekker TJ, et al. Discordances in ER, PR and HER2 receptors after neoadjuvant chemotherapy in breast cancer. Cancer Treat Rev. 2011;37(6):422–30. [PubMed: 21177040]
173. Dull AB, George AA, Goncharova EI, et al. Identification of compounds by high-content screening that induce cytoplasmic to nuclear localization of a fluorescent estrogen receptor alpha chimera and exhibit agonist or antagonist activity in vitro. J Biomol Screen. 2014;19(2):242–52. [PubMed: 24051224]
174. Barton VN, D'Amato NC, Gordon MA, et al. Androgen Receptor Biology in Triple Negative Breast Cancer: a Case for Classification as AR+ or Quadruple Negative Disease. Horm Cancer. 2015;6(5–6):206–13. [PubMed: 26201402]
175. Guo Z, Dai B, Jiang T, et al. Regulation of androgen receptor activity by tyrosine phosphorylation. Cancer Cell. 2006;10(4):309–19. [PubMed: 17045208]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

176. Lin SY, Makino K, Xia W, et al. Nuclear localization of EGF receptor and its potential new role as a transcription factor. Nat Cell Biol. 2001;3(9):802–8. [PubMed: 11533659]
177. Wang YN, Hung MC. Nuclear functions and subcellular trafficking mechanisms of the epidermal growth factor receptor family. Cell Biosci. 2012;2(1):13. [PubMed: 22520625]
178. Van Der Heide LP, Hoekman MF, Smidt MP. The ins and outs of FoxO shuttling: mechanisms of FoxO translocation and transcriptional regulation. Biochem J. 2004;380(Pt 2):297–309.
179. Wulfkuhle JD, Speer R, Pierobon M, et al. Multiplexed cell signaling analysis of human breast cancer applications for personalized therapy. Journal of Proteome Research. 2008;7:1508–17. [PubMed: 18257519]
180. Network CGAR Weinstein JN, Collisson EA, et al. The Cancer Genome Atlas PanCancer analysis project. Nat Genet. 2013;45(10):1113–20. [PubMed: 24071849]
181. Hicks DG, Kushner L, McCarthy K. Breast cancer predictive factor testing: the challenges and importance of standardizing tissue handling. J Natl Cancer Inst Monogr. 2011;2011(42):43–5. [PubMed: 21672896]
182. Emmert-Buck MR, Bonner RF, Smith PD, et al. Laser capture microdissection. Science. 1996;274(5289):998–1001. [PubMed: 8875945]
183. Espina V, Wulfkuhle JD, Calvert VS, et al. Laser-capture microdissection. Nat Protoc. 2006;1(2): 586–603. [PubMed: 17406286]
184. Baldelli E, Haura EB, Crinò L, et al. Impact of upfront cellular enrichment by laser capture microdissection on protein and phosphoprotein drug target signaling activation measurements in human lung cancer. Proteomics Clinical Applications. 2015;9:928–37. [PubMed: 25676683]
185. Lin J, Marquardt G, Mullapudi N, et al. Lung cancer transcriptomes refined with laser capture microdissection. The American Journal of Pathology. 2014;184:2868–84. [PubMed: 25128906]
186. Silvestri A, Colombatti A, Calvert VS, et al. Protein pathway biomarker analysis of human cancer reveals requirement for upfront cellular-enrichment processing. Laboratory Investigation; a Journal of Technical Methods and Pathology. 2010;90:787–96. [PubMed: 20195244]
187. Mueller C, deCarvalho AC, Mikkelsen T, et al. Glioblastoma Cell Enrichment Is Critical for Analysis of Phosphorylated Drug Targets and Proteomic-Genomic Correlations. Cancer Research. 2014;74:818–28. [PubMed: 24346432]
188. Carr SA, Abbatiello SE, Ackermann BL, et al. Targeted peptide measurements in biology and medicine: best practices for mass spectrometry-based assay development using a fit-for-purpose approach. Mol Cell Proteomics. 2014;13(3):907–17. [PubMed: 24443746]
189. McDonnell LA, Rompp A, Balluff B, et al. Discussion point: reporting guidelines for mass spectrometry imaging. Anal Bioanal Chem. 2015;407(8):2035–45. [PubMed: 25432304]
190. Yildiz-Aktas IZ, Dabbs DJ, Bhargava R. The effect of cold ischemic time on the immunohistochemical evaluation of estrogen receptor, progesterone receptor, and HER2 expression in invasive breast carcinoma. Mod Pathol. 2012;25(8):1098–105. [PubMed: 22460807]
191. McGrath TA, McInnes MDF, van Es N, et al. Overinterpretation of Research Findings: Evidence of “Spin” in Systematic Reviews of Diagnostic Accuracy Studies. Clin Chem. 2017;63(8):1353–62. [PubMed: 28606911]
192. Greco T, Zangrillo A, Biondi-Zoccai G, et al. Meta-analysis: pitfalls and hints. Heart Lung Vessel. 2013;5(4):219–25. [PubMed: 24364016]
193. Vivot A, Boutron I, Beraud-Chaulet G, et al. Evidence for Treatment-by-Biomarker interaction for FDA-approved Oncology Drugs with Required Pharmacogenomic Biomarker Testing. Sci Rep. 2017;7(1):6882. [PubMed: 28761069]
194. Nedelkov D, Kiernan UA, Niederkofler EE, et al. Population proteomics: the concept, attributes, and potential for cancer biomarker research. Mol Cell Proteomics. 2006;5(10):1811– 8. [PubMed: 16735302]
195. Dong C, Guo Y, Yang H, et al. iCAGES: integrated CAncer GEnome Score for comprehensively prioritizing driver genes in personal cancer genomes. Genome Med. 2016;8(1):135. [PubMed: 28007024]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

196. Karn T, Pusztai L, Rody A, et al. The Influence of Host Factors on the Prognosis of Breast Cancer: Stroma and Immune Cell Components as Cancer Biomarkers. Curr Cancer Drug Targets. 2015;15(8):652–64. [PubMed: 26452382]

197. Bond MR, Hanover JA. A little sugar goes a long way: the cell biology of O-GlcNAc. J Cell Biol. 2015;208(7):869–80. [PubMed: 25825515]

198. Matsuda A, Higashi M, Nakagawa T, et al. Assessment of tumor characteristics based on glycoform analysis of membrane-tethered MUC1. Lab Invest. 2017;97(9):1103–13. [PubMed: 28581490]

199. Potapenko IO, LQders T, Russnes HG, et al. Glycan-related gene expression signatures in breast cancer subtypes; relation to survival. Molecular Oncology. 2015;9:861–76. [PubMed: 25655580]

200. Ruhaak LR, Miyamoto S, Lebrilla CB. Developments in the identification of glycan biomarkers for the detection of cancer. Mol Cell Proteomics. 2013;12(4):846–55. [PubMed: 23365456]

201. Stowell SR, Ju T, Cummings RD. Protein glycosylation in cancer. Annu Rev Pathol. 2015;10:473–510. [PubMed: 25621663]

202. Yen T-Y, Macher BA, McDonald CA, et al. Glycoprotein profiles of human breast cells demonstrate a clear clustering of normal/benign versus malignant cell lines and basal versus luminal cell lines. Journal of Proteome Research. 2012;11:656–67. [PubMed: 22106898]

203. Haukaas TH, Euceda LR, Giskeodegard GF, et al. Metabolic clusters of breast cancer in relation to gene- and protein expression subtypes. Cancer & Metabolism. 2016;4:12 [PubMed: 27350877]

204. Willmann L, Erbes T, Halbach S, et al. Exometabolom analysis of breast car lines: Metabolic signature. Sci Rep. 2015;5:13374. [PubMed: 26293811]

205. Willmann L, Schlimpert M, Halbach S, et al. Metabolic profiling of breast cancer: Differences in central metabolism between subtypes of breast cancer cell lines. J Chromatogr B Analyt Technol Biomed Life Sci. 2015;1000:95–104.

206. Antuofermo E, Miller MA, Pirino S, et al. Spontaneous mammary intraepithelial lesions in dogs--a model of breast cancer. Cancer Epidemiol Biomarkers Prev. 2007;16(11):2247–56. [PubMed: 17982119]

207. Mohammed SI, Meloni GB, Pinna Parpaglia ML, et al. Mammography and ultrasound imaging of preinvasive and invasive canine spontaneous mammary cancer and their similarities to human breast cancer. Cancer Prev Res (Phila). 2011;4(11):1790–8. [PubMed: 21803985]

208. Shinoda H, Legare ME, Mason GL, et al. Significance of ERalpha, HER2, and CAV1 expression and molecular subtype classification to canine mammary gland tumor. J Vet Diagn Invest. 2014;26(3):390–403. [PubMed: 24760135]

209. Hines WC, Kuhn I, Thi K, et al. 184AA3: a xenograft model of ER+ breast adenocarcinoma. Breast Cancer Res Treat. 2016;155(1):37–52. [PubMed: 26661596]

210. Kumar R, Singh J. Biosimilar drugs: Current status. Int J Appl Basic Med Res. 2014;4(2):63–6. [PubMed: 25143877]

211. Arpino G, Bardou VJ, Clark GM, et al. Infiltrating lobular carcinoma of the breast. Breast Cancer Research. 2004;6:R149–R56. [PubMed: 15084238]

212. Iorfida M, Maiorano E, Orvieto E, et al. Invasive lobular breast cancer. Breast Cancer Research and Treatment. 2012;133:713–23. [PubMed: 22399188]

213. Jacobs VR, Augustin D, Wischnik A, et al. Concordance rates of biomarkers uPA and PAI-1 results in primary breast cancer vs. consecutive tumor board decision and therapy performed in clinical hospital routine: Results of a prospective multi-center study at certified breast centers. Breast. 2016;29:208–12. [PubMed: 27344290]

214. Prat A, Parker JS, Karginova 0, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast cancer research: BCR. 2010;12:R68. [PubMed: 20813035]

215. Vranic S, Tawfik 0, Palazzo J, et al. EGFR and HER-2/neu expression in invasive apocrine carcinoma of the breast. Mod Pathol. 2010;23(5):644–53. [PubMed: 20208479]

216. HercepTest, Code K5204. 8th ed. Denmark: Dako Denmark A/S and Dako Nor America, Inc.; 2015 p. 56.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Key issues

- Breast cancer is a heterogeneous disease, both between patients and within a patient. Breast cancer subtypes reflect different underlying biology, creating a variety of functional phenotypes.
- Intra-patient tumor heterogeneity limits the clinical accuracy of current breast cancer subtyping due to tissue sampling methods and the spatial and temporal variations in the breast cancer microenvironment.
- Breast cancer subtypes should be considered dynamic, working classifications due to on-going genomic and proteomic biomarker discoveries. Receptor cross-talk between and within receptor families affects downstream signal transduction.
- HER2 receptor signaling can be active despite low HER2 expression by IHC or FISH. A truncated HER2 extracellular domain remains hyperactive due to homodimer formation of the catalytically active internal domain. Phosphorylated HER2 and truncated HER2 profiling should become standard clinical practice for assessing treatment efficacy.
- Pre-analytical and analytical variables must be addressed by researchers, clinical trial sponsors, and funding agencies. Tissue fixation methods, cold ischemia time, whole tissue lysates versus selected cell population lysates, and adherence to standard operating procedures effect biomarker data and ultimately patient management and outcomes.
- Revealing the multiple biological interactions within the tumor microenvironment requires upfront tumor/stroma enrichment by macrodissection or laser capture microdissection. Tumor-stromal interactions and the presence of tumor subclones cannot be assessed in homogenized tissue lysates. It is not possible to determine which cell population contributed a particular DNA, RNA, or protein molecule unless the tumor, stroma, immune cells, and normal adjacent cells are analyzed as distinct specimens.
- Biomarker studies for investigating breast cancer subtypes should be designed to include minorities, various ages, and more importantly, strive to procure specimens representing a variety of breast cancer histomorpholog
- Quality control initiatives to increase the reproducibility of pre-clinical studies are gaining support from the proteomic community. Mass spectrometry consensus guidelines are available from Clinical Laboratory Standards Institute. Proteomic reporting standards have been developed by the Human Proteome Organization. Metrics for data submission are required by journals.
- Emerging biomarkers require verification and/or validation in large patient cohorts. Research sponsors and funding are needed to discover new

biomarkers/drug targets, but also to support on-going validation of promising biomarkers with translation to clinical use.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Step I: Histomorphology		
In situ carcinoma	Lobular	
	Ductal	Solid
		Cribriform
		Comedo
		Papillary
Invasive carcinoma	Lobular	
	Ductal	Medullary
		Mucinous (colloid)
		Tubular
		Apocrine
		Papillary
		Inflammatory

Step II: Protein Biomarker Menu	
Estrogen receptor	HER2
Progesterone receptor	EGFR
Androgen receptor	Ki-67
Cytokeratins 5/6/7/8/17/18/19	Claudin 3,4,7
CD44/CD24 (low or high)	ALDH1
E-Cadherin	P-Cadherin
Caveolin 1 and 2	Grb7
uPA/PAI-1	

Step III: Genomic/Proteomic Biomarker and Clinical Data Integration			
Intrinsic subtype (Gene expression)	Proteomic expression	Proliferation index	Treatment
Luminal A	ER+ and/or PR+ HER2 neg Cytokeratins /8/18 pos	Ki-67 low	Endocrine therapy
Luminal B	ER+ and/or PR+ Her2 neg or pos Cytokeratins 5/8/18 pos	Ki-67 high	Endocrine therapy Chemotherapy Anti-HER2 if HER2 positive
HER2+	ER neg/HER2+ Cytokeratin 5+/6 neg		Anti-HER2 Chemotherapy
Basal-like	ER neg/PR neg/HER2 neg EGFR+ Cytokeratins5/6/17pos P-cadherin+ Caveolin1 and 2+		Chemotherapy Radiation Angiogenesis inhibitors Combination therapy
Normal-like	ER+/neg HER2 neg Cytokeratins 5/6 pos	low	
Claudin low	Triple negative Low claudin 3, 4 and 7		Chemotherapy Radiation
Molecular apocrine	AR+ ER neg/PR neg HER2+ or EGFR+	low	Chemotherapy Radiation Anti-HER2

Figure 1. Current scheme for assigning breast cancer subtypes.

Histopathology is integrated with proteomic and genomic biomarkers to characterize breast tumors into clinically relevant subtypes. The first step is pathologist review of the tumor histology which is described as in situ or invasive, with the corresponding architectural pattern [2–4,11,211,212]. If the tumor is infiltrating ductal carcinoma, the stage of cellular differentiation is also described. Next selected protein biomarkers are semi-quantitatively scored based on the immunohistochemical staining pattern [4,18,26,83,213–215]. For ER +/PR+/HER2neg, lymph node negative tumors the Predictive Analysis of Microarray 50 gene signature (Prosigna, PAM50 assay) provides a predictive risk of recurrence score [40,41,44]. Integration of proteomic biomarker scores with gene expression signatures and clinical information aids therapy escalation and de-escalation decisions [33,37].

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

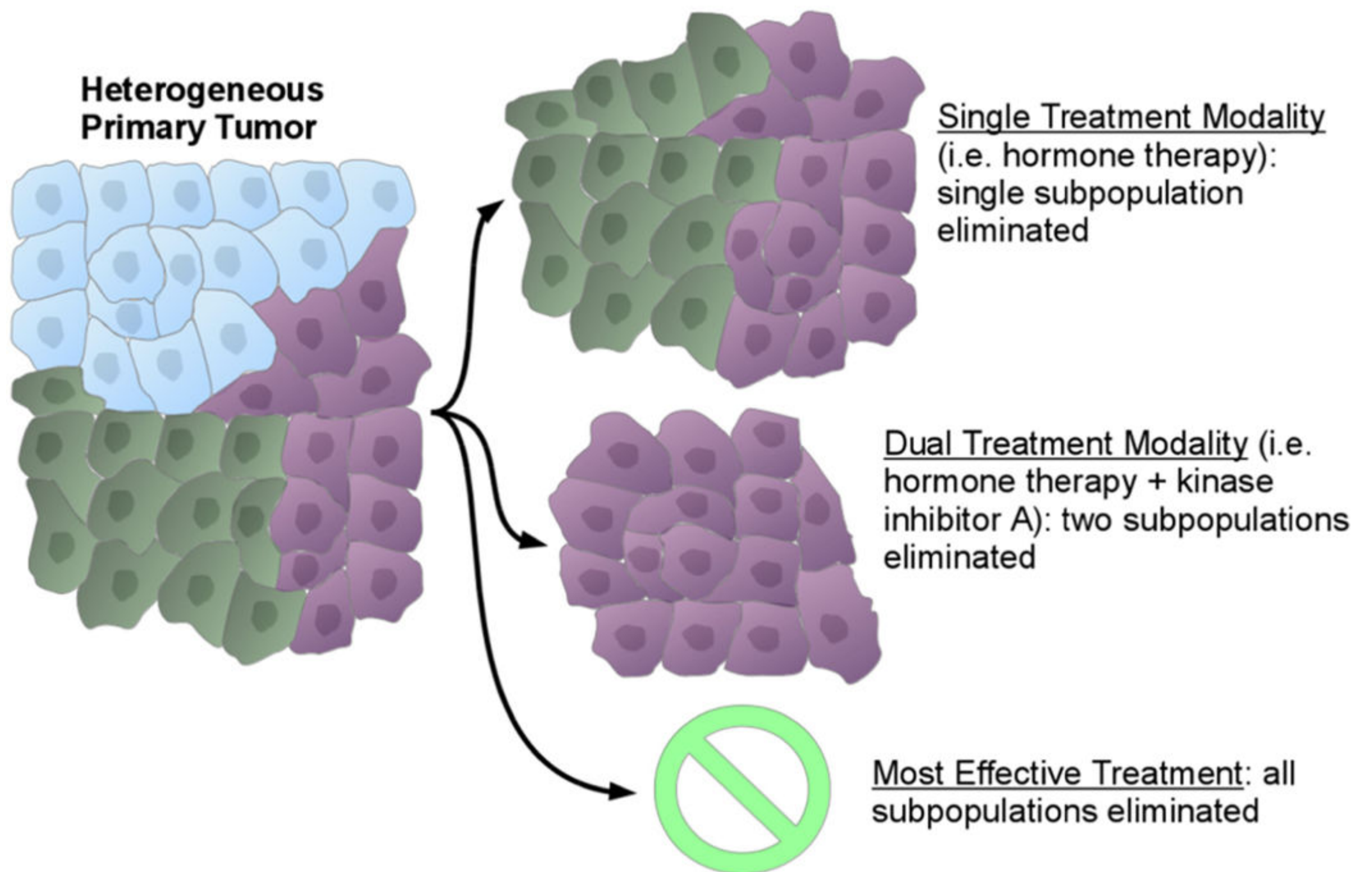


Figure 2. Intra-tumor heterogeneity in breast cancer

Breast tumors may contain cellular sub-clones harboring a variety of genomic and proteomic alterations. Clonal cooperation and emergence of sub-clones during treatment have a profound influence on the biological phenotype of the tumor and treatment resistance. Combination treatment to simultaneously, or sequentially, eliminate every clonal population could potentially provide greater ther efficacy.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

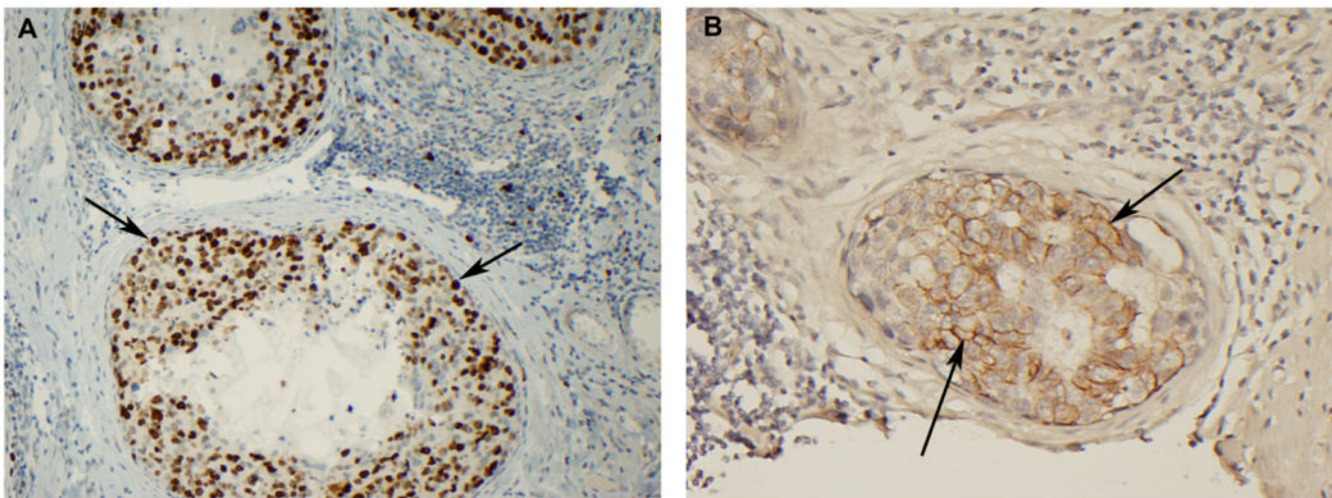
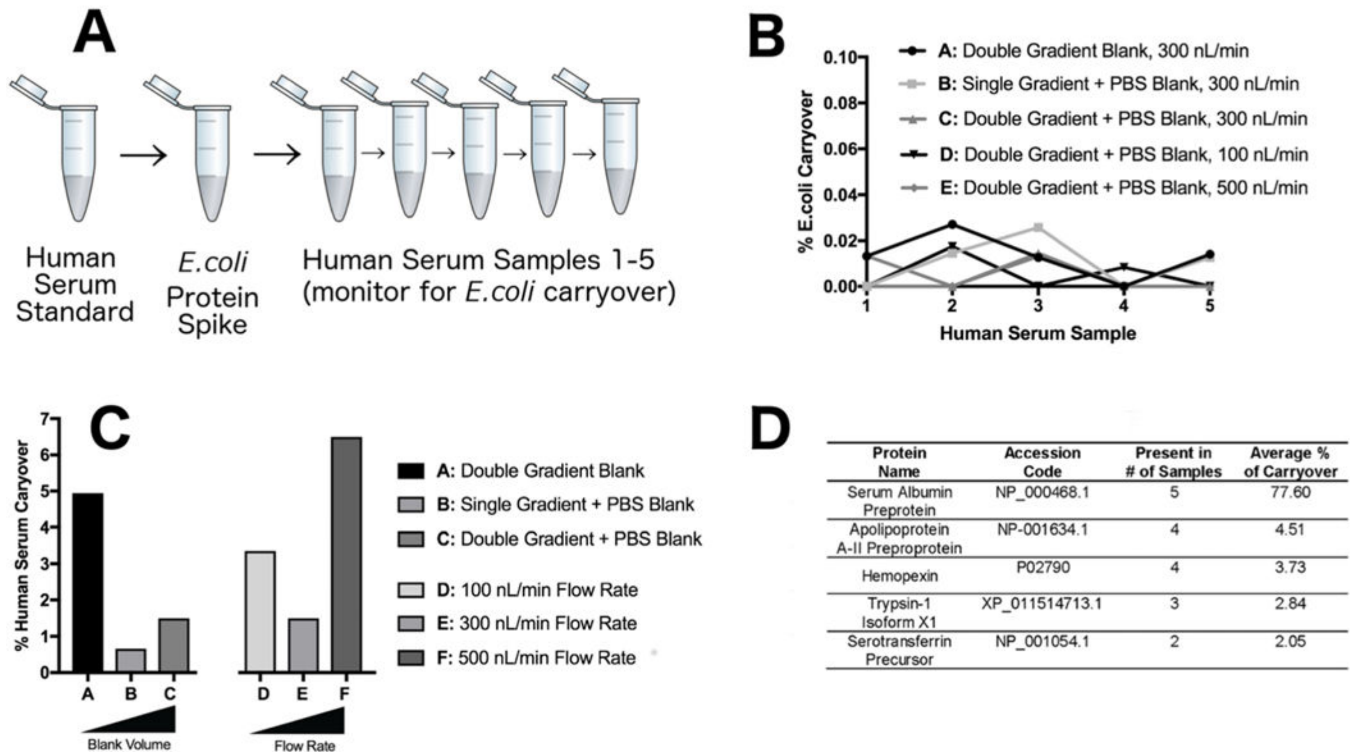


Figure 3. IHC provides protein biomarker cellular context and subcellular location

Example IHC staining patterns for (a) Ki-67 and (b) HER2 using Ki-67 antibody clone MIB-1 and Dako HercepTest™, respectively. Positive staining, which appears brown, occurs via deposition of diaminobenzidine at the site of the antigen-antibody interaction. The subcellular location of the staining is used for quality control; Ki-67 should be localized in the nucleus, whereas HER2 should be localized to the plasma membrane. IHC scoring can be qualitative (0, 1+, 2+, 3+) or semi-quantitative, based on the intensity of the stain and the proportion of positively stained cells. All IHC biomarker scoring is interpreted within the context of appropriate positive and negative controls. HER2 scoring using the HercepTest™ (Agilent/Dako) is considered negative if the membrane staining score is 0 or 1+. A score of 2+ is considered weakly positive. A score of 3+ is strongly positive. Note: HerceptTest™ staining is recommended for invasive cancers and any cytoplasmic staining should not be scored. The DCIS image depicted in panel B was for research use only and to illustrate IHC staining that is localized to the plasma membrane *[216]* (Ki-67 magnification 100x, HER2 magnification 200x).



Protein Name	Accession Code	Present in # of Samples	Average % of Carryover
Serum Albumin Preprotein	NP_000468.1	5	77.60
Apolipoprotein A-II Preproprotein	NP-001634.1	4	4.51
Hemopexin	P02790	4	3.73
Trypsin-1 Isoform X1	XP_011514713.1	3	2.84
Serotransferrin Precursor	NP_001054.1	2	2.05

Figure 4. Protein carry-overin liquid chromatography mass spectrometry potentially confounds results.

(a) Experimental procedure to assess protein carry-over in LC-MS/MS based on insertion of blank samples or adjustment of mobile phase flow rate. One human serum sample, used as a control, was analyzed by LC-MS/MS to ensure no *E. coli* contamination. A subsequent *E. coli* sample, consisting of recombinant *E. coli* DXP reductoisomerase (DXR) prepared in *E. coli* BL21 (DE3) RIL Codon Plus cells and purified to >90% purity, was analyzed for high and low abundance proteins. Carry-over of *E. coli* proteinswere examined in 5 subsequently analyzed human serum samples. All human serum samples were from the same serum aliquot. (b) Carry-over of *E. coli* proteins in subsequent human serum samples. No significant carryover of *E. coli* proteins were observed in any subsequent human serum samples using any of the five defined experimental protocols. (c) Carry-over of human serum proteins in *E. coli* samples. Carry-over was observed in all samples, with the greatest carry-over seen at the lowest volume in the blanking protocol and with the highest flow rates. (d) Human serum proteins carried over in *E. coli* samples. Albumin was found in all five *E. coli* samples, while apoliopprotein A-II preprotein, hemopexin, and trypsin-1 Isoform X1 were also found in the majority of the *E. coli* samples.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1.

Proteomic methodologies and emerging biomarkers for deciphering breast cancer subtypes/prognosis.

Protein biomarker	Breast cancer subtype	Specimen type/study set	Significance of protein biomarker	Reference
Immunohistochemistry (IHC)				
Programed cell death 1 (PD-1) protein Lymphocyte activation gene 3 (LAG-3)	TNBC	FFPE Early stage tumors	Immunotherapy benefit	[54] Bottai et al.
EGFR Cytokeratin 5/6	TNBC	FFPE surgical specimens	Increased levels associated with poor prognosis or recurrence	[86] Maeda et al.
Androgen Receptor (AR) DNAPKcs DNAPKcs Ser2056	TNBC	MDA-MB- 453 cell line	Recurrence post radiation therapy	[89] Speers et al.
CD44	All subtypes	Frozen breast tumors and FFPE	Stem cell phenotype associated with TNBC, EGFR+, and CK5/14+	[91] Honeth et al.
CD44 isoforms: splice variant CD44v8-v10	Basal-like	Frozen, stage II breast tumors, FFPE, and cell lines	Association with HER2+, ER+, and PR+	[92] Olsson et al.
Forkhead box transcription factor A1 (FOXA1) GATA binding protein 3 (GATA3)	Hormone+/Her2 negative	FFPE	Increased levels associated with good prognosis	[106] Hisamatsu et al. [107] Hisamatsu et al.
Heterochromatin protein 1 family (HP1β)	Ki67+	FFPE	Treatment response to PARP +carboplatin and prognosis	[113] Lee et al.
Reverse phase protein microarray/antibody array				
Mammalian target of rapamycin (mTOR) Extracellular regulated kinase (ERK)	TNBC	Frozen tumors, cell lines, animal models	Treatment options	[96] Montero et al.
GADD45 MAPK DNA damage proteins	TNBC	Frozen tumors	Biological features of inflammation, hormone responsive, or DNA damage response	[97] Masuda et al.
HER2 Tyr1248 EGFR Tyr1173 and Tyr1168 HER3	HER2+/HER2 negative	Frozen tumors and FFPE	Treatment options for HER2 negative patients	[99] Wulfkuhle et al.
Cleaved Caspase 9 53BP1 AMPKa GATA3 Rad51 P90RSK Thr359 and Ser363	Luminal A	Frozen tumors	Prognosis and functional differences within luminal A tumors	[102] Aure et al.
NME/NM23 Nucleoside Diphosphate Kinase 1 (NDKA) Ribosomal Protein S6 (RPS6) Caveolin-1	ER+	Frozen invasive tumors	Risk stratification	[103] Sonntag et al.
ERa Ser118 Fibronectin GATA3 Yes associated protein 1 (YAP-1) Elongation initiation factor 4B (elF4B) Histone 2 AX (H2AX)	Invasive lobular carcinoma	Frozen tumor and FFPE from TCGA and METABRIC cohorts	2 sub-groups of invasive lobular carcinoma: hormone+ and immune related	[105] Michaut et al.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Protein biomarker	Breast cancer subtype	Specimen type/study set	Significance of protein biomarker	Reference
Phosphatidylinositol- 3,4,5-trisphosphate dependent Rac exchange factor 1 (P-REX1)		TCGA breast cohort Cell lines	PI3Kinase inhibitor response	[111] Dillon et al.
Mass Spectrometry				
Immunoglobulins Apo-A1 Actin 15-hydroxyprostaglandin dehydrogenase (15- PGDH) Brain fatty acid binding protein (FABP7) hydroxymethylglutarate-CoA synthase 2	Apocrine breast cancer	Interstitial fluid Breast tissue	Identify functional phenotype of apocrine breast cancer	[118] Celis et al. [119] Celis et al. [120] Celis et al. [121] Gromov et al.
(HMGCS2)				
Mitochondrial isoleucyl-tRNA synthetase Hemopexin 14–3-3σ Lumican p53 SMAD NF-κ	ER+	Breast tumors and normal breast tissue	Identify functional protein pathways	[110] Cha et al.
Collagen Filamin-C Neprilysin Tenascin	Normal Luminal Triple negative Luminal/HER2	Cell lines	Determine functional phenotypes in breast cancer cell line models	[123] Bateman et al.
EGFR splice variants STAT3 splice variants MAPK1 splice variants	HER2+	Cell lines	Facilitate drug testing	[126] Menon et al.
Trypbottomhanyl-tRNA synthetase (TrpRs) Desmoplakin (DP) Thrombospondin-1 (TSP1)	TNBC	Frozen tumors, treatment naive	Identify drug targets	[127] Campone et al.
RAC2 RAB61 Biliverdin reductase A (BLVRA), Inorganic pyrophosphatase (IPYR)	TNBC	FFPE	Prognosis	[128] Gamez-Pozo et al.
Peroxidasin Lamininy2 Fibronectin EphA2 Transglutaminase 2 (TGM2)	Basal-like	MCF10A cell line	Prognosis and PI3Kinase inhibitor treatement	[129] Young et al.
AXL receptor tyrosine kinase (AXL)	TNBC	Cell lines	Prognosis	[132] Wu et al. [133] Wu et al.
Interferon regulatory factor 3 (IRF3) EGFR CD44 tRNA synthetase	HER2+ TNBC ER+	FFPE	Identify functional protein differences within subtypes	[142] Tyanova et al.
11 protein signature with catenin alpha-1	Frozen tumors, treatment naive	TNBC	Prognosis and chemotherapy decisions	[144] Liu et al.