Lysate Microarrays Enable High-throughput, Quantitative Investigations of Cellular Signaling

Mark Sevecka*1,†, Alejandro Wolf-Yadlin*1 and Gavin MacBeath1,2

1 Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford St., Cambridge, Massachusetts 02138, USA
2 Department of Systems Biology, Harvard Medical School, 200 Longwood Avenue, Boston, Massachusetts 02115, USA.

Correspondence should be addressed to G.M.
Email: gavin_macbeath@harvard.edu; Tel: (617) 432-6902; Fax: (617) 432-6990.

* These authors contributed equally to this work.
† Present address: Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, Massachusetts 02142, USA.

Running title: Lysate Microarrays Enable Systems Biology

This supplemental file includes:
• Supplemental Figures S1, S2, S3, S4, S5
  Captions to Supplemental Tables S1, S2, S3, S4
  (Tables are included as separate Excel files)
• Supplemental Experimental Procedures
• Supplemental Analysis

• Supplemental References
Supplemental Figure S1. Antibody hits on lysate microarrays are specific to a subset of biological contexts.

(A) Histogram of number of antibodies scoring as hits in different numbers of time-course lysate sets. 
(B) Histogram of number of antibodies scoring as hits in different numbers of Flp-In cell line time-course lysate sets. The Flp-In cell lines each expressed a different RTK, but are otherwise isogenic.
Supplemental Figure S2. Correlation between microarray and western blotting data.

(A) Distribution of Pearson correlation coefficients (ρ), showing two populations separated at approximately ρ = 0.75. The population with ρ > 0.75 is exponential in shape and represents true hits. The population with ρ < 0.75 is evenly distributed and likely arose from random noise.

(B) Cross-reactive signals are observed mainly on microarrays. For each antibody-context pair, we determined the y-intercept of a linear fit between the western blot data (x-axis) and microarray data (y-axis), and normalized it relative to the maximum microarray signal within the pair. Shown is a histogram of all normalized y-intercept values. Red indicates additional cross-reactivity on western blots; blue indicates additional cross-reactivity on microarrays.
Supplemental Figure S3. Graphical summary of validation status of all antibodies assessed in this study.

White squares represent non-hits on microarrays. Black squares (■) show microarray hits not tested by western blotting. Red squares (■) are microarray hits rejected based on western blot cross-reactivity or absence of bands. Blue squares (■) show microarray hits with acceptable western blots but poor Pearson correlation with microarray data (ρ < 0.75). Green squares (■) represent microarray hits with acceptable western blots and good Pearson correlation with microarray data (ρ ≥ 0.75). Column headings in blue font indicate lysate sets for which every microarray hit was tested by western blotting. Column headings in black font indicate lysate sets for which only a subset of microarray hits was tested by western blotting.
Supplemental Figure S4. Proposed model for non-equivalence of microarray and western blot data.

Two components contribute to total microarray signals: target-specific signals (detectable by western blotting) and an additional cross-reactive term, which depends on the identity of the antibody and the cellular background and is not observable on western blots. Depending on the relative magnitudes of both terms, microarray signals appear compressed to different extents. Microarrays are useful for signal quantification when the remaining spread in signal exceeds assay noise.
Supplemental Figure S5. Self-organizing map analysis.

(A) Component planes showing signal intensities of each map unit at each time point. Color maps indicate z-scored values.

(B) U-matrix in unit map representation, showing individual distances and map units (average distances).
CAPTIONS TO SUPPLEMENTAL TABLES S1, S2, S3, S4

Supplemental Table S1. List of all lysate sets used in this study.

Supplemental Table S2. List of all antibodies used in this study.

Supplemental Table S3. All microarray data collected in this study.
All values are the average of two analytical replicates, corrected for non-linearity and normalized relative to β-actin intensity.

Supplemental Table S4. All western blot data collected in this study.
All values have been normalized relative to β-actin intensity.
SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Cell lines and reagents

The A431, A549, HeLa, HMEC, MDA-MB-453, MDA-MB-468, SK-BR-3, T-47D and ZR-75-1 cell lines were obtained from the American Type Culture Collection (ATCC, Manassas, VA). HT-29 and Jurkat cells were a generous gift from Douglas Lauffenburger. Flp-In-293 cells expressing six different receptor tyrosine kinases were described previously (1). Human recombinant brain-derived neutrophic factor (BDNF), epidermal growth factor (EGF), fibroblast growth factor 1 (acidic, aFGF), hepatocyte growth factor (HGF), insulin-like growth factor 1 (IGF1), interleukin 7 (IL-7), platelet-derived growth factor BB (PDGF-BB) and tumor necrosis factor α (TNF-α) were purchased from Peprotech (Rocky Hill, NJ). Human recombinant insulin was obtained from Sigma-Aldrich (Saint Louis, MO). Anisomycin and staurosporin were purchased from BIOMOL (Plymouth Meeting, PA). 2% SDS lysis buffer with protease and phosphatase inhibitors was prepared as described (2). Pan and phospho-specific antibodies were obtained from Abcam (Cambridge, MA), Abgent (San Diego, CA), BD Biosciences (San Jose, CA), Cell Signaling Technology (Danvers, MA), EMD Chemicals (Gibbstown, NJ), Epitomics (Burlingame, CA), Invitrogen (Carlsbad, CA), Lab Vision/Neomarkers (Fremont, CA), Millipore (Billerica, MA), Panomics (Fremont, CA), ProteinTech Group (Chicago, IL), Santa Cruz Biotechnology (Santa Cruz, CA) and USBiological (Swampscott, MA). Mouse monoclonal anti-β-actin antibody was purchased from Santa Cruz. Rabbit monoclonal anti-β-actin antibody was purchased from Cell Signaling Technology. Infrared fluorophore-labeled secondary antibodies were obtained from LI-COR Biosciences (Lincoln, NE). Whereas pan-specific antibodies were purchased from a variety of vendors, we focused largely on a single commercial supplier for antibodies targeting post-translational modifications. This is because the production of such
antibodies requires greater technical expertise and, in our experience, PTM-specific antibodies from this supplier consistently outperformed those from other vendors. Detailed information about all antibodies used in this study can be found in Table S2 online.

**Correction for assay non-linearity**

To correct microarray data for any non-linear relationship between antigen concentration and signal intensity, we employed a modified version of the strategy we outlined previously (3). Each microarray contained eight-point serial two-fold dilutions of control lysates, as well as spots containing lysis buffer only, to generate an antibody-specific calibration curve. Because lysates from any particular biological context, at a single time point of treatment, are unlikely to contain high expression levels and high levels of phosphorylation of all proteins targeted by our large set of antibodies, we used serial dilutions of six different lysates, covering a number of biological contexts: A431 cells stimulated with EGF for 5 and 60 min, EGFR-expressing Flp-In 293 cells stimulated with EGF for 5 and 60 min, A431 cells stimulated with anisomycin for 60 min and HeLa cells stimulated with TNF-α for 15 min. We then separately fit the 680 nm and 800 nm channel signal intensities from the six dilution series on each microarray to a Boltzmann sigmoidal equation:

\[ I = d - \frac{a}{1 + e^{bx-c}} \]  

where \( x \) represents the relative concentrations of the lysates in the serial dilution (1, 0.5, 0.25, …), and \( I \) is the signal intensities measured. Beginning with the dilution series that contained the highest non-saturated spot intensity, and continuing in order of decreasing maximum spot intensity, a non-linear fit was attempted for each dilution series to determine fit parameters \( a, b, c, \) and \( d \). Once a goodness-of-fit of \( R^2 > 0.975 \) was reached, the inverse of function (1) was
applied to all spot intensities within this microarray using the fit parameters obtained. Where none of the six dilution series produced an acceptable goodness-of-fit, data were not corrected for non-linearity.

**Calculation of signal upregulation vectors**

Signal intensities within each microarray were mean-normalized to enable statistical comparisons across different arrays. We then normalized signal intensities from target proteins relative to the β-actin signal intensities from the same microarray spots to account for any differences in lysate concentration or spotting. Lastly, we averaged data from duplicate spots. Data from each antibody and biological context were then organized into vectors in two separate ways. For the 20 time course treatments, each data vector consisted of six data points, corresponding to the six different time points of treatment. For comparisons across cell lines, each vector comprised 17 data points, corresponding to the 17 different unstimulated cell lines. We calculated two measures of signal upregulation for each vector: (1) signal difference was defined as the difference between the highest and lowest signal intensity, and fold-upregulation was calculated as the ratio of the highest and lowest signal intensity of each data vector.

**Immunoblotting**

Lysates were separated electrophoretically on 48-well polyacrylamide gels (Invitrogen), and transferred to polyvinylidene difluoride (PVDF) membranes using a dry-transfer apparatus (Invitrogen). Included on each blot was one lane of a protein molecular weight reference standard (Invitrogen). Blots were blocked in 5% dry milk/1× PBS for 1 h at 4 ºC, and washed three times for 5 min each with PBST. Primary and secondary antibody incubation steps were
carried out as described for lysate microarrays, using 1:10,000 anti-β-actin antibody, 1:1,000 pan-or phospho-specific antibody, and 1:5,000 anti-rabbit-680 and anti-mouse-800 antibodies. Immunoblots were then scanned in the 680 nm and 800 nm channels using an Odyssey imager at 84 µm resolution. Using the Odyssey software, signal intensities of the bands were determined using equally sized rectangular integration areas. For each blot and each channel, we also selected a representative area of the blot outside the protein transfer area, and integrated background signal intensity. To assess the specificity of antibody binding, we first identified the dominant band in each blot, as well as the next-lightest and next-heaviest bands within the molecular weight marker lane. An antibody was considered to be specific for its intended target if the true molecular weight of the target protein fell within the range of the true molecular weights of these two reference proteins. Molecular weights of target proteins were as reported by the antibody manufacturer, accounting for any post-translational modifications.

To identify predictors of antibody performance, antibodies were grouped into several categories: (a) rabbit antibodies, (b) mouse antibodies, (c) polyclonal antibodies, (d) monoclonal antibodies, (e) pan-specific antibodies, (e) PTM-specific antibodies, (g) single-phosphorylation-site epitopes, (h) multiple-phosphorylation-site epitopes, (i) phosphotyrosine (pTyr) epitopes, and (k) phosphoserine/threonine (pSer/pThr) epitopes. Antibodies recognizing mixed epitopes of both pTyr and pSer/pThr sites were grouped together with pSer/pThr-detecting antibodies, as they behave similarly based on our previous experience. Because these designations are not mutually exclusive, each antibody falls into multiple groups. For each category (a) through (k), we then determined the fraction of successful western blot validations of microarray hits.
SUPPLEMENTAL ANALYSIS

Our experimental evidence suggests that the cross-reactive signal observed on lysate microarrays is mainly a function of the cell line used. To test this hypothesis, we compared the variation in cross-reactive signal between two groups of antibodies: (1) all 207 antibodies that produced no microarray hits across any of the RTK-expressing (but otherwise isogenic) Flp-In cell lines, and (2) all 160 antibodies that produced no microarray hits across any of the other cell lines in the ‘time courses’ dataset. Based on our experience in comparing microarray and western blot data for such non-hit antibodies (data not shown), we reasoned that signals from these antibodies are likely to be exclusively cross-reactive (with any signal spread due to low-level changes in target proteins or assay noise), with only little contribution from basal levels of target protein. We therefore approximated the cross-reactive signal term for each antibody and lysate set by the minimum signal value observed across the time course treatment. We then defined two metrics to capture variation in cross-reactive signal for each antibody: (1) the standard deviation of the cross-reactive term across lysate sets, given as a percentage of the mean, and (2) the ratio of the highest and the lowest cross-reactive terms observed across lysate sets. By applying both metrics to both groups of antibodies, we obtained for each group two distributions that capture the variation in cross-reactivity. We then used the two-sample Kolmogorov-Smirnov test to determine the significance of any observed differences between the distributions. Using both metrics, we rejected the null hypothesis that the two distributions are from the same continuous distribution or that cross-reactive signal terms are lower across unrelated cell lines than across isogenic lines with \( P \)-values \( P < 10^{-19} \) and \( P < 10^{-26} \), respectively. Conversely, when we tested the null hypothesis that cross-reactive signal varies more across unrelated than across isogenic cell lines, the null hypothesis was accepted for both metrics with \( P \)-values of \( p = 0.97 \) and
$P = 1.00$, respectively. These data lend support to our model that the pan-protein complement in a sample is the primary source of antibody cross-reactivity on lysate microarrays.
SUPPLEMENTAL REFERENCES

