Genome-Wide screen for Protein: Phosphorylated Protein Interactions using Multi-Array™ Technology

John H. Kenten, Steven Cheng, Paul Grulich,
Ilia V. Davydov, Nisar Pampori, Pankaj Oberoi,
Stefanie Nelson, Ryan Swenerton,
Yassamin J. Safiran, Laura K. Schaefer,
Hans A. Biebuyck and Jacob N. Wohlstadter

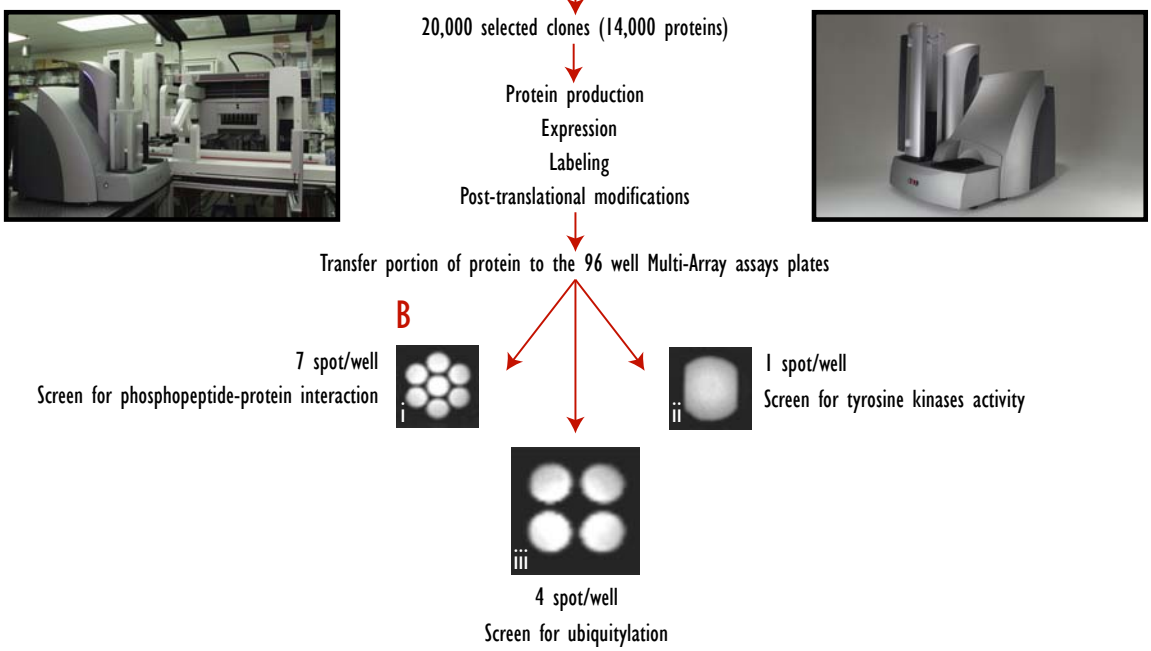
Meso Scale Discovery™
A division of Meso Scale Diagnostics, LLC.

Genome-Wide screen for Protein: Phosphorylated Protein Interactions using Multi-Array™ Technology

❶ Abstract

We demonstrate a proteome wide, high-throughput approach to the discovery of protein binding partners for specific phosphorylation sites on proteins. We selected 18,000 clones encoding 14,000 different proteins (Unigene) for expression and labeling from our library of 61,000 cDNAs, each calculated to be full-length (24,000 proteins). We measured the degree of binding of individual proteins to each member of arrays of specific phosphopeptides using MSD's Multi-Array technology. Characterization of specific binding resulted from the exposure of each protein to an array of select phosphopeptides, 14 in total, from phosphorylation sites on EGFR, NFKBIA, CREB, MYC, RAF1, CDC25C and GLI3, allowing the interrogation of more than 200,000 unique interactions. We qualified hits from this screen using phosphopeptide competition studies to demonstrate the significance and strength of the protein – phosphopeptide binding interaction. Further, by examining the pull down of putative protein partners by intact native phosphorylated proteins, we extrapolate from peptide – protein binding to protein-protein binding. This approach recovered most of the previously known protein-binding partners in blinded fashion (e.g. EGFR pY1068 - GRB2, NFKBIA, pS32pS36 - BTRC, RAF1 pS621 - YWAH family) and uncovered new and significant protein interactions with these phosphopeptides. Our method can be modified to screen for many other putative protein binding interactions on a proteome-wide scale particularly where control of these interactions results from a post-translational modification like phosphorylation or ubiquitylation. We think our demonstrations in tota provide significant validation of Multi-Array technologies as an engine for high throughput, functional biological discovery.

❷ Multi-Array Technology

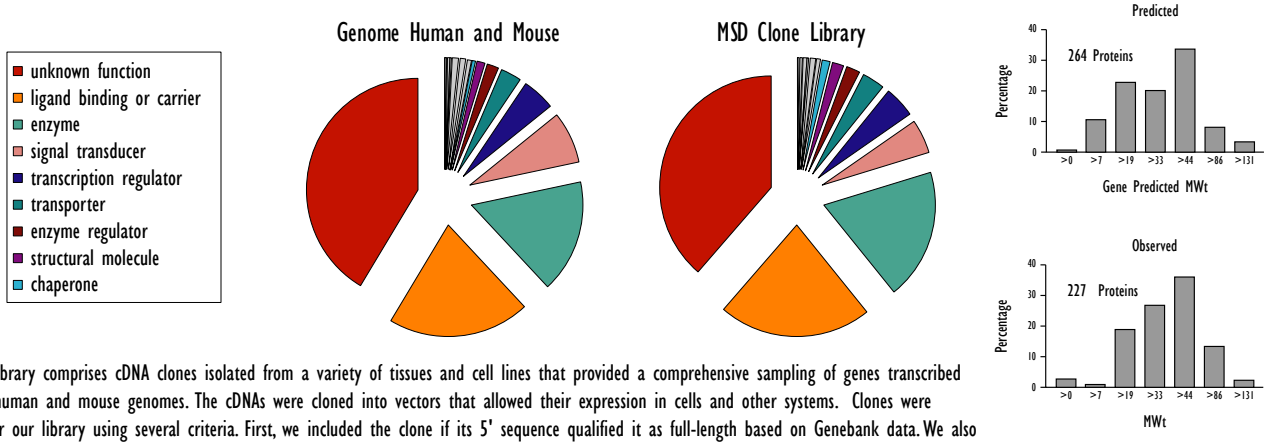


(A) We carried out a proteome wide screen using measurements of 20 different functional assays to survey a number of functional characteristics of each protein of the library including their (i) ability to act as binding partners to specific phosphopeptides, (ii) as substrates for tyrosine kinases or, generally, (iii) ubiquitylation. Common to each assay was the use of proteins prepared, including their modification and labeling, prior to distribution of each protein into individual wells of Multi-Array plates using automation (Figure at left) for subsequent measurement on the SECTOR Imager 6000 (Figure at right). (B) The Multi-Spot™ plates feature one or more spots in each well of a 96-well plate that allowed us to carry out experiments of many different types and in many different formats (see examples below). This poster focuses on the results from one part of the screen, the analysis of proteins ubiquitylated by rabbit reticulocytes lysates using Multi-Spot plates.

Meso Scale Discovery™
A division of Meso Scale Diagnostics, LLC.

Genome-Wide screen for Protein: Phosphorylated Protein Interactions using Multi-Array™ Technology

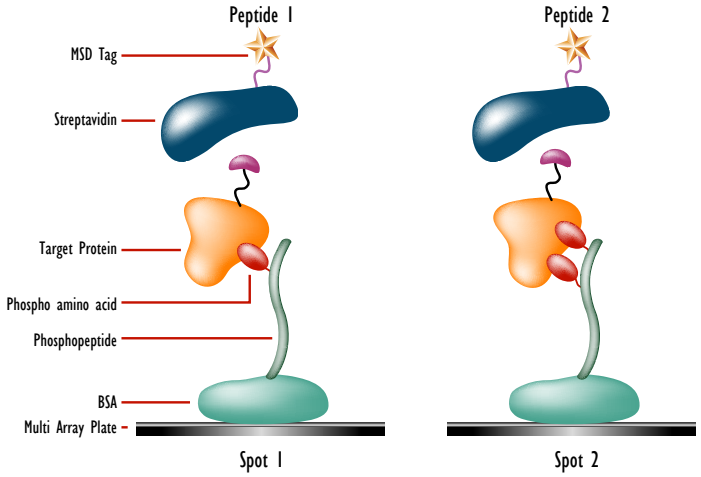
❸ Functional diversity of the MSD clone library



The MSD library comprises cDNA clones isolated from a variety of tissues and cell lines that provided a comprehensive sampling of genes transcribed from the human and mouse genomes. The cDNAs were cloned into vectors that allowed their expression in cells and other systems. Clones were selected for our library using several criteria. First, we included the clone if its 5' sequence qualified it as full-length based on Genebank data. We also included clones in our library if they were from a unique Unigene cluster up to a maximum of 3 clones for any given cluster. We thus effectively normalized the clone collection to ensure an optimal representation of genes within the genome, avoiding the generation of a library with few genes in many copies.
We compared our collection of genes to human and mouse genome annotations based on gene onotologies (pie charts). This comparison generated a profile that largely matched the MSD library to these genomes suggesting a highly representative collection. We also expressed a sample of the library and demonstrated that we produced the expected gene products for >70% of the genes tested (see bar graphs) validating the functionality of the library.

❹ Detection Methodology: Scheme

In order to develop a facile methodology that capitalized on the Multi-Array technology we explored a number of formats to immobilize or detect the interactions of peptides with the proteins expressed from our library. We selected the format illustrated in the figure from among these possibilities, using peptide sequences between 20 and 30 amino acids in length that encompassed the phosphorylation site(s) of interest as the capture agent. In many cases we used peptide sequences previously demonstrated to be useful. In cases where no capture peptide existed a priori, we selected sequences based on the relative position of the phosphorylation site(s) to the N and C terminus and the homologies between species around these sites. In either case, these peptides included a C or N terminal Cys residue to facilitate the coupling of the peptide to BSA. The resultant BSA-peptide conjugate was spotted onto Multi-Array plates to form regions having seven different peptides in each 96-well of the plate. These plates were incubated with individual proteins produced from the library so that results from one well reflected the interaction between each of seven distinct peptides and a distinct protein expressed from the library.



Meso Scale Discovery™
A division of Meso Scale Diagnostics, LLC.

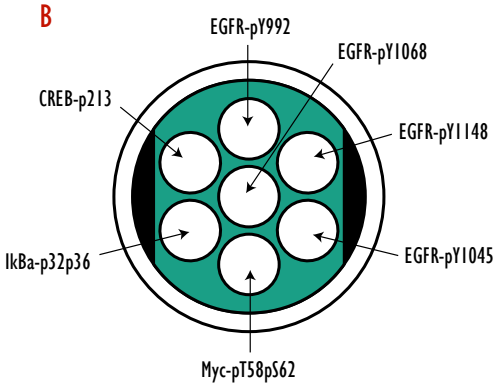
Genome-Wide screen for Protein: Phosphorylated Protein Interactions using Multi-Array™ Technology

5 Detection Methodologies: Multi-Array Assay Formats

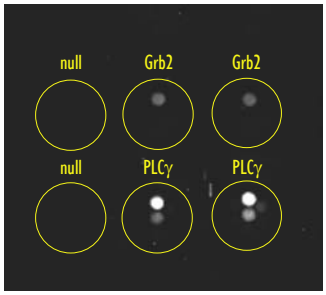
A

Plate 1	Plate 2
EGFR-pY992	Raf1-pS621
EGFR-pY1068	CDC25C-pT48
EGFR-pY1148	Gli3-1
EGFR-pY1045	Gli3-2
IkBa-pS32pS36	Gli3-3
CREB-pS133	Gli3-4
Myc-pT58pS62	Gli3-5

B



C

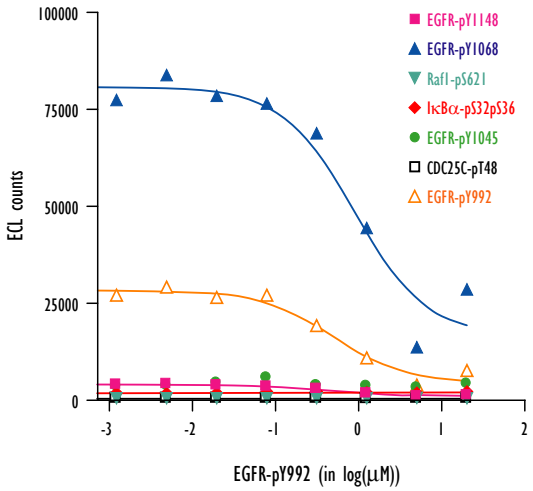


Imager 6000 Image(s)

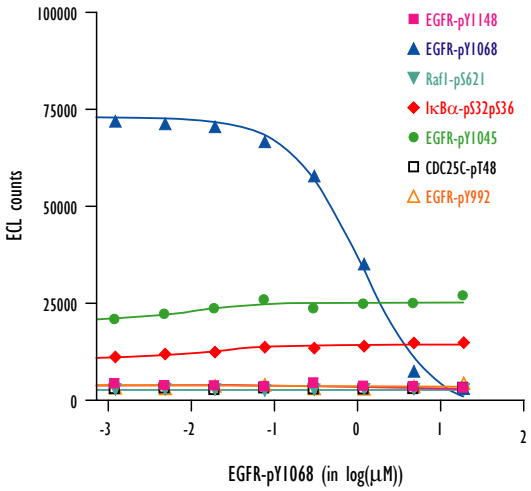
The peptides we used in this study are listed in Figure A. We spotted two plates with 7 spots per well to give us a total of 14 peptides that we tested for binding to the proteins that we produced in our expression system. The organization of these spots for Plate 1 is illustrated in Figure B. Test results are shown in Figure C where we examined interactions with two proteins, Grb2 and PLC-gamma, already known to interact with one of the specific phosphopeptide motifs to validate the spotting and assay approach and further characterize the system. In the screen, we ran Plate 1 and removed the supernatants from this plate for subsequent incubation with Plate 2. We confirmed that this workflow was practical in development and used it because it allowed for significant improvements to the speed and costs of the screen.

6 Specificity of Phosphopeptide Assay

Competitative binding of PLCγ to different phosphopeptides



Competitative binding of Grb2 to different phosphopeptides



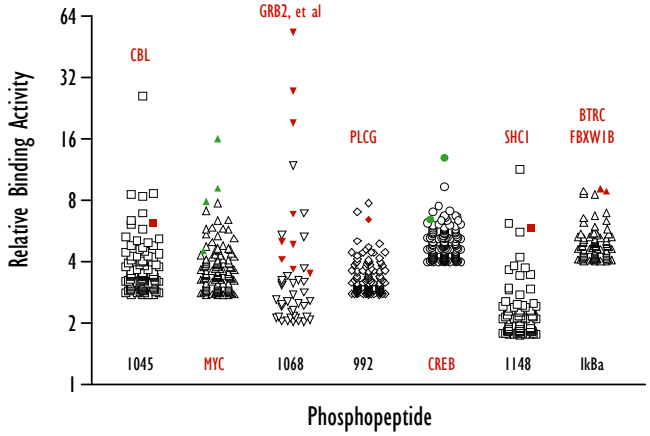
We were concerned that we maintain the specificity of the interactions between the peptide and the proteins from our expression system. We used known protein-peptide interactions to validate our assay protocols and demonstrate the recovery of specific binding partners, confirming these interactions with free peptide competition studies. The figures below illustrate the results from one study of the final assay format with the known binding partners PLC-gamma to EGFR pY1068 and pY992, and Grb2 to EGFR pY1068. Here we show the specific binding and the competition with the cognate peptides to demonstrate the retention of specificity with the assay format and the protocol selected for the study. We also used this approach to qualify further other putative binding interactions (see below).

Meso Scale Discovery™
A division of Meso Scale Diagnostics, LLC.

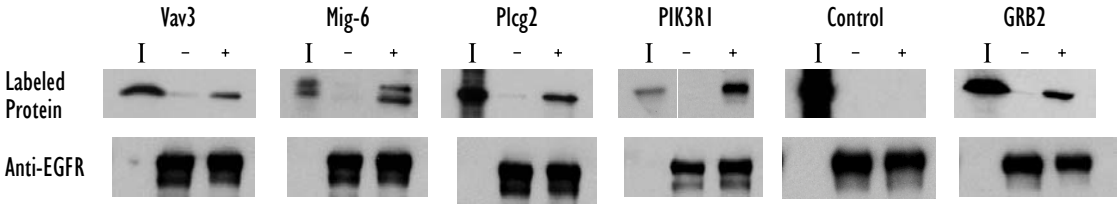
Genome-Wide screen for Protein: Phosphorylated Protein Interactions using Multi-Array™ Technology

7 Hits from Primary Screen of 18,000 cDNA clones

We selected hits based on the signal at a given peptide spot relative the lowest peptide spots in that well (estimate for non-specific background). This method generated a value that was an estimate of the 'Relative Binding Activity', reflecting the specific binding activity that we had for a given protein in the screen. These selected hits were then subjected to retesting and peptide competition. Solid red symbols are known confirmed hits; solid green symbols are novel confirmed hits; open symbols indicate datum that did not reconfirm.



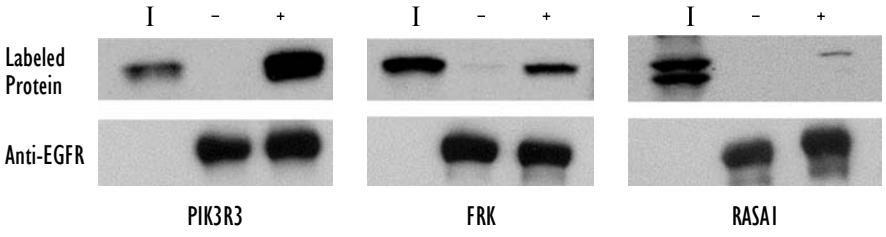
8 Secondary screen of primary hits from screen



In this study we confirmed hits from the primary screen using whole EGFR expressed in cells. We activated EGFR by the addition of EGF (+) or with mock activation (-). These cells were lysed after treatment and the EGFR immunoprecipitated and used to pull down the labeled proteins selected in the primary screen (upper gel panel). The level of EGFR used in each pull down was checked via western blot using anti-EGFR (lower gel panel). These clones represent the clones that were confirmed in this secondary screen. Catna1 was used as a negative control.

9 Interactions of novel hits with native EGFR: Pull down studies

We selected a number of novel hits from the screen that we confirmed using pull down, gel-based assays with native proteins and studies using various gene-reporter constructs. We looked at the binding of three novel hits to activated EGFR as an example. The proteins on the upper portion of the gel are the labeled proteins that were hits from the screen (PIK3R3, FRK, RASA1). The lower portion is a blot of EGFR added in the binding reaction. This is the same assay format as in section 8). Here we show that these hits selected from our screen were in fact able to bind to intact full length EGFR following its activation and phosphorylation in a biological system.

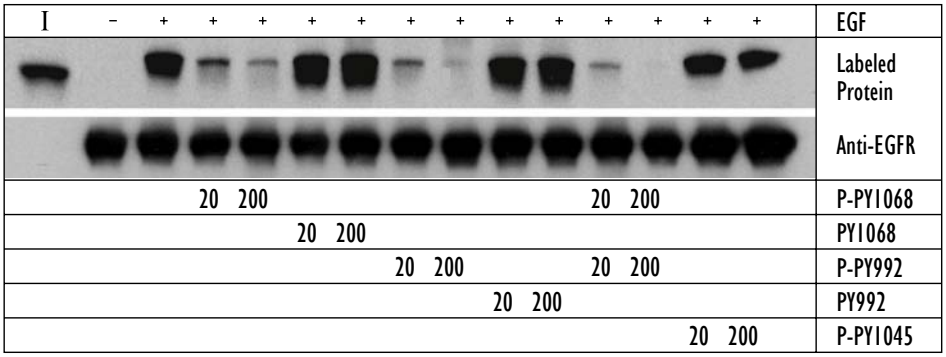


Meso Scale Discovery™
A division of Meso Scale Diagnostics, LLC.

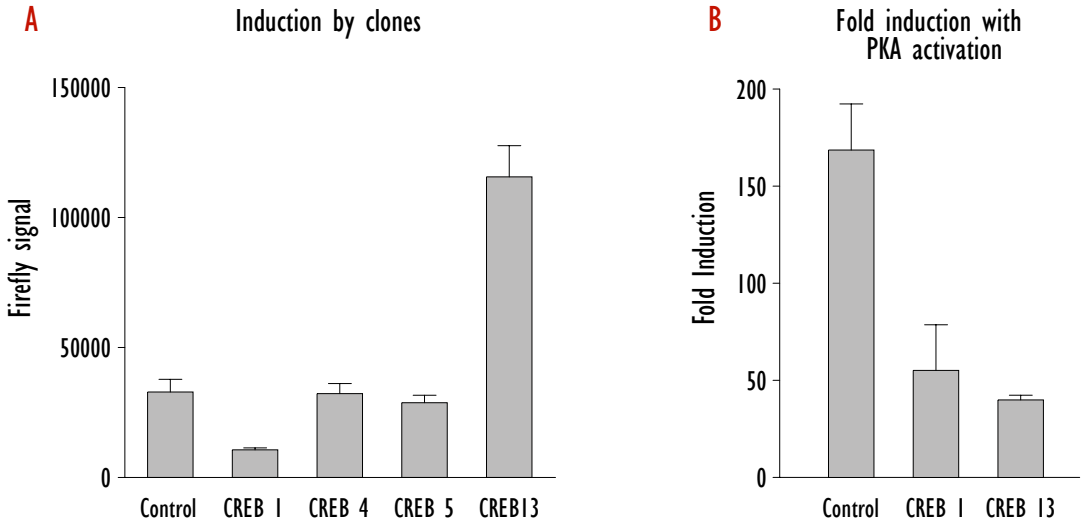
Genome-Wide screen for Protein: Phosphorylated Protein Interactions using Multi-Array™ Technology

⑩ Interactions of novel hits with native EGFR: Peptide competition

To further explore the interaction of one of the novel hits we carried out a peptide competition study to look at the specificity of the interaction and also to investigate which of the phosphorylation sites on native EGFR was involved in the binding interaction. Studies using PIK3R3 as an example demonstrated that this protein was able to bind to both phosphotyrosines at 1068 and 992, but not to the phosphorylation site at 1045. These competition studies also demonstrated that the interaction was phosphorylation dependent as the unphosphorylated peptides failed to compete with this binding.



⑪ Transfections of CREB hits



We determined if genes identified by their binding to a consensus creb phosphopeptide (p133) influence the response of a cre-luciferase reporter. In study A we see that CREB1 and CREB13 have very different effects on the signal from the cre-lucifease reporter with CREB1 lowering and CREB13 activating this reporter. The study B shows the effect of CREB1 and 13 on the induction of cre-luciferase signal by PKA co-transfection. Here we see that both block this measure of induction of the cre reporter plasmid. This blunting of the induction by PKA occurs in different ways: CREB1 blocks the cre signal both with and without PKA stimulation. CREB13 blunts this induction by inducing without PKA and lowering the PKA effect of stimulation, thus resulting in lowering the fold induction.

Meso Scale Discovery™
A division of Meso Scale Diagnostics, LLC.

Genome-Wide screen for Protein: Phosphorylated Protein Interactions using Multi-Array™ Technology

⓬ Validation and Discovery

We used our library of clones to investigate binding to a number of phosphopeptides in addition to the EGFR motifs already described. In the case of EGFR we found many well known interactors like GRB2, PLC-gamma, CBLC, PIK3R1(p85), MIG6, VAV3 and STAT3 among others. We identified the sites for three new protein interactions with this receptor (PIK3R3, FRK and RASA1; the current literature supports a general role for these genes in EGFR signaling) and confirmed their significance using classical pull-downs with phosphorylated EGFR from cell lysates. We also examined the serine phosphorylation sites of other proteins. We identified the two F-box proteins from our library, beta-TRCP 1 and 2 (BTRC and FBXW1B), for example, proteins known to interact with phosphorylated Ikappa B alpha (pS32,pS36). Similarly, we also found a series of the well-known interacting partners for RAF1 (pS621) and the 14-3-3 series of proteins (YWHAZ, Ywhah, YWHAG, YWHAB, Ywhab and YWHAE). In the case of the serine/threonine phosphorylation sites we were able to distinguish proteins that interact with MYC (pT58pS62) and CREB (p133) specifically. Once again these interactions were confirmed in part with initial pull down studies and transfection based reporter assays that suggested real biology. We hope to fully validate these interactions and complete sequencing to confirm their identity soon.

⓭ Conclusions

The overall success of this combination of a highly curated library of proteins with the Multi-Array technology can best be judged by the ability of our methodology to pick out well-known pairs of interactions from a collection of proteins in a blinded fashion. This blinded "discovery" was demonstrated most convincingly by the recovery of known interactions like the EGFR binding partners for GRB2, PLCG2, CBLC, STAT3, SHC1, MIG6 and others; the RAF1 binding interactions to YWHAB family of proteins (beta, gamma, epsilon, zeta and eta); and the NFKBIA binding to BTRC and FBXW1B. Our ability to reuse the labeled proteins from one binding step increased the speed of the screen and improved its economics, effectively creating screens through large numbers of different peptides for many possible interactions. In this example, we used 14 different peptides and made more than 200,000 determinations on 14000 different proteins using a process that was readily scalable beyond even this level. We are currently developing arrays at higher densities that should further increase the scope of our surveys. The combination of the protein library and the speed of the Multi-Array technology took the primary and secondary screens from idea to execution rapidly with comparatively modest effort. Our screens uncovered a significant number of new and significant interactions that we continue to validate.

Meso Scale Discovery™
A division of Meso Scale Diagnostics, LLC.