Implementation of a Multiplex and Quantitative Proteomics Platform for Assessing Protein Lysates Using DNA-Barcoded Antibodies*

Jinho Lee‡‖, Gary K. Geiss§‖, Gokhan Demirkan§, Christopher P. Vellano‡, Brian Filanoski§, Yiling Lu‡, Zhenlin Ju‖, Shuangxing Yu‡, Huifang Guo‡, Lisa Y. Bogatzki§, Warren Carter§, Rhonda K. Meredith§, Savitri Krishnamurthy¶, Zhiyong Ding‡‡, Joseph M. Beechem§, and Gordon B. Mills‡‖

Molecular analysis of tumors forms the basis for personalized cancer medicine and increasingly guides patient selection for targeted therapy. Future opportunities for personalized medicine are highlighted by the measurement of protein expression levels via immunohistochemistry, protein arrays, and other approaches; however, sample type, sample quantity, batch effects, and “time to result” are limiting factors for clinical application. Here, we present a development pipeline for a novel multiplexed DNA-labeled antibody platform which digitally quantifies protein expression from lysate samples. We implemented a rigorous validation process for each antibody and show that the platform is amenable to multiple protocols covering nitrocellulose and plate-based methods. Results are highly reproducible across technical and biological replicates, and there are no observed “batch effects” which are common for most multiplex molecular assays. Tests from basal and perturbed cancer cell lines indicate that this platform is comparable to orthogonal proteomic assays such as Reverse-Phase Protein Array, and applicable to measuring the pharmacodynamic effects of clinically-relevant cancer therapeutics. Furthermore, we demonstrate the potential clinical utility of the platform with protein profiling from breast cancer patient samples to identify molecular subtypes. Together, these findings highlight the potential of this platform for enhancing our understanding of cancer biology in a clinical translation setting. Molecular & Cellular Proteomics 17: 10.1074/mcp.RA117.000291, 1245–1258, 2018.

Molecular omics studies have moved to the forefront of cancer research, and recent technological advances in measuring and assessing DNA, RNA, and protein features have helped shape large-scale molecular analysis projects that have significantly enhanced our understanding of cancers and their complexities (1–6). Several studies have reported that the capacity to predict protein expression levels and activity from DNA or RNA analysis is limited. Somatic DNA mutation or mRNA amounts show partial concordance with protein levels (phosphorylation levels) because genomic information is limited in its ability to fully capture protein function mediated by post-translational modification, splicing, or half-life (7–10). Functional proteomic analysis across 11 diseases from The Cancer Genome Atlas (TCGA) projects illustrates distinct pathway activation patterns underlying tumor subgroups that were not identified by genomic and transcriptomic analyses (11), and further analysis of patient samples suggests that protein markers alone and in combination with RNA signatures can predict outcomes (12, 13). These findings support the importance of directly characterizing the functional proteome in addition to events at the DNA and RNA level.

Although several technologies are routinely used for assessing protein expression, they are limited to varying degrees by throughput, sample type and amount, analyte plexity, time, cost, and analytical resources (14, 15). For example, although immunohistochemistry (IHC)† is a gold standard for identifying predictive and prognostic markers in pathology labs (16, 17), each slide can typically only be stained with one or a small number of antibodies at a time. To increase the breadth of protein measurement, other useful high-throughput technologies have been introduced such as mass spectrometry, mass cytometry (multiplexed flow cytometry, CyToF), and Reverse-Phase Protein Array (RPPA). Nevertheless, accurately interpreting data from these platforms can still be challenging from both a technical and analytical standpoint (18, 19), which precludes their incorporation into clinical practice. There is therefore an unmet need to measure expression of multiple total and phospho-protein targets quickly and efficiently from a single sample using methods broadly available to the research community.

From the ‡The University of Texas M.D. Anderson Cancer Center, Department of Systems Biology, 1300 Moursund St., Houston, Texas 77030; §NanoString Technologies, Inc., 530 Fairview Ave N., Seattle, Washington 98109; ¶The University of Texas M.D. Anderson Cancer Center, Department of Pathology, 1515 Holcombe Blvd, Houston, Texas 77030
Received August 31, 2017, and in revised form, February 17, 2018
Published, MCP Papers in Press, March 12, 2018, DOI 10.1074/mcp.RA117.000291

This is an Open Access article under the CC BY license.
Although several approaches have been developed to measure expression of multiple proteins from a single fixed tissue (20, 21), studies examining easy-to-use and clinically translatable multiplex protein expression methods in lysate samples are limited. We therefore directed our efforts to developing a proteomics platform tailored to cell lysate samples. In order to quantify protein expression levels in cell lysates, we adapted a digital color-coded reporter system that counts single molecules without a need for PCR amplification steps (22). This approach extends a proof-of-concept study whereby multiple antibodies were conjugated to unique oligonucleotide (oligo) tags and used to stain intact cells in a single multiplexed protein assay (23, 24). This study was limited however to cell-based assays and lacked comprehensive validation testing to assess the functionality of conjugated antibodies and reproducibility across different users and antibody lots (25). We therefore capitalized on this technology and extended this approach to digitally quantify protein expression in cell lysates. Here, we present the development and validation of a lysate-based proteomics platform involving two oligo-labeled antibody panels using the nCounter platform for detecting oligos (barcodes) conjugated to the antibodies. Similar technologies using oligo-conjugated antibodies have been introduced previously, such as microsphere-based suspension assay (Luminex system, Austin, TX), a Proximity Ligation Assay (PLA), Aptamer-based assay (Somalogic, Boulder, CO) or Meso Scale Discovery (MSD, Kenilworth, NJ) assay. However, these methods still involve technical limitations based on bead-encoding (Luminex) or designing target-specific oligos with chemical modification (Somalogic), or may show less accuracy in quantitatively measuring protein levels because of specific probe amplification (PLA) or fluorescent intensity detection (MSD) (26–28).

Our validation pipeline includes comparisons to orthogonal lysate-based approaches including RPPA, Western blotting (WB), and Enzyme-Linked immunosorbent Assay (ELISA) using samples with high dynamic range of target expression. The broad utility of the platform is supported by the development of both nitrocellulose (NC) and plate-based protocols that generate highly concordant and reproducible results across different antibody panels. Eliminating the need for intact cells and improvements in assay workflow increase the flexibility and utility of the platform. Along with protocol development, we demonstrated the feasibility of this new platform to perform the assays with as low as 250 ng (~1000 cells) of protein within 36 h. Predicted pharmacodynamic effects were readily detected from samples treated with targeted therapeutics, and profiling of breast cancer patient tumors identified molecular subtypes corresponding to known breast cancer subgroups. Collectively, these results highlight the validity and utilization of this protein platform and its potential use in both research and clinical settings.

**EXPERIMENTAL PROCEDURES**

**Cell Culture and Reagents**—Cancer cell lines used in this study were obtained from the National Cancer Institute (NCI) (MCF7, T47D) and the American Type Culture Collection (ATCC) (CEBPs5 kit (MCF10A and HCC1954), IGROV-1, SKBr3, MDA-MB-468, MDA-MB-231, CAMA-1, and OVCA433 cells) were generously provided by Dr. Robert C. Bast (MD Anderson Cancer Center; MDACC). The MCAS cell line was a kind gift from Dr. Joan S. Brugge (Harvard Medical School). Cells were cultured in RPMI 1640 medium containing 5–10% fetal bovine serum (FBS) if not specified. HCT116 AKT1/2 parental and double knock-out (DKO) cell lines were a generous gift from Drs. Bert Vogelstein and Kenneth Kinzler (Johns Hopkins University, Baltimore, MD, USA) and were cultured in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% fetal calf serum (FCS). MCF10A cells were grown in DMEM/F12 containing 5% horse serum, 20 ng/ml of EGF, 0.5 mg/ml of hydrocortisone, 100 ng/ml of chola toxin, and 10 μg/ml of insulin. Cell line identity was confirmed by STR profiling using the MDACC Characterized Cell Line Core Facility. Compositions and growth factors used in this study included EGF (R&D Systems, Minneapolis, MN), the mTOR inhibitor AZD2014 (Selleck Chemicals, Houston, TX), the PI3K p110 inhibitor BYL719 (Stand Up to Cancer: SU2C), the dual PI3K/mTOR inhibitor BEZ235 (Dr. William Borman, MDACC) and the EGF/HER2 dual inhibitor Neratinib (Selleck Chemicals). RPPA data used for cross-platform comparisons included the antibodies listed in Supplemental Table S1. Goat antimouse and rabbit IgG HRP-conjugate (Bio-Rad, Hercules, CA, #172-1011 and #172-1019) were used for WBs.

**Antibody Conjugation and Purification**—

**Amine Labeling**—Custom oligonucleotides (60 nt) containing an internal photocleavable spacer (4-(4,4′-Dimethoxytrityloxy)butymetimimidomethyl)-1-(2-nitrophenyl)-ethyl]-2-cyanoethyl-(N,N-disopropyl)-phosphoramidite) and 5’ amine were purchased from I.D. Technologies (Corvallis, IA). The structure of the photocleavable spacer can be viewed on the IDT website (www.idtdna.com/Catalogue/Modifications/Product/1707). The 60 nt oligos are derived from nonmammalian sequences and have the following characteristic that ensure compatibility with the nCounter Analysis System: (1) Each oligo has two binding regions, one for the capture probe and one for each reporter barcode. The 25 nt capture probe region is present on the 5’ end of the oligo and is the same for all oligos attached to antibodies (i.e. universal). The reporter binding regions is 35 nt long and is unique for each antibody, (2) the melting temperatures for both binding regions is matched to the range of 77–83 °C, (3) each 60 nt sequence is screened for direct and inverted repeats and has a GC content between 40–60%, (4) the 60 nt sequence contains no more than 3 consecutive guanosine residues, (5) the oligo sequences were also screened for alignment to other nucleic acid sequences that are part of the nCounter Analysis System chemistry, and (6) oligos are purified by HPLC by the manufacturer.

Antibodies were purchased in PBS without carrier proteins (see Supplemental Tables S2 and S3 for sources). Oligos were first converted to activate sulfhydryl-reactive species by addition of N-succinimidyl 3-(2-pyridyldithio)propionate crosslinking reagent (SPDP, Molecular Bioscience, Boulder, CO, #67432). Briefly, 200 μM oligonucleotide was adjusted to basic pH by addition of 1 M sodium bicarbonate (1:10 volume, pH 9.0). A 10x molar ratio of SPDP in DMSO was added and the solution was incubated for 20 min at room temperature.
temperature (RT). Reactive groups on SPDP were reduced by adding DTT to a final concentration of 20 mM and incubated for 20 min at 37 °C, after which DTT and free SPDP were removed by purification on a P-6 column (Bio-Rad, #150–3134).

Primary amines on the antibody were converted to reactive maleimide groups as follows: carrier-free antibody in PBS was diluted to 1 mg/ml and brought to basic pH by adding 1 mM sodium bicarbonate (1:10 volume, pH 9.0). A 10-fold molar excess of succinimidyl-4-[N-maleimidomethyl] cyclohexane-1-carboxy-[6-aminocaproate] (LC-SMCC, Molecular Bioscience, #97248) was added slowly to the antibody solution while mixing. After incubation for 20 min at RT, free LC-SMCC was removed by purification on a P-6 column using the manufacturer’s recommendations. The modified antibody was added to the modified oligo solution (1:3 ratio) and incubated for 2 h in the dark at RT. Antibody-oligo conjugates were made double-stranded by addition of 1.2X complementary oligonucleotides and incubated for 1 h. The reaction was quenched by adding 1 mM N-ethylmaleimide in DMSO (1:500 volume). Antibody-oligo conjugates were concentrated to ~200 μl using 30K MWCO filters (Millipore, Burlington, MA, UFC503096) and loaded onto a 1.5 × 50 cm Bio-gel A (Bio-Rad, #1510440) column washed with PBS (pH 7.2) at 2 mM sodium azide. The fraction collection volume was ~1 ml. Fractions containing conjugated antibodies were identified via SDS-PAGE and subsequent staining with both protein and DNA dyes and were then combined and concentrated. Concentrations were measured using a Bradford assay and the degree of labeling was estimated by the nCounter Analysis platform based on molar ratio of oligos to protein. Multiplexed antibody mixes were made by combining conjugates into a single antibody staining mix.

Cysteine Labeling—Oligos were converted to reactive maleimide groups as follows: 200 μM oligonucleotide was adjusted to basic pH by addition of 1 mM sodium bicarbonate (1:10 volume, pH 9.0). A 10-fold molar excess of succinimidyl-4-[N-maleimidomethyl] cyclohexane-1-carboxy-[6-aminocaproate] (LC-SMCC, Molecular Biosciences, #97248) in DMSO was added and mixed. The solution was incubated for 20 min at RT. Free LC-SMCC was removed by purification on a P-6 column (Bio-Rad) using the manufacturer’s recommendations.

Endogenous sulfhydryl groups in the antibody were liberated by reduction with 2 mM TCEP (Pierce, Waltham, MA, #77720) diluted in PBS and the solution incubated for 20 min at 37 °C. Free TCEP was removed by purification on a P-6 column using the manufacturer’s recommendations. The modified antibody was added to the modified oligo solution and incubated for 2 h at RT in the dark. Antibody-oligo conjugates were made double-stranded by addition of 1.2X complementary oligonucleotides (IDT) and incubated for 1 h. Antibodies were concentrated, purified and mixed as described above.

Cell Perturbation Assays—MDA-MB-468 cells (1 × 10^6 cells) were incubated in 6-well plates (35 mm) overnight in low serum (2% FBS) prior to EGF stimulation (20 ng/ml for 30 min) and Neratinib treatment (10 μM for 4 h). For the PI3K pathway inhibitor study, MCF10A, MCAS, and HCC1954 cells (1 x 10^5 cells) were seeded into 96-well plates, incubated in 2% FBS media overnight, and then treated with a dose-range (0 μM, 0.013 μM, 0.04 μM, 0.12 μM, 0.36 μM, 1.1 μM, 3.3 μM, and 10 μM) of AZD2014, BYL719, and BEZ235 for 72 h. Following addition of 5 μl PrestoBlue (Invitrogen, Carlsbad, CA) cell viability reagent to each well, plates were further incubated for 4–8 h and fluorescence readings of all cells were measured at 530 nm excitation/604 nm emission (Fluoroskan, Labsystems) to assess viability. The IC50 doses were then determined for each inhibitor across each cell line using a dose-response model (DRM) (24) implemented in open source software R package “drc” (supplemental Table S4). A four-parameter log-logistic function of the DRM model was applied to fit the dose-response data. To control constrained optimization, we used mean least square estimation and 1000 iterations. For the nCounter lysate assay, 3 x 10^5 cells were cultured in 6-well plates with media supplemented with 5% FBS. Before compound treatment, complete media was replaced with low serum media and cells were incubated overnight. Cells were then treated with their respective IC50 doses of each inhibitor for 24 h and subjected to WB, RPPA, and nCounter analysis.

WB Analysis—Cells were lysed with ice cold RPPA buffer (1% Triton X-100, 50 mM HEPES, pH 7.4, 150 mM NaCl, 1.5 mM MgCl2, 1 mM EGTA, 100 mM NaF, 10 mM Na4P2O7, 1 mM Na3VO4, and 10% glycerol, containing freshly added protease and phosphatase inhibitors from Roche Applied Science, Penzberg, Germany, Cat. # 05056489001 and 04068637001, respectively). The lysed samples were incubated on ice for 20 min and the supernatants were collected by centrifugation at 13,000 rpm for 10 min. Lysate samples were denatured with 1X SDS sample buffer containing 9% 2-mercaptoethanol (Boston BioProducts, Ashland, MA, #BP-111R) and boiled at 95 °C for 5 min. Samples were subjected to SDS-PAGE and transferred to nitrocellulose membranes (0.45 μm, Amersham Biosciences, Little Chalfont, United Kingdom). Membranes were blocked with 5% BSA diluted in TBS-T (Tris-buffered saline, 0.1% tween 20) for 1 h at RT and incubated with primary antibody diluted in blocking buffer at 4 °C overnight. For DNA labeled-antibody incubation, Dextran Sulfate (MW 200 kDa, 5 mg/ml) was added to the blocking buffer to reduce nonspecific binding. After washing with TBS-T (3 ×), membranes were incubated with secondary antibody (1:5000) for 1 h at RT, washed again with TBS-T (3 ×), and then incubated with HRP-enhanced chemiluminescence (ECL) reagent (GE Healthcare, Little Chalfont, United Kingdom) for 1 min. Protein signals were detected by autoradiography.

Patient Sample Collection—For clinical sample testing, 9 fresh breast tumor tissue samples were collected from surgical excision specimens of patients newly diagnosed with breast cancer at MDACC (Houston, Texas) under an institution review board approved research protocol (PA14–1036). Collected tumor tissue samples were immediately stored in tubes with Precellys ceramic beads (1.4 mm, Cayman Chemical, Ann Arbor, MI, Cat. #10011152) at −80 °C. Lysate samples were prepared using RPPA lysis buffer as described previously (25). The histologic diagnosis and biologic characterization of the invasive mammary carcinoma using standard prognostic and predictive markers was established by Dr. Savitri Krishnamurthy (MDACC), a breast cancer pathologist and collaborator for the study.

nCounter Lysate Protocol Development and Optimization—Cell lysates were prepared using RPPA lysis buffer as described above or Tris-SDS lysis buffer (2% SDS, 100 mM Tris pH 6.8, 50 mM DTT). After denaturing proteins with 1X SDS sample buffer and boiling at 95 °C for 5 min, the residual SDS was removed using spin columns (Pierce, #87776 or #88305) to improve nCounter signals. For the NC protocol, a total of 250–500 ng/well of lysate was loaded onto 16-well nitrocellulose membrane embedded slides (FlexWell, GraceBio Lab, Bend, OR) and samples were dried for 15 min at RT. Samples were then incubated with blocking buffer (5% BSA/TBS-T including 100 μg/ml of salmon sperm DNA and 5 mg/ml of Dextran Sulfate) at RT for 1 h. For the plate-based protocol, lysate samples were incubated in high-absorbance ELISA plates (NUNC, Roskilde, Denmark) overnight at 4 °C and blocked with cell staining buffer (Biolegend, San Diego, CA) with the same concentrations of salmon sperm DNA and Dextran Sulfate listed above for 1 h at RT. Nitrocellulose slides were then incubated with a pooled mix of the DNA-conjugated antibodies (100 ng/ml per antibody) overnight at 4 °C, whereas plates were incubated with the pooled antibody mix (50 or 100 ng/ml per antibody) for 2 h at RT. After primary antibody incubation, wells of the slides/plates were
washed for 10 min with TBS-T (6x/H11003) with gentle agitation and then incubated with RLT buffer (Qiagen, Hilden, Germany) for 10 min at RT to retrieve the DNA barcodes. Optical color-coded probe pairs (Tagset) were hybridized with target DNA barcodes and subsequent cartridge preparation and barcode readout were performed using the nCounter Analysis platform according to the manufacturer’s instructions.

**Data Analysis**—NanoString’s nSolver™ 3.0 software was used to analyze data from the nCounter Analysis System. Raw data were normalized with positive hybridization controls and then to the geometric mean of all protein targets within each sample across the experimental data set. Clustering analysis was performed using Gene Cluster3.0 and TreeView1.6 (29).

**RPPA Analysis**—Lysate samples were prepared using RPPA lysis buffer and analyzed as previously described (30). Signal intensities were processed with the loading control normalization integrated in the R package SuperCurve (31).

**RESULTS**

**Study Design and Workflow**—In this study, we demonstrate lysate based validation and proof of concept studies showing that barcoded antibodies can be applied for protein profiling in lysate samples from cancer cell lines and patient tumors. The overall barcoding, testing, and validation workflow for the protocols and antibody panels is outlined in Fig. 1. Unique 60 nt DNA sequence tags were designed and conjugated to carrier-free antibodies using photocleavable linkers (see Experimental Procedures). Both an NC slide and a plate-based protocol were independently developed to allow for broad use of the system. The antibody panel validated using the NC slide protocol includes many key cancer signaling total and phospho-proteins, with an emphasis on breast cancer mark-
ers (supplemental Table S2). A second antibody panel with additional PI3K and MAPK pathway targets was developed in parallel and validated using the plate-based protocol (supplemental Table S3). For the NC slide protocol panel validation, barcoded antibodies were tested via WB on cell line samples with high dynamic range of target expression to assess functionality of the barcoded antibody compared with standard unlabeled antibodies. If no bands of the predicted molecular weight were observed, then the antibody “failed” and was removed. If a dominant band of the right size was detected with the predicted trend of expression and with minimal background, then the antibody passed to the next phase of the validation process and was added to an antibody mix and tested in nCounter on the same sample set (Fig. 1, left). Geomean normalized nCounter data was compared with both the WB results and RPPA data generated from different lysates from the same sample set, and a final validation call was given for each respective antibody. Antibodies were categorized as “Valid” if the nCounter data corresponded to WB analysis and showed high correlation with RPPA data (Pearson’s r $\geq 0.65$), “Use with Caution” if the data still captured the predicted expression trends within the sample in the presence of minor background signal or less abundant non-specific bands, or “Failed” if the nCounter results were discordant with the WB or RPPA data. The final panel used in all analyses excludes antibodies classified as “Failed,” and only includes one antibody target (Bcl-2) that is “Use with Caution” (supplemental Table S2).

**Validation of DNA-barcoded Antibodies**—Representative examples of data generated by the different NC slide protocol validation approaches are shown for p-EGFR (Y1173), HER2, Progesterone Receptor (PR), and GATA3 antibodies (Fig. 2). Signaling trends in the WB were compared between standard (with carrier), unlabeled carrier-free (Non-BC), and barcoded (BC) antibodies, and then barcoded antibody signals were compared with nCounter and RPPA data (generated using standard antibodies). MDA-MB-468 (MDA468) cells overexpress EGFR, and induction of p-EGFR is observed as expected in the presence of EGF stimulation across all platform assays (Fig. 2A). SKBr3 is a HER2-amplified breast cancer cell line and we indeed observed markedly higher HER2 signals in this line compared with MCF7 and MDA-MB-468 cells across all platforms (Fig. 2B). MCF7 is a luminal breast cancer cell line that expectedly exhibited higher expression levels of PR and GATA3 compared with SKBr3 and MDA-MB-468 cells (Fig. 2C, 2D). Results from the other NC slide panel antibodies are presented in supplemental Fig. S1, and indicate that barcoding and running in a multiplex mix do not affect the functionality of the antibodies when comparing with WB and RPPA results.

A parallel approach was taken to validate the antibodies in the plate-based protocol (Fig. 1, right). Antibodies were screened against a variety of cell line lysates and benchmarked against WB and ELISA. Overall expression patterns were similar for WB, direct ELISA, and the nCounter Analysis system (supplemental Fig. S2). Strong correlations ($r \geq 0.98$) in relative protein expression between the nCounter plate-based assay and ELISA indicate that the two assays are comparable (supplemental Fig. S3). We then assessed the performance of the plate-based assay relative to commercially available sandwich ELISA kits by examining five antigens in lysates from A431 cells treated with either EGF or the PP1/PP2A phosphatase inhibitors calyculin A and okadaic acid. In all cases, the sensitivity of the nCounter assay was equal to or better than the commercial sandwich ELISA (supplemental Fig. S4), suggesting that multiplexing of antibodies does not interfere with antigen recognition. To further address potential concerns that arise when combining 31 antibodies in a single assay, we ran antibodies both one at a time or as a combined mix on a single pooled sample lysate of A431 cells + calyculin A and okadaic acid, A431 cells + EGF, and MCF7 cells. The correlation of log$_2$-normalized signal was high ($r = 0.99$), indicating that the risk of reduced specificity or interference between antibodies in the multiplexed format is low (supplemental Fig. S5). Taken together, these data show that the results of a single multiplexed assay are comparable to running multiple ELISA-based assays.

In order to further assess the functionality of phospho-antibodies, we tested both panels using the respective nCounter lysate protocols with the A431 cell lysates described above. Phosphatase inhibitor treatment stabilized expression of the phospho-Ser/Thr AKT, p38, ERK1/2, MEK1, and S6 proteins but not the phospho-tyrosine (Tyr) proteins p-EGFR, p-cMET, and p-HER2 in the NC slide panel as expected based on the Ser/Thr phosphatase used, further supporting that these labeled antibodies function as they should (supplemental Fig. S6A). Similar findings were observed with the phospho-Ser/Thr and phospho-Tyr antibodies of the plate-based protocol panel (supplemental Fig. S6B).

**Reproducibility of Technical, Biological, and Across-batch Replicates**—Reproducibility across replicates and “batches” is one of the most important factors in evaluating performance of a molecular assay platform, particularly in the pathology lab. We assessed reproducibility in our protein platform by testing lysates from technical (same lysate) and biological (different lysate) replicates of MDA-MB-468 cells treated with either vehicle (DMSO), EGF, or EGF + Neratinib across two separate sample cartridges/batches (Fig. 3A). The correlations between technical (Pearson’s $r = 0.98$) and biological ($r = 0.94$) replicates were markedly high across the Control, EGF, and EGF + Neratinib treatment groups (Fig. 3B), suggesting that multiple replicates may not be needed from a technical and potentially a biological perspective. Batch effects stem from nonbiological factors causing variation between experimental runs and are common to almost all high-throughput molecular assays (32, 33). To test the potential for batch effects in our system, we measured protein expression
levels of the same lysate samples across two cartridges run independently at different times. For this we compared the fold-change signals of EGF- and EGF/H11001 Neratinib-treated MDA-MB-468 cells to control cells (Fig. 3C), as well as basal protein expression of a panel of unmodified and engineered cancer cell lines (Fig. 3D). ERK1/2 was blotted in one sample set as an expression reference control. Molecular weight markers are shown at the right side of each blot. Each antibody signal was compared with expression data from nCounter (n = 2, average geomean values) and RPPA (normalized linear values, n = 1). Error bars represent standard error of the mean (S.E.). GATA3 (FL): GATA3-Full-length; GATA3 (T): GATA3-Truncated.

Fig. 2. Validation of DNA-labeled antibodies. A, WBs for Phospho-EGFR (Y1173) antibody across MDA-MB-468 cells treated with DMSO (Ctrl) or 20 ng/ml EGF for 30 min, and B, HER2, C, Progesterone Receptor (PR), and D, GATA3 antibodies across breast cancer cell lines. Samples were blotted with the standard “with carrier” antibodies (left), carrier-free nonbarcoded antibodies (middle; Non-BC), and carrier-free DNA-barcoded antibodies (right; BC). ERK1/2 was blotted in one sample set as an expression reference control. Molecular weight markers are shown at the right side of each blot. Each antibody signal was compared with expression data from nCounter (n = 2, average geomean values) and RPPA (normalized linear values, n = 1). Error bars represent standard error of the mean (S.E.). GATA3 (FL): GATA3-Full-length; GATA3 (T): GATA3-Truncated.
Fig. 3. Reproducibility of nCounter protein platform results. A, MDA-MB-468 cells stimulated with EGF (20 ng/ml, 30 min) in the presence or absence of Neratinib (NRT; 10 μM, 4 h) were prepared in two independent sets (biological replicates, B1 and B2) and loaded twice (technical replicates, T1 and T2) in two separate cartridges (#1 and #2). B, Technical replicates for both B1 and B2 samples (left panel) and biological replicates (right panel) from Cartridge #1 were tested in nCounter for determining the reproducibility of the platform. Pearson’s correlation coefficients (r) were calculated and indicated in the plots. Cartridge-to-cartridge (“batch”) effects were assessed by hierarchical clustering of the fold-change ratio of treated versus untreated MDA-MB-468 protein expression C, and geomean normalized protein expression values from a broad cell line sample panel D, across 2 cartridges (#1 and #2). Pearson’s correlation coefficients (r) between cartridges are shown above the sample dendrograms. DKO; double knock-out. *Bcl-2; Epitomics clone was used for a plate protocol.
(SKBr3 and HCC1954) breast cancer cell lines clustered with one another. Together, our data shows that the results from the NC slide protein assay are highly reproducible within and across cartridges for both technical and biological replicates, abrogating the need for normalization algorithms which have the potential to over-correct the data and reduce its information content.

In addition to biological sampling and cartridge affects, other sources of variation were considered and examined. First, the associations between results generated by different users and using different antibody lots were analyzed. We tested reproducibility between users by performing the plate-based lysate protocol across 3 different samples and users, whereby all other assay reagents and lots were held constant. The correlation of log₂ protein signal was universally high between users for the A431, A431 + EGF, and BT474 cell lysates tested, ranging from $r = 0.97$ to $r = 0.99$, indicating the robustness of the assay protocol (supplemental Fig. S7). For antibody lot testing, we compared two antibody mixes across the same set of samples. The two mixes consisted of the same antibody clones that were coupled to their respective oligonucleotide sequences at different times and by different personnel. A subset of the antibodies were from the same vendor lot of unconjugated antibody whereas others represented different source lots. Although there are differences in the absolute signal between antibody lots presumably because of slight differences in conjugation efficiency or affinity (supplemental Fig. S8A and S8B), the correlation of relative expression levels between lysate samples across two lots of antibody remains high. The Pearson’s correlation coefficients for the log₂ fold-change measurements of A431 + Calyculin A and BT474-M1 to untreated A431 across all antibodies are both above 0.97 (supplemental Fig. S8C). In order to assess which aspects of the antibody labeling process might contribute to variation in signal between lots, we compared signals between antibodies from different vendor lots and separate conjugations (supplemental Fig. S8A) and those from the same vendor lot and separate conjugations (supplemental Fig. S8B). Variation in coupling efficiency is most likely the largest driver of signal differences between lots, but these correlations remain exceptionally high ($r = 0.92$–0.96) across the 3 lysates tested. Variations in antibody source lot are well documented however (34), and further studies are needed to address the degree of involvement in this platform.

**Cross-platform and Protocol Evaluation—**RPPA is a well-established, high-throughput proteomics platform that quantitatively measures protein expression levels in lysates from cell lines, tissues, and other sample types. We therefore compared protein expression signals between nCounter and RPPA across multiple cancer cell line lysate samples (Fig. 4). We report a strong positive correlation between log₂-transformed nCounter and RPPA data ($r = 0.81$ for NC slide protocol) across 7 cell lines for the same antibody targets (Fig. 4A), whereby 13 out of 19 antibodies between the two platforms were the same clone (see supplemental Tables S1 and S2). Individual target correlation coefficients were also calculated and plotted for all cell lines combined (Fig. 4B) and by cell line (Fig. 4C). The highest correlation value between platforms was 0.99 and the lowest value was 0.66, supporting the feasibility of cancer signaling pathway analysis with the established panel. Further, we also show strong correlations between the NC slide and the plate protocols ($r = 0.86$, supplemental Fig. S9) and between the plate protocol and RPPA ($r = 0.75$, supplemental Fig. S10). Taken together, both nCounter protein lysate protocols were highly concordant with each other and displayed similar signaling trends with an orthogonal array-based proteomics platform.

**Biological and Clinical Application of Protein Lysate Platform—**We next tested our multiplexed antibody panel in a biological model of drug response. Aberrations in the PI3K pathway are common in cancers, leading to an increase in signaling activity that has been the target of many small molecule inhibitors in clinical development (35). We therefore used our platform to assess pharmacodynamic changes in protein signaling induced by PI3K pathway inhibitors. For this, MCF10A, MCAS, and HCC1954 cell lines were treated with PI3K/mTOR signaling pathway inhibitors AZD2014 (mTORC1/2), BEZ235 (dual PI3K/mTOR), and BYL719 (PI3Kδ) for 24 h, with MCF10A serving as a negative control because it is an immortalized but nontransformed human breast epithelial cancer cell line that did not respond phenotypically to inhibitor treatment (Fig. 5 and supplemental Table S4). MCAS and HCC1954 cell lines were tested because they harbor PI3K H1047R mutations and exhibit subsequent PI3K-mediated AKT activation. As expected, inhibitor treatment induced down-regulation of the downstream targets p-AKT and p-S6 in HCC1954 and MCAS cells that was observed in both WBs and nCounter, whereas basal signals of these proteins were predictably very low and showed marginal changes upon treatment in MCF10A cells (Fig. 5A) (36). We assessed the signaling response of the cell lines to each compound and observed that most treated samples tested using the two different protocols cluster together following normalization to untreated samples, with the nonresponding MCF10A group forming a separate cluster from the other lines (Fig. 5B). Marked pharmacodynamic changes in p-S6 and p-AKT are observed in HCC1954 and MCAS cells that appear generalizable across each inhibitor, with the degree of change slightly different across specific samples. Protein expression signals across the two protocols are highly correlated using both control-normalized ($r = 0.92$) and unnormalized ($r = 0.90$) results, further supporting the clustering analysis and cross-utility of the nCounter protocols (Fig. 5C).

We next examined whether the antibody panel could identify or discriminate between tumor subtypes using patient samples. We collected frozen lysate samples from 9 breast cancer patient tumors with known biological subtypes and
Fig. 4. Cross-platform and protocol comparison. A, Lysates from multiple cancer cell lines (see Experimental Procedures) were tested in RPPA \( n = 1 \) and nCounter \( n = 2 \), average using the NC slide protocol and Pearson’s correlation between log2-expression is shown. B, C, Pearson’s correlation coefficients \( r \) were calculated between log2-expression for individual antibodies (Total ERK1/2 was excluded as an internal control). The Table indicates targets for which one cell line drives the correlation, either because of markedly high or low expression compared with the other lines.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Cell line</th>
<th>Expression</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bcl-2</td>
<td>MCF7</td>
<td>High</td>
<td>High expression in luminal</td>
</tr>
<tr>
<td>(p-) HER2</td>
<td>SkBr3</td>
<td>High</td>
<td>HER2 amplified</td>
</tr>
<tr>
<td>EGFR</td>
<td>MDA-MB-468</td>
<td>High</td>
<td>EGFR amplified</td>
</tr>
<tr>
<td>p-MEK1</td>
<td>MDA-MB-231</td>
<td>High</td>
<td>Ras mutated</td>
</tr>
<tr>
<td>p-p38</td>
<td>MDA-MB-231</td>
<td>High</td>
<td>Ras mutated</td>
</tr>
<tr>
<td>PTEN</td>
<td>MDA-MB-468</td>
<td>Low</td>
<td>PTEN null</td>
</tr>
</tbody>
</table>

A Barcoded Antibody Based Proteomics Platform

Molecular & Cellular Proteomics 17.6

1253
Fig. 5. **nCounter** protein platform readily detects pharmacodynamic changes induced by targeted therapeutics. The IC_{50} values of AZD2014, BEZ235, and BYL719 were calculated for each cell line at 24 h treatment. MCF10A (10 μM AZD2014, 10 μM BEZ235 and 10 μM BYL719), MCAS (0.62 μM AZD2014, 10 μM BEZ235, and 10 μM BYL719) and HCC1954 (0.45 μM AZD2014, 1.13 μM BEZ235 and 0.9 μM BYL719) respectively. A, The effects of each compound on expression levels of downstream targets p-S6 (S235/S236) and p-AKT (S473) were analyzed by nCounter (n = 2, average; top) and WB (bottom). Error bars represent standard error of the mean (S.E.). Total AKT was used as a loading control. B, Unsupervised hierarchical clustering of relative expression values of treated samples using nCounter data generated from...
analyzed by nCounter using the NC slide protocol (supplementary Table S5). We combined the resulting expression data with that from cell lines in order to assess potential subtypes (Fig. 6). Data from untransformed (MCF10A), luminal-like/epithelial (CAMA-1, T47D, MCF7, MDA-MB-453, MDA-MB-175), HER2-amplified (SKBr3, HCC1954) and triple-negative/basal-like (MDA-MB-468, MDA-MB-231, MDA-MB-157) cell lines were merged with patient sample data (Fig. 6). Protein clustering showed that patient samples were aligned with distinctive subtypes of breast cancer using well-characterized cell line samples as references. For example, luminal-like MCF7, T47D, and CAMA-1 cell lines expectedly grouped together with patient samples exhibiting high expression of the luminal markers ER, PR, and GATA3 and low expression of EGFR and HER2. We also observed that only one patient sample (#9) showed relatively high expression of HER2 among other samples and was grouped with both HER2-amplified breast cancer cell lines. Overall, these results demonstrate the capacity of our platform to accurately and quantitatively measure pharmacodynamic signaling changes to targeted therapeutics and assess critical prognostic and predictive protein markers, creating an opportunity to translate the technology to clinical practice and trials.

Fig. 6. Protein profiling of breast cancer patient tumor lysates and alignment with known subtypes. Lysate samples from 9 frozen breast cancer patient tumors were run on nCounter using the NC slide protocol. Unsupervised hierarchical clustering was performed with the data from these samples and breast cancer cell lines run previously. Known subtypes of the breast cancer cell lines are presented.

DISCUSSION

Our study describes a digital, highly multiplexed, amplification-free approach for measuring proteins in cell lysates that can be broadly used across the research community. The platform involves combining multiplexed DNA-labeled antibodies, the nCounter system (22, 23), and easy-to-use planar-based protocols similar to ELISA or dot-blot methods. The platform shows high sensitivity and specificity equivalent to other single antibody-based assays (such as WB, ELISA, and RPPA) with an input of ~1,000 cells per experiment. The assay is highly reproducible between technical and biological replicates, independently run batches, and different users and lots of labeled antibody panels. Our study showed pharmacodynamic changes in phospho-protein expression levels that were induced by compounds targeting the PI3K and/or mTOR signaling pathways and provides an example of how protein analysis from nCounter can be utilized to evaluate and validate therapeutic approaches in cancer treatment. Furthermore, we also demonstrated the method’s potential clinical application by using expression data from breast cancer cell lines to help classify 9 breast cancer patient tumors based on their protein expression profiles. Taken together, the results

the NC slide (n = 2) and the plate (n = 3) protocols. Values were normalized by control values of each cell line and averaged across technical experimental data sets. C, Cross-correlation analysis between the plate and NC slide protocols using data from all samples (top) and data normalized to untreated controls (bottom). Pearson’s correlation coefficients (r) are presented.
support the broad utility of the platform for use in pharmacodynamic analysis and analysis of clinically-relevant cancer patient samples.

The ideal clinical protein assay would combine attributes that are currently suboptimal in existing platforms such as multiplexing (IHC, ELISA, and flow cytometry), reproducibility and accuracy (IHC, WB, and array-based assays), timeliness (array-based assays), ease of use and data analysis (mass spectrometry and CyToF), and minimal sample input (mass spectrometry, and antibody or bead-based immunoassays) (28, 37–40). Our approach addresses many of these concerns but there are additional challenges that will need to be considered and additional work that will be required to fully realize its clinical potential. For example, because our assay relies on a single antibody per antigen, ongoing validation of specificity is critical and antibody clone availability may be limited. In this work, we have selected highly validated monoclonal antibodies based on approaches that have been successful in other assays such as RPPA and IHC (38, 41). In addition, the cost for generating carrier-free antibodies required for labeling with DNA barcodes is relatively expensive; however, the overall amount of antibody required is lower than most existing assays and costs can be amortized for selected protein panels. Finally, this protein panel was developed based on total lysate sample detection and thus lacks single cell or spatial resolution information provided in some other approaches (such as CyToF, IHC, and single cell sequencing). Intra-tumoral heterogeneity is a major challenge in treating cancer, and the development of a proteomics platform that can resolve spatial expression will provide an important tool for clinically-relevant cancer research. The spatially-resolved version of this technology, termed Digital Spatial Profiling, has recently been presented by several groups in abstract form (42, 43) and may serve as an informative complement to our current lysate-based assay.

The goal in extracting as much biological information from limited amounts of sample or tissue is to measure protein, RNA and DNA in a single assay. Recent reports have shown this is possible for RNA and protein analytes but multiplexing is usually limited (44) or the techniques involve complicated workflow procedures or analysis (45–47). Importantly, the nCounter system can accommodate RNA- and DNA-targeted probes to quantify all three analytes simultaneously from the same sample, known as 3D Biology™ Technology with a single digital readout. The degree of multiplexing for all 3 analytes is only limited by the number of currently available barcodes (800) and the validation strategies required for clinical use. Approximately 2 million on-target molecules can be read out per sample at costs that are at or below that of clinical sequencing. Studies are ongoing to develop this approach across several sample types and biological systems. Technical reproducibility, ease of use, and simple digital read-out give this barcoded multi-omic platform a promising potential for utilization in clinical oncology practice.

* This work was supported in whole or in part by NanoString Technologies, Inc. which also supports ongoing work supervised by G.B. Mills at The University of Texas MD Anderson Cancer Center, the MD Anderson Cancer Center RPPA and CCSG-CCLC core facility National Cancer Institute (NCI) grant #CA016672, and the Stand Up To Cancer Dream Team Translational Research Grant, a Program of the Entertainment Industry Foundation (SU2C-AACR-DT0209). The frozen breast cancer patient samples were prepared under the supervision of S. Krishnamurthy (#00005011 Research Grant from Caliber ID). G.D., G.K.G., B.F., L.Y.B., W.C., R.K.M., and J.M.B. are employees and stockholders of NanoString Technologies Inc. This work was performed in part under a sponsored research agreement between The University of Texas MD Anderson Cancer Center and NanoString Technologies, Inc. Funding for studies performed under this agreement was provided by NanoString Technologies, Inc.

S This article contains supplemental material.

To whom correspondence should be addressed: The University of Texas MD Anderson Cancer Center, Department of Systems Biology, 1300 Moursund St., Houston, TX 77030. E-mail: gmills@mdanderson.org, jlee17@mdanderson.org, or ggissi@nanostring.com.


REFERENCES


