



Published in final edited form as:

Nat Methods. 2014 February ; 11(2): 149–155. doi:10.1038/nmeth.2763.

Demonstrating the feasibility of large-scale development of standardized assays to quantify human proteins

Jacob J. Kennedy¹, Susan E. Abbatiello², Kyunggon Kim³, Ping Yan¹, Jeffrey R. Whiteaker¹, Chenwei Lin¹, Jun Seok Kim⁴, Yuzheng Zhang¹, Xianlong Wang¹, Richard G. Ivey¹, Lei Zhao¹, Hophil Min³, Youngju Lee⁴, Myeong-Hee Yu⁴, Eun Gyeong Yang⁴, Cheolju Lee⁴, Pei Wang¹, Henry Rodriguez⁵, Youngsoo Kim³, Steven A. Carr², and Amanda G. Paulovich¹

¹Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. N., Seattle, WA 98109

²Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142

³Department of Biomedical Engineering, Seoul National University College of Medicine, 28 Yongon-Dong, Seoul 110-799 Republic of Korea

⁴Center for Theragnosis, Korea Institute of Science and Technology, Seoul 136-791, Republic of Korea

⁵Office of Cancer Clinical Proteomics Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892

Abstract

The successful application of MRM in biological specimens raises the exciting possibility that assays can be configured to measure all human proteins, resulting in an assay resource that would promote advances in biomedical research. We report the results of a pilot study designed to test the feasibility of a large-scale, international effort in MRM assay generation. We have configured, validated across three laboratories, and made publicly available as a resource to the community 645 novel MRM assays representing 319 proteins expressed in human breast cancer. Assays were multiplexed in groups of >150 peptides and deployed to quantify endogenous analyte in a panel of breast cancer-related cell lines. Median assay precision was 5.4%, with high inter-laboratory correlation ($R^2 > 0.96$). Peptide measurements in breast cancer cell lines were able to discriminate amongst molecular subtypes and identify genome-driven changes in the cancer proteome. These results establish the feasibility of a scaled, international effort.

Users may view, print, copy, download and text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: http://www.nature.com/authors/editorial_policies/license.html#terms

Corresponding Author: Amanda Paulovich, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. N. E2-154, PO Box 19024, Seattle, WA 98109-1024, 206-667-1912, Fax 206-667-2277, apaulovi@fhcr.org.

AUTHOR CONTRIBUTIONS

J.J.K., S.E.A., K.K., J.R.W., P.W., Y.K., S.A.C. and A.G.P. conceived and designed the experiments. J.J.K., S.E.A., K.K., J.S.K., R.G.I., L.Z., Y.L. and H.M. performed the experiments. J.J.K., S.E.A., K.K., P.Y., C.L., Y.Z., and X.W. analyzed the data. M.H.Y., E.G.Y., C.L., H.R., Y.K., S.A.C. and A.G.P. contributed reagents/materials/analysis tools. J.J.K., J.R.W. and A.G.P. wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

INTRODUCTION

Rapid advances in technology have enabled extraordinarily deep proteomic coverage^{1, 2}. This deep coverage comes at the expense of throughput, due to extensive sample processing requirements. Thus, for interesting discovery proteomic leads to be actionable, investigators must be able to verify the results in larger clinical or biological studies³, requiring targeted methods of analysis enabling higher throughput. Unfortunately, conventional technologies (e.g. ELISA, IHC, Western blotting) are low in throughput, unable to avoid nonspecific interferences, not routinely multiplexed, not quantitative (aside from ELISA), and do not use internal standards (and thus are not readily standardized across laboratories)⁴. Thus proteomics currently lacks critical tools required for success.

Multiple Reaction Monitoring (MRM) Mass Spectrometry (MS) is positioning itself to dramatically improve quantitative proteomics. MRM-MS is an assay platform used for decades in clinical reference laboratories to quantify small molecules⁵ (e.g. metabolites in newborn screening) and is being rapidly taken-up by the biology and clinical research communities for quantifying peptides released via proteolysis of biospecimens^{6, 7}. MRM-MS was recently selected as the “method of the year” by *Nature Methods*⁸, given its potential to promote rapid advances in protein-based research, potentially replacing Western blotting and providing the critical missing link between discovery proteomics and downstream implementation of proteomic findings^{9, 10}.

MRM-MS is a targeted technique that is completely different from the mass spectrometry approaches widely used in discovery proteomics. MRM is performed on specialized instruments that enable targeting of specific analyte peptides of interest and provides exquisite specificity and sensitivity^{11–14}. Background interferences can be detected and avoided, and the use of spiked-in, stable isotope-labeled standards enables precise relative quantification of endogenous analytes in commonly used biospecimens¹⁵. The National Cancer Institute has invested heavily in the standardization and analytical validation of MRM-based quantification of peptides through its Clinical Proteomic Tumor Analysis Consortium (CPTAC)¹⁶, which has demonstrated robust analytical performance for MRM analyses across laboratories and instrument platforms¹⁷.

For the MRM-based assay technology to meet its potential to promote rapid advances in protein-based biomedical research, the ability to analyze MRM-based assays to quantify any human protein (with sufficient sensitivity and throughput) must be made readily available to the target user community (i.e. basic and translational scientists) in the form of validated assays that can be analyzed in individual laboratories or readily implemented in proteomic core facilities. Towards this end, global assay development projects have been proposed^{18–21}, and peptide spectral databases^{22, 23} (e.g. <http://www.srmatlas.org>) as well as open-source, vendor-neutral software tools^{24–28} (<https://panoramaweb.org>) are being rapidly developed to support such efforts.

In this study, we tested the feasibility and usefulness of a large-scale, international collaborative effort in MRM-MS assay generation targeting the human proteome, modeling what a global assay development effort might look like. Our approach was to develop a

panel of 645 MRM assays covering 319 proteins (~1.5% of the basic human proteome) differentially expressed amongst human breast cancer subtypes from start to finish (i.e. including reagent generation, assay development, analytical validation, assay deployment on biospecimens, and distribution of data and SOPs as a community resource) using state-of-the-art technology and multiplexing capabilities. The results demonstrate feasibility of an international, scaled project to develop MRM assays to all human proteins. We also demonstrate that MRM-based targeted proteomic measurements can recapitulate known biological subtypes of breast cancer, identify genome-driven changes in the cancer proteome, and provide complementary information to that encoded in mRNA or copy number profiles.

RESULTS

Empirical selection of targets

To model what an international global assay development effort might look like, 3 performance sites (Seattle, Boston, and Seoul) cooperated to develop 645 MRM assays representing 319 target proteins expressed in human breast cancers. Breast cancer was chosen as a model system because extensive genomic characterizations have been used to describe well-defined molecular subtypes^{29–31} and because a panel of highly characterized breast cancer cell lines^{32–34} was readily available for the study. Although we focused on breast cancer (and on cell lysates) to provide a framework for this pilot, the assays we developed are limited neither to application in cell lysates nor to breast cancer; they are generalizable.

To generate an empirical dataset for selection of target analytes for MRM assay development, unfractionated protein lysates derived from a panel of human breast cancers and breast cancer-derived cell lines (Supplementary Table 1) were analyzed by shotgun LC-MS/MS analysis. Over 64,000 unique peptides (representing 9,996 proteins) were identified at a peptide FDR < 0.005 in the combined cell line and tissue data. To enrich for targets that might vary in expression level amongst breast cancer subtypes and thus be of biological interest, potential MRM targets were rank-ordered by differences in their signal intensities amongst the breast cancer subtypes represented in the cell line panel, and identification was required in both the cell lysate and corresponding cancer tissue. Finally, a target list for MRM assay development was constructed from the filtered list of peptides that were detectable from neat cellular lysate by MRM on a triple quadrupole mass spectrometer. From this rank-ordered list, a set of 318 proteins (represented by 642 proteotypic peptides) were selected for assay development (Supplementary Table 2). These proteins were shown to be enriched for breast cancer-specific targets as 73 of these 318 proteins (23%) were also included in a list of 1000 proteins of potential functional importance in breast cancer³⁵. Although not observed by shotgun LC-MS/MS analysis, 3 peptides to ESR1 were also included, for a total of 319 proteins represented by 645 proteotypic peptides. The selected proteins map to a variety of cellular compartments and span a range of biological processes (Supplementary Fig. 1).

Development and characterization of multiplexed assays

For each analyte, synthetic light and heavy stable isotope-labeled peptides were prepared, and optimum MRM transitions and instrument parameters were determined as described in the online methods. The 645 individual peptide assays (Supplementary Table 2) were distributed randomly amongst 4 multiplex assay groups (each containing between 156 and 169 peptides). To avoid any bias for performance amongst assay groups, we ensured that each multiplex group contained an equivalent distribution of analyte intensities and retention times (Supplementary Fig. 2) in addition to LLOQs and CVs (see below and Supplementary Fig. 3). One of the multiplex assays was randomly chosen to be acquired at all 3 performance sites (the “inter-laboratory” assay), whereas each of the remaining 3 multiplex assays were acquired at only 1 performance site (the “site-specific” assays).

The analytical performance of the assays was evaluated at each site by generating response curves in a cell lysate matrix. For the 645 peptides in the study, 1,938 individual reverse response curves were generated [(483 site-specific assays + 486 inter-laboratory assays = 969 total) x 2 matrix dilutions]. All response curves were plotted (Supplementary Results), and assay figures of merit were calculated (Supplementary Table 3). The majority of assays featured a linear range >3 orders of magnitude. The median assay LLOQs for the inter-laboratory assay group were 0.40, 0.61 and 0.52 fmol/μg (at a cell lysate matrix protein concentration of 1.0 μg/μL), with median CVs of 3.5%, 5.0% and 4.4% for sites 1, 2, and 3, respectively. At this concentration, the site-specific assay groups had median assay LLOQs of 0.37, 0.65 and 0.40 fmol/μg, with median CVs of 3.5%, 5.4% and 4.4% for sites 1, 2, and 3, respectively.

An assay was deemed successful if it was precise (%CV ≤20% at the lowest concentration point in the linear range of the assay) and specific (detection of ≥1 transition of the light and ≥2 transitions of the heavy peptide and perfect co-elution of heavy and light peptides). Of the 645 assays attempted, 622 (96%) met these criteria and were considered to be successful. Furthermore, 599 (93%) had ≥2 transitions and 534 (83%) had all three transitions meeting these criteria.

Deploying MRM assays in breast cancer-related cell lines

Next, we determined the robustness of the assays when deployed in a common biological setting, characterizing human cell lines. Protein lysates from 30 human cell lines representing breast cancer (or normal breast epithelial cells; Supplementary Table 1) were prepared at a single site and distributed to all performance sites for MRM analysis (Fig. 1). Each lysate was digested in triplicate, so assay variability incorporates the complete processing variability.

A total of 174,420 individual assays were analyzed [(483 site-specific assays + 486 inter-laboratory assays = 969 total) x 30 cell lines x triplicate process replicates x 2 dilutions]. One quantifying transition was selected to calculate endogenous levels for each measurement above the analyte LLOQ (Supplementary Table 4 and Supplementary Fig. 4), from the transitions with no interfering signal (Supplementary Table 5). An assay was considered to be informative if the empirically determined concentration of the analyte was

above the assay LLOQ (i.e. indicates sufficient sensitivity); 93% (897 of 969) of the assays attempted met these criteria. Endogenous analyte was measured for all 319 proteins. At the individual peptide level, 609 out of 645 peptides (94%) were detected in at least one cell line and 547/645 (85%) were measured in at least half of the cell lines. The empirical concentrations of the endogenous peptides derived from the same protein showed very high correlation (median of 0.93) in the individual cell lines (Supplementary Fig. 5).

To evaluate precision, the CV across the complete process triplicates was calculated for all endogenous measurements above the LLOQ (Supplementary Table 6 and Supplementary Fig. 4). At the three sites, the median assay CVs for the inter-laboratory assay group were 5.0%, 7.3% and 5.1%, with 95% of the results having CVs less than 15%, 25% and 17% for sites 1, 2, and 3, respectively (Fig. 2a). The site-specific assay groups had median assay CVs of 4.7%, 6.3% and 4.7%, with 95% of the results having CVs less than 14%, 20% and 17% for sites 1, 2, and 3, respectively (Fig. 2b). The median CV for all measurements was 5.4%.

The empirically determined endogenous concentration of all analytes constituting the inter-laboratory 152-plex assay, which was analyzed at all 3 laboratories, was used to determine the correlation and agreement across the performance sites. Those measurements that were above the LLOQ at ≥ 2 sites (90% of measurements) were compared to determine the reproducibility of the measurements across sites. The correlation was excellent, with correlation coefficients ranging from 0.96 to 0.99 (Fig. 2c). There was also excellent agreement in the results amongst the sites, as demonstrated by the slopes from the linear regression of the correlation plots, which ranged from 0.95 to 1.07. This is also demonstrated by a histogram of the percent difference between site measurements (Fig. 2d). The mean percent difference was 0.9%, with 95% percent of the data within 22% difference and 75% percent of the data within 6.6% difference.

MRM results recapitulate known breast cancer subtypes

The empirically determined MRM-based measurements of 319 proteins were used for hierarchical clustering of the 30 cell lines. The cells lines formed 2 major clusters (Supplementary Fig. 6) which exactly match the clustering results previously observed for these cell lines using mRNA levels^{32–34} (into luminal and basal subtypes, largely correlated with estrogen receptor (ER) expression), demonstrating that MRM-based analyses can recapitulate the known molecular subtypes of breast cancer.

MRM results are clearly complementary to genomic data

We next asked whether the MRM data revealed any novel information about breast cancer that could not be determined using the genomic profiles of the cell lines. First, to identify proteins that are differentially expressed amongst the molecular subtypes of breast cancer, a Wilcoxon rank test was performed using the MRM dataset. When the false positive rate³⁶ (FDR) was controlled at 0.01, 4 proteins were found to be differentially expressed between HER2⁺ vs. HER2⁻ cell lines, 83 proteins were differentially expressed between ER⁺ vs. ER⁻ cell lines, and 118 proteins were differentially expressed between basal vs. luminal cell lines (Supplementary Table 7).

To determine if similar association patterns for this set of proteins can be observed based on their gene expression (mRNA) data (or if the proteomic data provided novel information), we made use of the genomic data of Neve *et al.* (2006)³³, which contains gene expression arrays for 28 of the 30 cell lines examined in our project. A total of 232 proteins quantified by MRM in this study also had corresponding gene expression measurements. A comparison between the proteins showing subtype-association at the mRNA and the proteomic level illustrates that candidate markers could be identified using the MRM data that were not detected based on RNA expression profiles (Supplementary Table 7 and Supplementary Fig. 7). Two, 7 and 11 proteins showed RNA expression levels significantly associated (P value ≤ 0.01) with HER2 (*ERBB2* gene product), ER and basal-luminal status, respectively, and did not show the same association patterns in their protein abundances, while 0, 44, and 56 proteins showed protein abundances significantly associated (using Wilcoxon rank test, FDR ≤ 0.01) with HER2, ER, and basal-luminal status, respectively, and did not show the same association patterns in their RNA expression signatures. These discrepancies demonstrate that protein profiling provides complementary information to genomic data (Fig. 3). To further demonstrate the complementary information that protein profiling provides, we focused on the 71 proteins whose protein abundances were significantly associated with HER2, ER, or basal-luminal status but whose RNA expression levels were not (i.e. the protein and mRNA data were discordant). Of these 71 proteins, 28 are believed to be functionally important in breast cancer, based on their inclusion in an independently curated set of 1000 human proteins of relevance to human breast cancer³⁵. This example demonstrates that information encoded at the proteomic level is different from that at the mRNA level, where no subtype-specific regulation of expression was observed.

Integrative analysis can identify potential disease genes

In prior studies of breast cancer, hundreds of genes were found to be associated with patient prognosis at the RNA expression level^{37–39}. Although these data suggest candidates, they are not sufficient to identify the primary drivers of clinical behavior of tumors, and many of these mRNA expression differences are not translated into differences at the protein level. Given the complementary information obtained from the mRNA and MRM proteomic results, we hypothesized that proteomic analyses may help identify clinically significant changes. The rationale for this hypothesis is twofold: i) changes observed in multiple independent datasets using orthogonal technologies (i.e. genomics and proteomics) are less likely to be false positives, and ii) having protein-level data should greatly augment the interpretation of genomic profiles by identifying changes that are ultimately expressed in the proteome, closer to the clinical phenotype.

We performed an integrative analysis and identified 31 proteins that show significant correlation (Bonferroni adjusted P value ≤ 0.0001) between the genomic³³ (i.e. DNA copy number and mRNA expression) and proteomic (MRM) data (Supplementary Table 8). Furthermore, amongst the 4 proteins associated with HER2 status, 2 have DNA copy number and gene expression information available, and both proteins (HER2 and GRB7) show significant concordance between genomic and proteomic signatures. Amongst the 118 proteins associated with basal-luminal status, 30 have corresponding genomic data, and only 10 (ABAT, ANXA1, PLOD3, CDKN2A, HER2, GALK1, CLTC, PRDX3, ALDOA and

DPYSL2) show significant concordance scores. Amongst the 83 proteins associated with ER status, 20 have corresponding genomic data, and only 5 (CLTC, PRDX3, ANXA1, ABAT, and PLOD3) show significant concordance scores (Fig. 4). Proteins whose expression is primarily regulated by gene expression showed agreement of measured protein levels to mRNA levels. 14 proteins were identified with protein levels significantly correlated (correlation > 0.7 and Bonferroni adjusted P value < 0.01) with its own gene expression. In other words, we can view this as a subset of proteins whose expressions are primarily regulated by RNA expression. Of these, 2, 7, and 3 proteins have abundances significantly associated with HER2, ER, and basal-luminal status, and 2, 4 and 2 respectively, showed the same association patterns in their RNA expression signatures. There was no proteins measured showing RNA expression levels significantly associated with HER2, ER and basal-luminal status which did not also have significantly associated protein abundances (Supplementary Fig. 7). Based on the above result, we conclude that the concordance of protein and mRNA levels for the subset of proteins whose expression is primarily regulated by gene expression is high, but not perfect.

Although the importance of amplification of the *ERBB2* locus (which also contains *GRB7*) in breast cancer is well established⁴⁰, the clinical relevance of the other 9 genes identified above is not known. As it has been shown that the genomic profiles of the cell lines in this study closely recapitulate those of primary breast cancers³³, we next tested whether these nine genes' expression levels were associated with outcome in 2 independent breast cancer datasets (referred to as van 't Veer *et al.*⁴¹ and Loi *et al.*⁴² datasets) that provide both survival outcome and genomic profiles for large sets of primary human breast cancers. When patients were stratified by either high or low expression levels for each of the 9 candidate genes, significant differences between Kaplan–Meier (KM) survival curves of the 2 patient groups were observed in both datasets for *CLTC*, *DPYSL2* and *ABAT* (Fig. 5). We next fit a multivariate cox proportional hazard model to further assess the association between gene expression and survival outcome, accounting for molecular subtype (PAM50)⁴³, age, tumor size, lymph node status, and other clinical covariates (Supplementary Table 9). Again, *CLTC* and *DPYSL2* were found to be significantly associated (P value_{*CLTC*}=0.029, 0.068 and P value_{*DPYSL2*}=0.067, 0.0048 in the 2 clinical datasets, respectively) with survival outcome. *ABAT* showed evidence of association with survival outcome in the Loi *et al.* dataset (P value_{*ABAT*}=0.012), but not in the van 't Veer *et al.* dataset. In summary, as a proof-of-principle, the above results illustrate the potential advantage of integrating quantitative proteomic data with genomic data to improve our understanding of which of the multitude of genomic alterations are most likely to be translated to the protein level, and thus most likely to contribute to clinical phenotypes.

DISCUSSION

Targeted proteomic assays covering the entire human proteome would alter the state of clinical and biomedical research, promoting rapid advances in protein-based biomedical research by allowing for better translation of basic findings into actionable results. To be useful, such assays must be easily implemented anywhere, with minimal adjustments, while maintaining a high level of performance.

All MRM assays developed in this study, including standard operating protocols (SOPs) for sample preparation and analyte-specific instrument parameters for data acquisition, have been made freely available as a resource for the community (see online methods). Each assay underwent rigorous analytical characterization and determination of analytical figures of merit, ensuring high quality standards for assay performance, as well as fit-for-purpose validation for the interrogation of human cell lines. The majority of academic centers now have proteomic core facilities with instrumentation to implement MRM-based assays, and all assays developed in this study are readily implementable in such facilities using the SOPs and Skyline files provided (see online methods, Supplementary Protocol). Furthermore, the cell lysate sample preparation is straightforward, does not require specialized equipment or expertise, and thus can be easily implemented in any modern biology laboratory. Although we have focused on breast cancer (and on cell lysates) to provide a framework for this study, the assays we develop are limited neither to application in cell lysates nor to breast cancer; they are generalizable.

The portability of MRM assays across laboratories and instrument platforms has been previously demonstrated in smaller studies aimed at a limited number of peptide analytes quantified by MRM-MS^{17, 44, 45}. In the present study, we substantially extend the work by demonstrating key requirements for a scaled effort, including a substantial increase in the number of assays configured, an unprecedented level of multiplexing analytically validated assays with internal standards (essential for a scaled effort), and successful international transfer of assays. Strict adherence and attention to SOPs enabled the assays to be highly reproducible, demonstrating international transferability of MRM assays, and thus the potential usefulness of a global MRM assay resource to the international community.

Of great interest and use to clinically-driven research, peptide measurements in individual breast cancer cell lines were able to discriminate between molecular subtypes, identify genome-driven changes in the cancer proteome, and provide information about cancer cell lines that was not encoded in genomic profiles. This demonstrates that panels of MRM assays can effectively contribute to biological characterization of molecular subtypes of cancer. Implementation of the assays to clinical samples (i.e. tumor tissue) will require overcoming at least two challenges: the limited yield of protein from a biopsy or surgical specimen and the microheterogeneity of cell types encountered in tumor tissue samples. The protein yields from core biopsies range from 80 to 400 micrograms, making it feasible for quantification of the analytes in this study; however these yields may be a challenge when lower abundance analytes are targeted and enrichment is required. Tissue microheterogeneity can be addressed by strict quality control of the input material (e.g. tumor cellularity), as has proven to be feasible in the application of gene expression profiles for breast cancer prognosis⁴⁶.

Together, the results of this study demonstrate the feasibility and usefulness of an international effort to develop, analytically validate, and distribute MRM-based assays to large suites of human proteins and demonstrates what could be done if various countries were willing to co-fund a scaled human protein quantification project^{18, 19}. One approach to realizing this potential is to develop analytically robust assays to groups of proteins based on biological pathways, cellular localization, or other logical groupings in an internationally-

coordinated fashion. Assay panels targeting whole pathways might be constructed for quantitative interrogation of biology.

This study targeted proteins accessible for MRM-based quantification using a very simple sample preparation protocol for generating cellular protein lysate, without biochemical fractionation or enrichment of the target analytes prior to MRM analysis. Assuming the success rate found in this study extends to the full range of human proteins whose endogenous levels are detectable by MRM from neat cellular lysates (i.e. without enrichment or fractionation), it is reasonable to estimate that several thousand human proteins might be quantified from cell line lysates by MRM alone (i.e. without enrichment). Note that this number is highly context-dependent. For example, although thousands of proteins may be quantifiable in cell lysates without enrichment, in a more challenging matrix (e.g. blood plasma) that number is in the hundreds. In all biospecimen types, the MRM assay success rate for quantifying endogenous levels of analyte is higher for more abundant proteins than for less abundant proteins. Thus, to achieve the vision of configuring MRM assays capable of detecting endogenous levels for the entire human proteome, enrichment strategies will be required for many proteins. For example, major classes of post-translational modifications (e.g. phosphorylation, etc.) are largely not accessible by MRM without enrichment. In the case of modifications, quantification using MRM may face limitations due to enrichment technologies (e.g. occasional difficulty enriching a specific modification or in generating an antibody to a specific modification) or peptide characteristics (modifications of interest must reside within proteotypic peptides with suitable size, chromatographic qualities, ionization properties, etc.) for analysis by mass spectrometry.

Analyte enrichment upstream of MRM can reduce sample complexity ($10^3 - 10^4$ enrichment), offering advantages of improved sensitivity, increased selectivity, and potential for increased throughput (via shorter LC-MRM-MS acquisition times). Enrichment can be achieved either biochemically⁴⁷⁻⁴⁹ (e.g. using chromatography) or through the use of analyte-specific antibodies for immuno-affinity enrichment⁵⁰⁻⁵⁴ (producing an immuno-MRM assay). Biochemical enrichments are generally costly and/or labor-intensive procedures that critically limit throughput and require specialized expertise (i.e. are not readily distributable to the general biology community). Immuno-affinity enrichment involves a single-step capture (immunoprecipitation) that is easily implemented in any modern research laboratory using existing expertise and infrastructure (and thus is highly distributable); the major limitations of this approach are a current lack of validated affinity reagents and the up-front cost and time required to generate renewable affinity reagents. Aside from costs, the production of high-affinity anti-peptide antibodies is associated with a 55% per peptide success and a >95% success rate on a protein level (when a multiplex immunization strategy is used⁵³).

ONLINE METHODS

Reagents—Urea (Ultra grade), iodoacetamide (IAM, Ultra grade), dithiothreitol (DTT), tris(2-carboxyethyl)phosphine (TCEP), EGTA, EDTA, and phosphatase inhibitor cocktails #1 and #2 were obtained from Sigma Aldrich. Trypsin Gold was purchased from Promega.

MS grade acetonitrile (MeCN) and water (Optima LCMS, #A955 and #W6, respectively) and ethanol (ETOH) were purchased from Thermo Fisher Scientific. Synthetic heavy and light peptides were purchased from Thermo Fisher Scientific, New England Peptide, and 21st Century Biochemicals.

Processing of breast cancer tissues—Surgically resected human breast cancers representing each breast cancer sub-type (ER⁺/HER2⁻, ER⁻/HER2⁺, ER⁺/HER2⁺, and ER⁻/HER2⁻) were obtained from the Fred Hutchinson Cancer Research Center's and the University of Washington's Breast Specimen Repository & Registry under IRB 5306. Marker determinations were made as part of the routine diagnostic workup of these tumors (e.g. IHC for ER and IHC ± FISH for HER2). Tissues were weighed while frozen and immediately transferred to ice-cold PBS to thaw. Once thawed, tissues were blotted to remove excess PBS, quickly chopped, and immediately transferred to ice-cold lysis buffer (25 mM Tris, 6 M Urea, 1 mM EDTA, 1 mM EGTA, 1 mM TCEP, 1% Sigma phosphatase inhibitor cocktail 1, 1% Sigma phosphatase inhibitor cocktail 2) at a mass ratio of 1:4 tissue:buffer. Tissues were homogenized in 1.5 or 0.5 mL microcentrifuge vials with disposable pestles, as was appropriate for tissue size. After disruption, samples were incubated in lysis buffer on a roller for 15 minutes at 4°C. Samples were then spun at 14k RPM for 10 minutes at 4°C, and the liquid phase was removed. To remove any residual debris, lysates were spun again at 14k RPM at 4°C and the liquid phase was again removed. Lysates were then aliquotted and stored at -80°C.

Lysates were pooled by mass based on a BCA (Pierce) measurement of lysate protein concentrations. Lysate pools for each tumor sub-type were processed in parallel. Briefly, lysates were reduced in 20 mM TRIS/20 mM DTT for 30 minutes at 37°C with shaking, followed by alkylation with 50 mM IAM in the dark at room temperature. Lysates were then diluted 1:10 with 100 mM TRIS, pH 8, before trypsin was added at a 1:50 trypsin:protein ratio by mass. After 2 hours, a second aliquot was added at 1:100 enzyme:substrate. Digestion was carried out overnight at 37°C with shaking. After 16 hours, the reaction was quenched with formic acid, final concentration 1% by volume. Digests were desalted using C-18 cartridges (Waters Cat. #WAT094225) with vacuum. The C-18 cartridges were washed with 3 volumes of 80% MeCN/0.1% formic acid (FA), then equilibrated with 4 washes of 0.1% FA. The digest was applied to the C-18 cartridge, then washed with 4 volumes 0.1% FA before being eluted drop by drop with 3 washes of 80% MeCN/0.1%FA. The eluate was then aliquotted by volume and digests were lyophilized, followed by storage at -80°C until use.

Growth of cell lines—All cell lines were obtained from American Type Culture Collection (ATCC). The cell lines were characterized and authenticated by ATCC using short tandem repeat (STR) DNA profiles. Individual cell lines were cultured as follows. Cells were thawed in 37°C water bath. The vials were wiped down with 70% ETOH, the cells were spun at 180 x g for 8 minutes at 4°C, the supernatant was discarded and the cells were resuspended in 10 mL medium (Supplementary Table 1). The cells were transferred to a 100mm x 20 mm plate (adherent cells) or T25 Flask (suspension cells) and incubated at 37°C, 5% CO₂. The adherent cells were split at 80–90% confluence (cell line-dependent) by

removing growth medium by aspiration, adding 2 mL 0.25% Trypsin-EDTA per 100 mm x 20 mm plate or 6 mL 0.25% Trypsin-EDTA per T175 Flask, and incubating cells at room temperature with occasional mixing until the cells lifted from the surface as seen under the microscope. Two to four volumes of medium containing serum were added to quench trypsin, and, if needed, a cell lifter was used to remove all cells from culture surface. The cells were pooled and additional medium was added as needed to obtain the optimal split ratio (cell line dependent - ranging from 1:2 to 1:6 as conditions warranted). The cells were dispensed into culture containers (100 mm x 20 mm plate for 10 mL final volume, T75 Flask for 12 to 15 mL and T175 Flask for 25 to 35 mL). The cells were grown, lifted from the surface as above, pooled in a 50 mL Falcon tube, and spun at 180 x g for 8 minutes at 4°C. The supernatant was removed and the cells were resuspended in freezing medium (90% growth medium + 10% DMSO) and aliquotted into 1.8 mL Cryovials. The vials were placed in a Nalgene freezing container (Cat. #5100-001), frozen overnight at -80°C, and transferred for storage in the vapor phase of a liquid nitrogen tank.

Preparation of protein lysates from cell lines—Individual and pooled cell lysates were prepared as needed. Cells were transferred to pre-cooled 50 mL tubes, and were spun at 180 x g for 8 minutes at 4°C and the supernatant was discarded. Cells from the same cell line were resuspended and pooled in 10 mL ice-cold DPBS, removing 50 µL for cell counting by Hemocytometer. Cells were washed twice by adding ice-cold DPBS to 50 mL, spinning cells 180 x g for 8 min at 4°C and discarding the supernatant. Lysis buffer was added to a final concentration of 0.5×10^8 cells / mL on ice, and the cell suspension was sonicated twice for 10 seconds (550 Sonic Dimembrator, Fisher Scientific; knob set to 5). The lysates were transferred by pipette tip to micro-centrifuge tubes and vortexed twice for 15 seconds, with a 10 minute rest on ice in between. The lysates were centrifuged at 20k x g for 10 minutes at 4°C and the supernatants transferred to 1.0 mL cryo-vials (Nunc Cat. #377267) and stored in liquid N₂. Lysate protein concentration was determined by BCA. For discovery profiling, lysates were pooled by protein mass based on their molecular subtype (ER⁺/HER2⁻, ER⁻/HER2⁺, ER⁺/HER2⁺, and ER⁻/HER2⁻). The lysates were reduced in 100 mM TRIS/20 mM TCEP for 30 minutes at 37°C with shaking, followed by alkylation with 50 mM iodoacetamide in the dark at room temperature. Lysates were then diluted 1:10 with 100 mM TRIS, pH8, before trypsin was added at a 1:50 trypsin:protein ratio by mass. After 2 hours, a second aliquot was added at 1:100 enzyme:substrate. Digestion was carried out overnight at 37°C with shaking. After 16 hours, the reaction was quenched with formic acid, final concentration 1% by volume. Digests were desalted using C-18 cartridges (Waters Cat. #WAT094225) with vacuum. The C-18 cartridges were washed with 3 volumes of 80% MeCN/0.1% FA, and then equilibrated with 4 washes of 0.1% FA. The digest was applied to the C-18 cartridge, and then washed with 4 volumes 0.1% FA before being eluted drop by drop with 3 washes of 80% MeCN/0.1% FA. The eluate was then aliquotted by volume, and digests were lyophilized and stored at -80°C until use.

Empirical identification of targets for MRM assay development—Breast cancer samples (tissue and cell line lysates) representing 4 molecular subtypes based on ER and HER2 status (Supplementary Table 1) were analyzed by shotgun LC-MS/MS. Four pooled tissue samples (one from each of the 4 subtypes) and ten pooled cell line samples (2 pools

each from the 4 subtypes and 2 non-tumor “normal”) were analyzed by 1D- and 2D-LC-MS/MS. Results from the 2D-LC-MS/MS analysis were used to measure the relative abundances of the proteins in the various subtypes and to distinguish those that were differentially expressed. From this list of candidates, proteins with two proteotypic peptides that were observed by 1D-LC-MS/MS (considered to be at a sufficient intensity to be seen reliably in MRM) and were observed in the corresponding pooled tissue sample were prioritized for assay development.

Two analytical systems were used for analysis of the tissue and cell line lysates. For the first system, the LC system consisted of a nanoAcquity HPLC (Waters) with high pH mobile phases of 20 mM ammonium formate at pH 10 in water (A1) and 100% acetonitrile (B1) and low pH mobile phases of 0.1% formic acid in water (A2) and 0.1% formic acid in acetonitrile (B2). For the 2D-LC-MS/MS analyses, 10 µg of protein digest was injected at high pH onto a 300 µm x 50 mm XBridge C18, 130Å, 5 µm column (Waters Cat. #186003682). A step gradient was used to elute the sample off of the high pH reverse-phase column in 6 distinct fractions using 11.1, 14.5, 17.4, 20.8, 45.0 and 65.0% B. Fractions were eluted from the high pH column into 20 µL/min of low pH (A2) and onto a 180 µm x 20 mm C18, 100Å, 5 µm, column (Waters Cat. #186006527) by the following method: hold 3% B for 0.5 min, gradient from 3 to step% B for 0.5 min, hold step% B for 4 min gradient from step to 3% B for 0.5 min, re-equilibrate at 3% B for 15 min. The flow rate was 2 µL/min. 2D-LC-MS/MS samples eluted from the trap column and separated by a 100 µm x 100 mm C18, 130Å, 1.7 µm, column (Waters Cat. #186003546) by the following method: gradient from 3 to 40% B for 120 min, gradient from 40 to 90% B for 2 min, hold 90% B for 10 min, re-equilibrate at 3% B for 20 min. The flow rate was 1000 nL/min. 1D-LC-MS/MS analyses were carried out as described without a trap column, using direct injection of 2 µg protein digest. The HPLC was coupled to an LTQ-Orbitrap Velos hybrid mass spectrometer using an Advance CaptiveSpray source (Michrom Bioresources) operated in positive ion mode. A spray voltage of 1700 V was applied to the nanospray tip. MS/MS analysis consisted of 1 full scan MS from 300–2000 m/z at resolution 30000 followed by 15 data dependent MS/MS scans. Dynamic exclusion parameters included repeat count 1, exclusion list size 500, and exclusion duration 15 seconds. For 1D-LC-MS/MS analyses, 5 replicate injections were performed.

The second analytical system consisted of an Eksigent nanoLC-Ultra 2Dplus (Eksigent Technologies) used for a direct injection of 1 µg of sample onto a 75 µm x 15 cm column made from a PicoFrit column (New Objective) packed with Magic C18 AQ 5 µm, 100 Å resin (Michrom Bioresources). The samples eluted from the column by the following method: gradient from 3 to 40% B for 120 min, gradient from 40 to 90% B for 2 min, hold 90% B for 10 min, re-equilibrate at 3% B for 20 min. The flow rate was 200 nL/min. The HPLC was coupled to an LTQ-Orbitrap hybrid mass spectrometer using an Advance CaptiveSpray source (Michrom Bioresources) operated in positive ion mode. A spray voltage of 1700 V was applied to the nanospray tip. MS/MS analysis consisted of 1 full scan MS from 300–2000 m/z at resolution 60000 followed by 10 data dependent MS/MS scans. Dynamic exclusion parameters included repeat count 1, exclusion list size 500, and exclusion duration 15 seconds. 5 replicate injections were performed.

Data were searched against version 3.69 of the Human International Protein Index (IPI) sequence database with decoy sequences using Spectrum Mill, OMSSA, and the X!Tandem database search engine with a previously described score plugin⁵⁵. All searches were performed with tryptic enzyme constraint set for up to two missed cleavages, oxidized methionine set as a variable modification and carbamidomethylated cysteine set as a static modification. For X!Tandem, peptide MH+ mass tolerances were set at ± 2.0 Da with post search filtering of precursor mass to 50 ppm and fragment MH+ mass tolerances were set at ± 0.5 Da. For OMSSA, peptide MH+ mass tolerances were set at ± 2.0 Da and fragment MH+ mass tolerances were set at ± 0.5 Da. For Spectrum Mill, peptide MH+ mass tolerances were set at 20 ppm and fragment MH+ mass tolerances were set at ± 0.7 Da. Identifications from three search engines were made with an FDR < 0.005 based on a decoy database search. The LC-MS/MS proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository⁵⁶ with the dataset identifier PXD000246. The intensity of each peptide in each subtype (including normal) was measured by two methods (MS1 and MS2). MS1 measures the precursor ion intensity for each acquisition (reported by SpectrumMill) and MS2 measures the sum of intensities of matched fragment ions for each acquisition⁵⁷. The intensity values were normalized by the median values of each individual acquisition and then log base 2 transformed. For each peptide, the reported intensity value for a given subtype was the maximum value between the two pools and between the MS1 and MS2 measurements.

Results from the 2D-LC-MS/MS analysis (where the method is more sensitive to detection of the analytes) were used to measure the relative abundances of the proteins in the various subtypes and to distinguish those that were differentially expressed. Absolute differential expression of a peptide was measured by the ratio of the highest intensity of the samples from a given subtype to the highest intensity of the samples from a comparator subtype (e.g. ER⁺ vs. ER⁻). Relative differential expression was measured by a differential score (DScore), which took the ratio and normalized it to the sum of the standard deviations of all intensities associated with the two comparator subtypes. Proteins were ranked according to the weighted average of the ratio and DScore of the individual peptides from that protein. From this list of candidates, proteins with two proteotypic peptides that were observed by 1D-LC-MS/MS (considered to be at a sufficient intensity to be seen reliably in MRM) and were observed in the corresponding pooled tissue sample (maximizing the value of the assay as a resource to the community as well as providing a suite of validated assays to be used to develop approaches for dealing with the difficulties of tissue-based analysis) were prioritized for assay development.

Qualification of potential targets on a triple quadrupole platform—The detectability of prioritized targets was verified on a triple quadrupole (QqQ) system. Proteotypic peptides and suitable transitions could be optimized directly from the discovery libraries, but to ensure that we could successfully develop an MRM assay, we confirmed detection using a triple quadrupole mass spectrometer prior to final peptide selection. Each candidate peptide was qualified by detection of endogenous analyte in a pool of all cell line lysates used in the profiling experiments using MRM (4000 or 5500 QTRAP (Sciex)).

Peptides were filtered to include only those that were proteotypic, had a hydrophobicity calculated by SSRCalc⁵⁸ between 8 – 57, did not contain methionine, and did not contain N-terminal cysteine or glutamine. For peptides meeting these criteria, the 10 most intense transitions were selected from a consensus spectral library of all identifications and exported into a scheduled MRM method based on retention time prediction using SSRCalc 3.0 (100A). The exported transition list was separated into multiple instrument methods so that there were no more than 1000 transitions per MRM analysis.

Verification of the MRM detectability of the endogenous peptide signals was performed using an Eksigent nanoLC-Ultra 2Dplus (Eksigent Technologies, Dublin, CA) coupled to a 5500 QTRAP mass spectrometer (ABSciex). Mobile phases consisted of 0.1% formic acid in water (A) and 90% acetonitrile with 0.1% formic acid (B). The pool of cell line lysates (1 µg) was loaded onto a 0.2 × 5 mm Chromolith CapRod RP-18e column (EMD Chemicals) for 3 minutes at 10 µL/min with 5% mobile phase B. The peptides were then separated by a 0.1 × 150 mm Chromolith CapRod RP-18e column (EMD Chemicals) by the following gradient method: hold 5% B for 9 min, gradient from 5 to 40% B for 100 min, gradient from 40 to 90% B for 1 min, hold 90% B for 5 min, re-equilibrate at 5% B for 14 min. The flow rate was 1000 nL/min. The trap column was back-flushed during the last 5 minutes of an acquisition using 5% B at 10 µL/min. The source employed was an Advance CaptiveSpray source (Michrom Bioresources). The MS was used in positive ion mode with typical source parameters consisting of a 1700 V ion spray voltage, curtain gas setting of 10, collision gas setting of medium, ion source gas 1 setting of 0, and an interface heater temperature of 110 °C. Collision energy (CE) and declustering potential (DP) parameters were set by Skyline, EP was set to 10, CXP was set to 10 and Q1 and Q3 set to unit/unit resolution (0.7 Da). The scheduled MRM option was used for all data acquisition with a target scan time of 2.25 seconds, and a 20 minute MRM detection window. The results of the MRM analysis were imported into Skyline. A positive identification of the peptide from the MRM was based on the relative intensities of the transitions compared to the relative intensities of the fragments in the shotgun MS/MS spectrum, yielding a Skyline dot product score greater than 0.95.

LC-MRM-MS for quantitative analysis

All quantitative MRM studies, including response curves and profiling cell lysates, were conducted using the following approach. The LC-MRM-MS analysis was performed by two-column switching on a nanoLC-MRM-MS system. The LC system consisted of an Eksigent nanoLC-Ultra 2Dplus connected to a cHiPLC-Nanoflex system (Eksigent Technologies). Mobile phases consisted of 0.1% formic acid in water (A) and 90% acetonitrile with 0.1% formic acid (B). 1 µL of sample was loaded onto and separated by a 75 µm × 15 cm ChromXP C18-CL 3 µm 120 Å column (Eksigent Technologies). Because the method employed two-column switching, the sample was loaded onto the column being regenerated (i.e. the column where flow was not directed to the MS) by the following method: gradient from 50 to 90% B for 2 min, hold 90% B for 5 min, gradient from 90 to 3% B for 1 min, re-equilibrate at 3% B for 42 min. The sample was injected onto the loading column at 25 minutes. The column was switched (i.e. flow was directed to the MS) and the sample was separated by the following gradient method: gradient from 3 to 7% B for 3 min, gradient from 7 to 25% B for 27 min, gradient from 25 to 40% B for 7 min, gradient from 40

to 60% B for 1 min, hold 60% B for 2 min, gradient from 60 to 3% B for 1 min, re-equilibrate at 3% B for 9 min. The flow rate was 300 nL/min and the column temperature was 40 °C. The LC system was coupled to a 5500 QTRAP mass spectrometer (ABSciex) by an Advance CaptiveSpray source (Michrom Bioresources). The MS was used in positive ion mode with parameters consisting of a 1200 V ion spray voltage, curtain gas setting of 10, nebulizer gas setting of 0, and an interface heater temperature of 110 °C. CE was set by Skyline, DP was set to 100, EP was set to 10, CXP was set to 10 and Q1 and Q3 set to unit/unit resolution (0.7 Da). Throughout the method, the actual cycle time remained at or below 2 seconds allowing for measurement of at least 10 points across the peaks.

Quantitative MRM assay development and characterization by response

curves—Heavy stable isotope-labeled standards (SIS) and matched light versions were synthesized and purified to >95% purity by HPLC. Heavy peptides incorporated a fully atom labeled ¹³C and ¹⁵N isotope at the C-terminal lysine (K) or arginine (R) position of each (tryptic) peptide, resulting in a mass shift of +8 or +10 Da, respectively. Peptides were quantified by amino acid analysis and aliquots were stored in 30% acetonitrile/0.1% formic acid at -80°C until use. Each synthetic peptide was analyzed by MS/MS on an LTQ mass spectrometer using infusion with an Advion TriVersa interface. The spectral library from these analyses was used to select transitions for optimization. Optimal collision energy for a hybrid triple quadrupole/linear ion trap mass spectrometer (5500 QTRAP) for each of the peptides was determined by injecting 50 fmol standard peptide solutions into a flow of 30% acetonitrile, 0.1% formic acid at a flow rate of 1 µL/min. Optimal values were determined by ramping the potentials and evaluating the results through DiscoveryQuant (Sciex). The top 3 transitions were selected for method development based on the presence of abundant y-ions at *m/z* greater than the precursor. In the absence of high *m/z* y-ions, the most abundant fragment ions were selected.

The analytical performance of the biomarker candidate assays was characterized by performing a response curve in a background matrix consisting of an equal mix (by protein mass) of all of the cell lines. Because the concentration of a number of analytes was expected to exceed the practical linear range of the assay, response curves were generated in the pooled lysate and a 1:10 dilution of the pooled lysate (1.0 µg/µL and 0.1 µg/µL background matrix). Digestion was performed in an automated fashion as described above using an Eppendorf epMotion 5075 automated pipetting system. A reverse curve was prepared in which the SIS peptide concentration was varied over 8 concentration points using 3-fold serial dilutions over the range 200–0.091 nM. Light peptide was also spiked into the cell lysate pool at 5nM to ensure that the heavy to light peak areas were within two orders of magnitude. Blanks and double blanks were prepared and analyzed in addition to the concentration points of the curve. Three process replicates were prepared and analyzed at the eight concentration points (along with blank and double-blank samples). Retention times on the LC platform were empirically determined using a mixture of the standard stable isotope-labeled synthetic peptides in a non-scheduled fashion. Once the retention times were known, scheduled MRMs were set up using a 150 second MRM detection window and a target cycle time of 1.5 seconds.

Peak integration was performed by Skyline, and the integrations were manually checked. Peak specificity between the light (or endogenous) and heavy (or SIS) MRM signal was defined as the detection of ≥ 1 transition from the endogenous peptide exactly co-eluting with ≥ 2 transitions from the stable isotope-labeled peptide, with the relative intensity of the light transition(s) deviating no more than 20% compared to the relative intensity of the corresponding heavy transitions. Linear regression was used to fit the serial dilution data points for each curve. Regression was performed using a $1/y$ weighting on all points having a correlation coefficient of >0.98 . Limits of detection (LOD) and lower limits of quantification (LLOQ) for each transition were obtained by using the average of the three blank measurements plus three times the standard deviation of the noise (for LOD) and ten times the standard deviation of the noise (for LLOQ), respectively. The upper limit of quantification (ULOQ) was determined by the highest concentration point of the response curve that was maintained in the linear range of the response. An assay was deemed successful if it was precise (%CV $\leq 20\%$ at the lowest concentration point in the working range of the assay) and specific, as defined above (≥ 1 light transition co-eluting with ≥ 2 heavy transitions). Assays that met these specificity/precision criteria were considered successful assays.

Deployment of analytically validated MRM assays on a panel of cell lines—

Digestion of the individual cell lines at two concentrations (1.0 $\mu\text{g}/\mu\text{L}$ and 0.1 $\mu\text{g}/\mu\text{L}$, by protein mass) was performed in an automated fashion as described above. After digestion, a mix of SIS peptides was spiked into the individual cell lysates at the concentrations shown in Supplementary Table 4. The “heavy” peptides were spiked at one of three possible concentrations, depending on the LLOQ of the peptide MRM assay as well as the expected endogenous signal. Spike levels were high enough above the LLOQ so as not to contribute unnecessarily to the assay CV and were designed to be close to expected endogenous levels so that the peak area ratio was not outside of the range of 100:1 and 1:100. Three complete process replicates were prepared and analyzed for the 30 individual cell lines. As with the response curves, retention times on the LC platform were empirically determined using a mixture of the standard stable isotope-labeled synthetic peptides in a non-scheduled fashion. Once the retention times were known, scheduled MRMs were set up using a 150 second MRM detection window and a target cycle time of 1.5 seconds. The quantitative LC-MRM-MS analysis was performed using the method described above. MRM data were processed using Skyline. All data were manually inspected to ensure correct peak detection, absence of interferences, and accurate integration. Reported peak areas are the sum of the peak area and background area reported by Skyline. Peak specificity between the light (or endogenous) and heavy (or SIS) MRM signal was defined as the detection of ≥ 1 transition from the endogenous peptide exactly co-eluting with ≥ 2 transitions from the stable isotope-labeled peptide, with the relative intensity of the light transition(s) deviating no more than 20% compared to the relative intensity of the corresponding heavy transitions. Peptide concentrations are reported from the results of one quantifying transition, defined as the transition with the lowest LLOQ with no interferences for a given peptide. When a working assay had two or more transitions that met these criteria, the quantifying transition was defined as the transition with the lowest CV at the LLOQ. Working assays were described as “informative” if it were sufficiently sensitive to precisely quantify the endogenous analyte in

at least one cell line. Endogenous levels were calculated by integrating peaks from the heavy and light signals and measuring the peak area ratio against the isotope-labeled analog, which was spiked at a known concentration in the lysates. Peak integration was performed by Skyline, and the integrations were manually checked. For the individual data sets, 8% of the integrations were changed after manual investigation. Precision was determined by measuring the coefficient of variation (CV, standard deviation divided by the mean) and expressed as a percent. Integration results were exported to the program R for linear regression and statistical analysis.

Cell lysate samples were analyzed in complete process triplicate (including digestion) at two cell lysate protein concentrations (1.0 $\mu\text{g}/\mu\text{L}$ and 0.1 $\mu\text{g}/\mu\text{L}$). All reported results are from the 1.0 $\mu\text{g}/\mu\text{L}$ data set unless the measurements were above the upper level of quantification (ULOQ) or missing. In those cases, the results from the 0.1 $\mu\text{g}/\mu\text{L}$ data set were normalized to the lysate protein amount and reported in place of the 1.0 $\mu\text{g}/\mu\text{L}$ results. Endogenous levels reported at 0.1 $\mu\text{g}/\mu\text{L}$, which make up 4.2% of the reported measurements, are marked with an asterisk in Supplementary Tables 4 and 5.

Results from the peptide quantification were used to report protein concentrations, assuming 100% recovery of the peptides. From the results of the two peptides associated with a given protein, the median values of the triplicate measurements were considered. Concentrations of 0.5*LLOQ or 2*ULOQ were imputed for missing values below LLOQ and above ULOQ, respectively. For a given protein, the peptide with the least number of missing values among the 30 cell lines was used to calculate protein concentration. If multiple peptides had the same number of missing values, the most intense peptide was used. Clustering was performed with these protein concentrations using the R function heatmap with complete linkage. All MRM data (response curves and cell line measurements) are available at <https://proteomics.fhcrc.org/CPAS/project/Published%20Experiments/International%20MRM%20Assays/begin.view?>

Cellular compartment and biological process characterization of proteins

included in the study—A total of 5,416 Gene Ontology (GO) annotations were obtained for the 319 proteins in the study from the Gene2GO dataset derived from NCBI Entrez Gene resources (<ftp://ftp.ncbi.nih.gov/gene/DATA/>). Broad classifications of the gene products based on cellular compartment and biological process were determined by cross-mapping to the Generic GO Slim dataset (<http://www.geneontology.org/GO.slims.shtml>). Cellular Compartment GO Slim annotations were available for 288 of the proteins in the study, which mapped to 26 compartment terms. Cellular Biological process GO Slim annotations were available for 188 of the proteins in the study, which mapped to 36 biological process terms. Heatmap plots were generated using the heatmap.2 R package.

Proteo-genomic integration—DNA copy number data and RNA expression data were obtained from Neve *et al.* (2006)³³. Gene expression arrays were available for 232 of the 319 proteins and 28 of the 30 cell lines targeted in this study. Differential expression amongst the molecular subtypes of breast cancer was determined by calculating the p-value and FDR (false discovery rate). The p-value reported is from the nonparametric Wilcoxon

rank sum test based on Z statistic. The FDR was calculated using R package qvalue⁵⁹, to control the family wise error due to multiple hypothesis testing.

Both DNA copy number and mRNA expression data were available for 90 of the 319 proteins and 27 of the 30 cell lines targeted in this study. Global normalization was performed for both raw CGH array and expression array data, and copy number amplifications and deletions were inferred using R package cghFlasso⁶⁰. Then for each gene, the average of the pairwise Spearman's rank correlations among its DNA copy number, RNA expression, and protein expression across the 27 cell lines was calculated. To assess the significance of these average correlation scores, permutation testing was performed (10k permutations), in which the null distributions were estimated through permuting the sample orders in the RNA expression data set and protein expression data set.

Survival analysis—The van 't Veer *et. al.* data⁴¹ were downloaded from the Stanford University public repository (<http://microarraypubs.stanford.edu/woundNKI/explore.html>). The dataset contains gene expression arrays, molecular subtype, and patient clinical outcome information for 244 breast cancer tumors. We normalized the raw data for each array to have median 0 and MAD (median absolute deviance) 1. To investigate the prognostic property of the candidate genes, we first partitioned the 244 samples into two groups by using the median gene expression level of each gene, and performed a Logrank test to compare the KM curves of the two patient groups. For the two genes giving significant p-values in the Logrank test, we further fit a multivariate cox proportional hazard model (R function *coxph*) to assess the association between gene expression and the survival outcome, accounting for the molecular subtype, age, tumor size, lymph node status, and whether the patient has received chemotherapy.

The Loi *et al.* dataset⁴², including gene expression data and clinical information for 414 breast cancer tumors, was obtained from (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE6532>). After performing the global normalization, R Package “genefu” (version 1.8.0) was used to determine the molecular subtypes of the breast cancer patients in the Loi *et. al.* data. As described above for the van 't Veer *et. al.* dataset, a logrank test and cox proportional hazard model were performed.

Public access to mass spectrometry data—The LC-MS/MS proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository⁵⁶ with the dataset identifier PXD000246 (username: review35086; password: ncZMftbc). All MRM data (response curves and cell line measurements) are available at <https://proteomics.fhrc.org/CPAS/project/Published%20Experiments/International%20MRM%20Assays/begin.view?>.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to the Fred Hutchinson Cancer Research Center – University of Washington Breast Specimen Repository and Registry (BSRR) for specimens used in this study. The BSRR is generously supported by the Breast Cancer Relief Foundation, The Foster Foundation, and Hutchinson Center funds. All BSRR specimens and data have been obtained in accordance with all applicable human subjects laws and regulations, including any requiring informed consent. The authors thank S. Skates, D. Ransohoff, and L. Anderson for helpful discussions. Research reported in this publication was supported in part by the Office of the Director, US National Institutes of Health (OD) and the US National Cancer Institute (NCI) with funds from the American Recovery and Reinvestment Act of 2009 under Grant RC2CA148286. The research was also partially supported by US National Institutes of Health Grant U24CA160034 from NCI Clinical Proteomics Tumor Analysis Consortium Initiative, and US National Institutes of Health Grant P50CA138293 from the NCI Specialized Programs of Research Excellence (SPORE). The content is solely the responsibility of the authors and does not necessarily represent the official views of the US National Institutes of Health. This research was also partially supported by the Proteogenomic Research Program through the National Research Foundation of Korea, funded by the Korea government [MSIP], and the correspondences to the Seoul site should be addressed to Y Kim.

References

1. Mann M, Kulak NA, Nagaraj N, Cox J. The coming age of complete, accurate, and ubiquitous proteomes. *Mol Cell*. 2013; 49:583–590. [PubMed: 23438854]
2. Lemeer S, Heck AJ. The phosphoproteomics data explosion. *Curr Opin Chem Biol*. 2009; 13:414–420. [PubMed: 19620020]
3. Rifai N, Gillette MA, Carr SA. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotechnol*. 2006; 24:971–83. [PubMed: 16900146]
4. Hoofnagle AN, Wener MH. The fundamental flaws of immunoassays and potential solutions using tandem mass spectrometry. *J Immunol Methods*. 2009; 347:3–11. [PubMed: 19538965]
5. Chace DH, Kalas TA. A biochemical perspective on the use of tandem mass spectrometry for newborn screening and clinical testing. *Clin Biochem*. 2005; 38:296–309. [PubMed: 15766731]
6. Picotti P, Bodenmiller B, Aebersold R. Proteomics meets the scientific method. *Nat Methods*. 2013; 10:24–27. [PubMed: 23269373]
7. Gillette MA, Carr SA. Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry. *Nat Methods*. 2013; 10:28–34. [PubMed: 23269374]
8. Method of the Year 2012. *Nat Methods*. 2013; 10:1. [PubMed: 23547284]
9. Addona TA, et al. A pipeline that integrates the discovery and verification of plasma protein biomarkers reveals candidate markers for cardiovascular disease. *Nat Biotechnol*. 2011; 29:635–643. [PubMed: 21685905]
10. Whiteaker JR, et al. A targeted proteomics-based pipeline for verification of biomarkers in plasma. *Nat Biotechnol*. 2011; 29:625–634. [PubMed: 21685906]
11. Lange V, Picotti P, Domon B, Aebersold R. Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol Syst Biol*. 2008; 4:222. [PubMed: 18854821]
12. Picotti P, Aebersold R. Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. *Nat Methods*. 2012; 9:555–566. [PubMed: 22669653]
13. Pan S, et al. Mass spectrometry based targeted protein quantification: methods and applications. *J Proteome Res*. 2009; 8:787–797. [PubMed: 19105742]
14. Liebler DC, Zimmerman LJ. Targeted Quantitation of Proteins by Mass Spectrometry. *Biochemistry*. 2013; 52(22):3797–806. [PubMed: 23517332]
15. Huttenhain R, et al. Reproducible quantification of cancer-associated proteins in body fluids using targeted proteomics. *Sci Transl Med*. 2012; 4:142ra94.
16. Rodriguez H, et al. Reconstructing the pipeline by introducing multiplexed multiple reaction monitoring mass spectrometry for cancer biomarker verification: an NCI-CPTC initiative perspective. *Proteomics Clin Appl*. 2010; 4:904–914. [PubMed: 21137031]
17. Addona TA, et al. Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. *Nat Biotechnol*. 2009; 27:633–641. [PubMed: 19561596]

18. Anderson NL, et al. A human proteome detection and quantitation project. *Mol Cell Proteomics*. 2009; 8:883–886. [PubMed: 19131327]
19. Legrain P, et al. The human proteome project: current state and future direction. *Mol Cell Proteomics*. 2011; 10:M111.009993. [PubMed: 21742803]
20. Picotti P, et al. A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis. *Nature*. 2013; 494:266–270. [PubMed: 23334424]
21. Aebersold R, et al. The biology/disease-driven human proteome project (B/D-HPP): enabling protein research for the life sciences community. *J Proteome Res*. 2013; 12:23–27. [PubMed: 23259511]
22. Picotti P, et al. A database of mass spectrometric assays for the yeast proteome. *Nat Methods*. 2008; 5:913–914. [PubMed: 18974732]
23. Remily-Wood ER, et al. A database of reaction monitoring mass spectrometry assays for elucidating therapeutic response in cancer. *Proteomics Clin Appl*. 2011; 5:383–396. [PubMed: 21656910]
24. MacLean B, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*. 2010; 26:966–968. [PubMed: 20147306]
25. Farrah T, et al. PASSEL: the PeptideAtlas SRMexperiment library. *Proteomics*. 2012; 12:1170–1175. [PubMed: 22318887]
26. Abbatiello SE, Mani DR, Keshishian H, Carr SA. Automated detection of inaccurate and imprecise transitions in peptide quantification by multiple reaction monitoring mass spectrometry. *Clin Chem*. 2010; 56:291–305. [PubMed: 20022980]
27. Reiter L, et al. mProphet: automated data processing and statistical validation for large-scale SRM experiments. *Nat Methods*. 2011; 8:430–435. [PubMed: 21423193]
28. Chang CY, et al. Protein significance analysis in selected reaction monitoring (SRM) measurements. *Mol Cell Proteomics*. 2012; 11:M111.014662.
29. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012; 490:61–70. [PubMed: 23000897]
30. Wood LD, et al. The genomic landscapes of human breast and colorectal cancers. *Science*. 2007; 318:1108–1113. [PubMed: 17932254]
31. Curtis C, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*. 2012; 486:346–352. [PubMed: 22522925]
32. Kao J, et al. Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. *PLoS One*. 2009; 4:e6146. [PubMed: 19582160]
33. Neve RM, et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell*. 2006; 10:515–527. [PubMed: 17157791]
34. Lehmann BD, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest*. 2011; 121:2750–2767. [PubMed: 21633166]
35. Witt AE, et al. Functional proteomics approach to investigate the biological activities of cDNAs implicated in breast cancer. *J Proteome Res*. 2006; 5:599–610. [PubMed: 16512675]
36. Storey JD. The positive false discovery rate: A Bayesian interpretation and the q-value. *Annals of Statistics*. 2003; 31:2013–2035.
37. Naderi A, et al. A gene-expression signature to predict survival in breast cancer across independent data sets. *Oncogene*. 2007; 26:1507–1516. [PubMed: 16936776]
38. Van Laere S, et al. Relapse-free survival in breast cancer patients is associated with a gene expression signature characteristic for inflammatory breast cancer. *Clin Cancer Res*. 2008; 14:7452–7460. [PubMed: 19010862]
39. Frings O, et al. Prognostic Significance in Breast Cancer of a Gene Signature Capturing Stromal PDGF Signaling. *Am J Pathol*. 2013; 182(6):2037–47. [PubMed: 23583284]
40. Menard S, Fortis S, Castiglioni F, Agresti R, Balsari A. HER2 as a prognostic factor in breast cancer. *Oncology*. 2001; 61 (Suppl 2):67–72. [PubMed: 11694790]
41. van 't Veer LJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature*. 2002; 415:530–536. [PubMed: 11823860]

42. Loi S, et al. Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. *J Clin Oncol.* 2007; 25:1239–1246. [PubMed: 17401012]
43. Parker JS, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol.* 2009; 27:1160–1167. [PubMed: 19204204]
44. Prakash A, et al. Platform for establishing interlaboratory reproducibility of selected reaction monitoring-based mass spectrometry peptide assays. *J Proteome Res.* 2010; 9:6678–6688. [PubMed: 20945832]
45. Prakash A, et al. Interlaboratory reproducibility of selective reaction monitoring assays using multiple upfront analyte enrichment strategies. *J Proteome Res.* 2012; 11:3986–3995. [PubMed: 22639787]
46. Paik S, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol.* 2006; 24:3726–3734. [PubMed: 16720680]
47. Keshishian H, Addona T, Burgess M, Kuhn E, Carr SA. Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. *Mol Cell Proteomics.* 2007; 6:2212–29. [PubMed: 17939991]
48. Stahl-Zeng J, et al. High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. *Mol Cell Proteomics.* 2007; 6:1809–17. [PubMed: 17644760]
49. Halvey PJ, Ferrone CR, Liebler DC. GeLC-MRM quantitation of mutant KRAS oncoprotein in complex biological samples. *J Proteome Res.* 2012; 11:3908–3913. [PubMed: 22671702]
50. Madian AG, Rochelle NS, Regnier FE. Mass-linked immuno-selective assays in targeted proteomics. *Anal Chem.* 2013; 85:737–748. [PubMed: 22950521]
51. Whiteaker JR, Paulovich AG. Peptide immunoaffinity enrichment coupled with mass spectrometry for peptide and protein quantification. *Clin Lab Med.* 2011; 31:385–396. [PubMed: 21907104]
52. Ackermann BL. Hybrid immunoaffinity--mass spectrometric methods for efficient protein biomarker verification in pharmaceutical development. *Bioanalysis.* 2009; 1:265–268. [PubMed: 21083162]
53. Whiteaker JR, et al. Evaluation of large scale quantitative proteomic assay development using peptide affinity-based mass spectrometry. *Mol Cell Proteomics.* 2011; 10(4):M110.005645. [PubMed: 21245105]
54. Anderson NL, et al. Mass spectrometric quantitation of peptides and proteins using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA). *J Proteome Res.* 2004; 3:235–44. [PubMed: 15113099]
55. Maclean B, Eng JK, Beavis RC, McIntosh M. General framework for developing and evaluating database scoring algorithms using the TANDEM search engine. *Bioinformatics.* 2006
56. Vizcaino JA, et al. The PRoteomics IDentifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res.* 2013; 41:D1063–9. [PubMed: 23203882]
57. Griffin NM, et al. Label-free, normalized quantification of complex mass spectrometry data for proteomic analysis. *Nat Biotechnol.* 2010; 28:83–89. [PubMed: 20010810]
58. Krokhin OV. Sequence-specific retention calculator. Algorithm for peptide retention prediction in ion-pair RP-HPLC: application to 300- and 100-Å pore size C18 sorbents. *Anal Chem.* 2006; 78:7785–7795. [PubMed: 17105172]
59. Storey JD. A direct approach to false discovery rates. *J R Statist Soc B.* 2002; 64(3):479–498.
60. Tibshirani R, Wang P. Spatial smoothing and hot spot detection for CGH data using the fused lasso. *Biostatistics.* 2008; 9:18–29. [PubMed: 17513312]

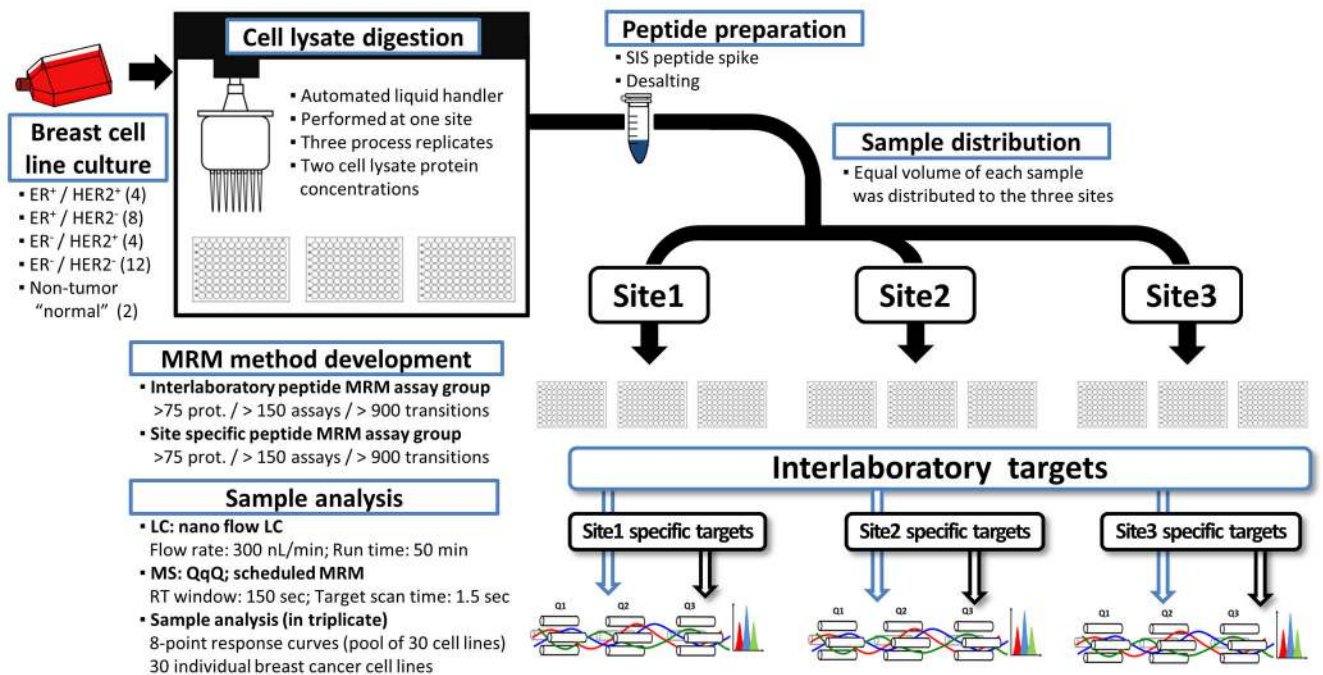


Figure 1. Overview of cell line sample preparation, distribution, and MRM analysis

Thirty cell lines related to breast cancer were prepared in complete process triplicate for analysis by quantitative LC-MRM-MS. For each cell line, 3 aliquots of each of 2 cell lysate protein concentrations (1.0 $\mu\text{g}/\mu\text{L}$ and 0.1 $\mu\text{g}/\mu\text{L}$) were digested by trypsin. A mixture of stable isotope-labeled standards was added prior to desalting the digested peptides. Aliquots were distributed to each performance site where two multiplexed assay groups (one inter-laboratory assay and one site-specific assay) were analyzed on a standardized analytical platform, as described in the Experimental Procedures. The inter-laboratory assay group successfully quantified the endogenous levels of 150 peptides (representing 79 proteins), whereas the site-specific assay groups successfully quantified the endogenous levels of between 147–160 peptides (representing 78–83 proteins; 240 overall) (Supplementary Table 2).

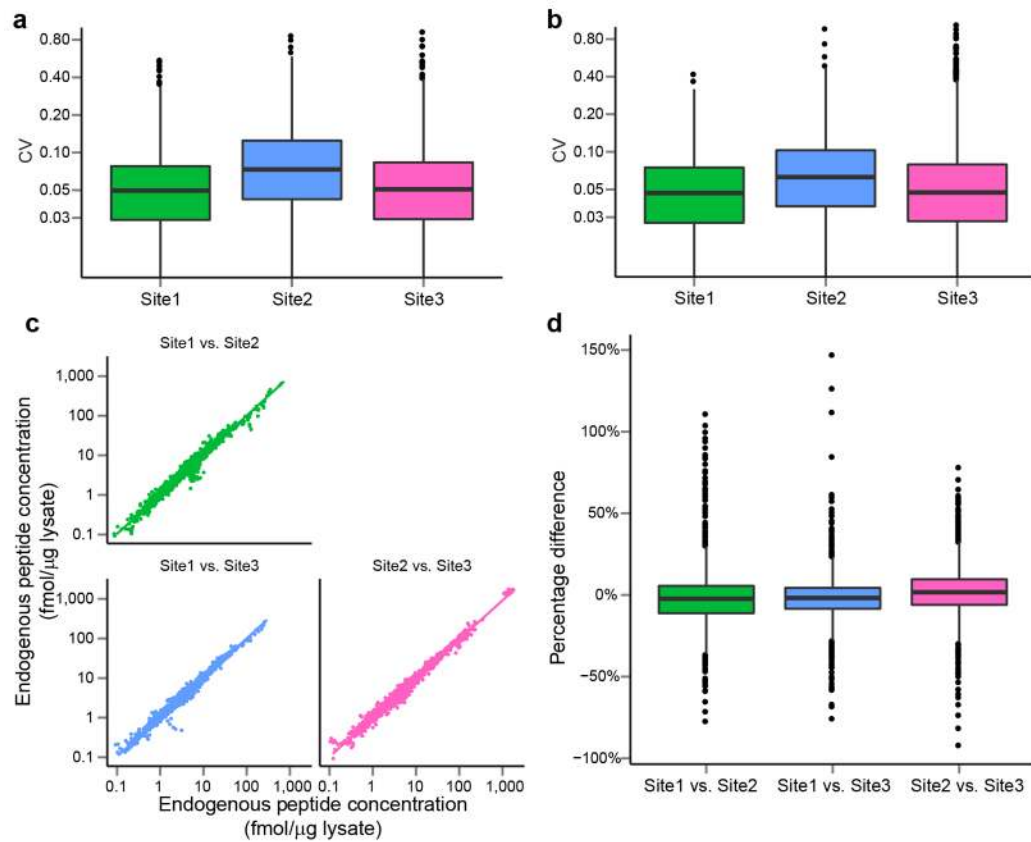


Figure 2. Analysis of cell lysates shows excellent precision of MRM-based measurements in a biological setting, and inter-laboratory assays show high correlation and agreement between sites

(a, b) CV values for the multiplexed assays measured in complete process triplicates, consisting of inter-laboratory targets (150 peptides, 79 proteins) (a) and site-specific target groups (b). At the three sites, the median assay CVs for the inter-laboratory assay group were 5.0%, 7.3% and 5.1%, with 95% of the results having CVs within 15%, 25% and 17%. The site-specific assay groups had median assay CVs of 4.7%, 6.3% and 4.7%, with 95% of the results having CVs within 14%, 20% and 17%. (c) Results for individual peptide measurements were correlated by plotting the peptide amounts measured at the Fred Hutchinson Cancer Research Center, Broad Institute, and Seoul National University - Korea Institute of Science and Technology. For each plot, the x-axis shows the log₁₀ amount of peptide measured at site 1 and the y-axis shows the log₁₀ amount of peptide measured at site 2. (d) A distribution of the percent difference for a pairwise comparison of results. Box plots show the median value plotted as a line with each box displaying the distribution of the inner quartiles and vertical lines show 95% of the data.

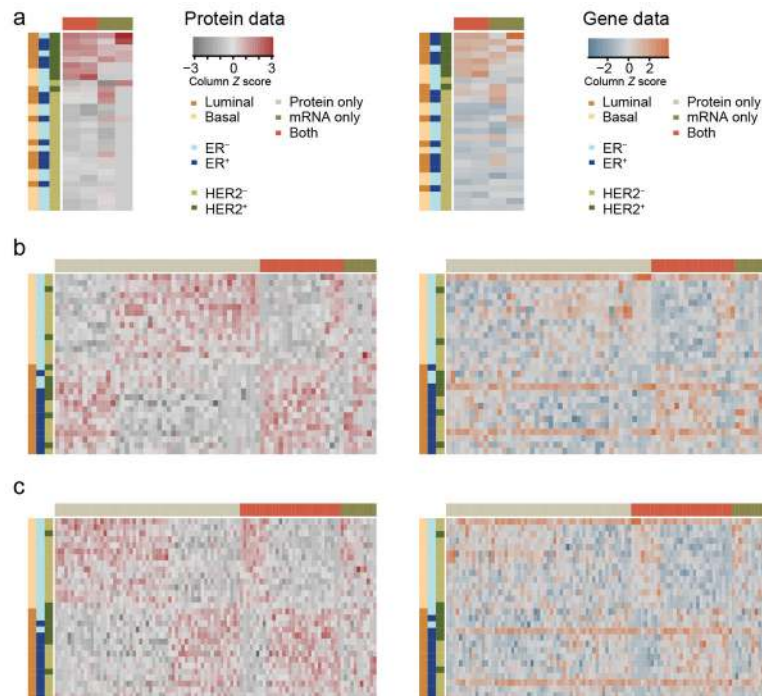


Figure 3. Heat maps for the protein expressions (left column) and RNA expressions (right column) show different genes significantly associated with HER2, ER and basal-luminal³³ status In each heat map, one row represents a sample and one column represents a gene. The color bar on the left side of each heat map illustrates the subtypes of cell lines. The color bar on the top of each heat map illustrates whether only the protein expression, or only the RNA expression, or both expressions of the gene were associated with the subtype. **(a)** Of the 4 genes have significantly different RNA expression levels between HER2⁺ and HER2⁻ cell lines; while only 2 out of the 4 have significantly different protein expression levels. **(b)** Of the 69 genes shown, 25 or 62 have significantly different RNA or protein expression levels between ER⁺ and ER⁻ cell lines respectively, with an overlap of 18 genes. **(c)** Of the 98 genes shown, 42 or 87 have significantly different RNA or protein expression levels between basal and luminal cell lines, with an overlap of 31.

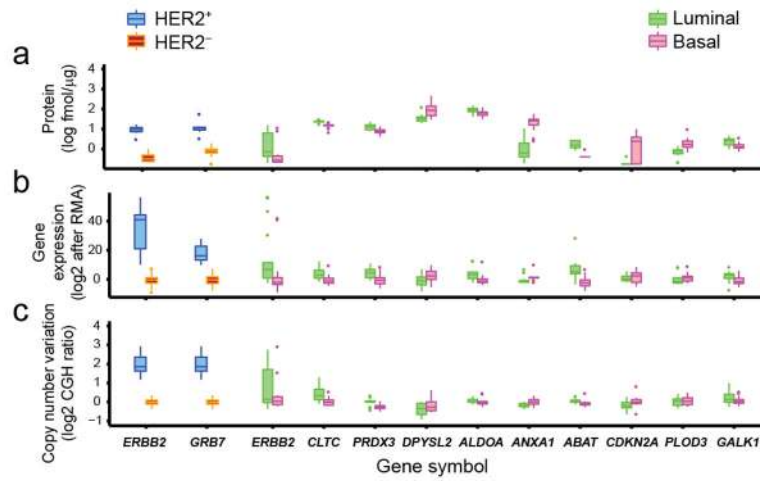


Figure 4. Distribution of protein expression levels, RNA expression levels, and DNA copy numbers of the twelve subtype-enriched genes show high concordance amongst genomic and proteomic datasets

(a, b, c) The protein expression levels measured by MRM (a), RNA expression levels (b) and DNA copy number variation (c) are shown for HER2+ and HER2- cell lines or for basal and luminal cell lines. Two proteins, ERBB2 and GRB7 at chr17, are products of HER2 amplicon genes that show good separation of HER2+ and HER2- groups. The other ten proteins show a difference between the basal and luminal subtypes; the corresponding *P* values from Wilcoxon rank test are all $\leq 1e-4$ with 10k iterations.

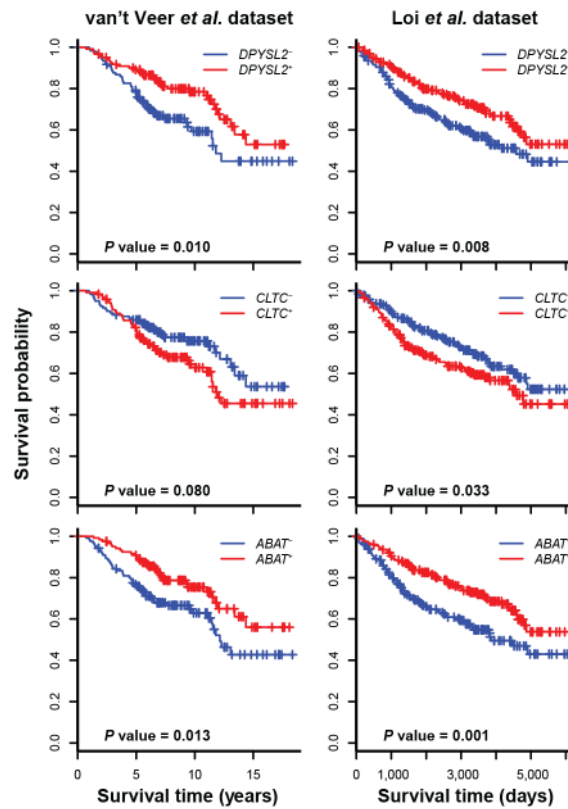


Figure 5. Kaplan-Meier (KM) survival curves of breast cancer patients are stratified by their expression levels of DPYSL2, CLTC or ABAT

Two independent breast cancer datasets^{41, 42} providing both outcome information as well as genomic profiles were used to determine whether the expression of candidate gene products identified in this study show association with outcome. The data are shown for *DPYSL2*, *CLTC* and *ABAT*. For each gene, the breast cancers were classified into high- or low-expressing groups, based on whether or not the expression of the candidate gene was greater than the median expression of the candidate gene. The *P* values from Logrank tests comparing the two KM curves are shown above each figure.