

MINIREVIEW

Proteomic strategies for the discovery of novel diagnostic and therapeutic targets for infectious diseases

Moushimi Amaya¹, Alan Baer¹, Kelsey Voss¹, Catherine Campbell¹, Claudius Mueller², Charles Bailey¹, Kylene Kehn-Hall¹, Emanuel Petricoin III² & Aarthi Narayanan¹

¹ National Center for Biodefense and Infectious Diseases, George Mason University, Manassas, VA, USA

² Center for Applied Proteomics and Personalized Medicine, George Mason University, Manassas, VA, USA

Proteomic approaches can be a useful means of identifying targets for identifying and treating a variety of infectious agents. This is especially true for agents responsible for emerging infections or those important to biodefense.

Keywords

reverse-phase protein microarray; mass spectrometry; protein–protein interactions.

Correspondence

Aarthi Narayanan, National Center for Biodefense and Infectious Diseases, George Mason University, Biomedical Research Lab 10650 Pyramid Place, MS 1J5 Manassas, VA 20110, USA.

Tel.: +1 703 993 9610

fax: +1 703 993 4280

e-mail: anaraya1@gmu.edu

Received 4 October 2013; revised 18 January 2014; accepted 23 January 2014. Final version published online 13 March 2014.

doi:10.1111/2049-632X.12150

Editor: Gerald Byrne

Abstract

Viruses have developed numerous and elegant strategies to manipulate the host cell machinery to establish a productive infectious cycle. The interaction of viral proteins with host proteins plays an important role in infection and pathogenesis, often bypassing traditional host defenses such as the interferon response and apoptosis. Host–viral protein interactions can be studied using a variety of proteomic approaches ranging from genetic and biochemical to large-scale high-throughput technologies. Protein interactions between host and viral proteins are greatly influenced by host signal transduction pathways. In this review, we will focus on comparing proteomic information obtained through differing technologies and how their integration can be used to determine the functional aspect of the host response to infection. We will briefly review and evaluate techniques employed to elucidate viral–host interactions with a primary focus on Protein Microarrays (PMA) and Mass Spectrometry (MS) as potential tools in the discovery of novel therapeutic targets. As many potential molecular markers and targets are proteins, proteomic profiling is expected to yield both clearer and more direct answers to functional and pharmacologic questions.

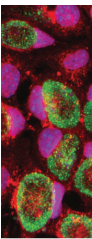
Introduction

Viruses have evolved an arsenal of mechanisms for infecting their host and to establish a productive infectious cycle. Hosts in turn have developed innate cellular defenses to block infection and replication of these viruses. It has been well established that the introduction of a virus into a host cell alters the host's proteome resulting in novel protein interactions. The novel interactions may not only occur between integral host components, but may also involve viral proteins. A comprehensive understanding of the molecular mechanisms underlying these protein–protein interactions (PPIs) would therefore enable the utilization of these mechanisms for the discovery of antiviral drugs and the identification of virally altered host proteins (Zheng *et al.*, 2011; Kshirsagar *et al.*, 2013; Mancone *et al.*, 2013).

“Omics” technologies have advanced tremendously over the last few years, with genomics and transcriptomics

pushing the gene expression field ahead (Braun and Gingras 2012; Mancone *et al.*, 2013; Noisakran *et al.*, 2008; Keating *et al.*, 2013). However, gene expression profile datasets are limited in that they do not necessarily correlate with steady state protein products and are often not truly representative of the interactome (Graves & Haystead, 2002; Noisakran *et al.*, 2008; Mancone *et al.*, 2013). For the purposes of this review, we consider host proteins that interact with one another and with viral proteins forming distinct complexes as an interactome. Furthermore, genomics and transcriptomics studies alone cannot elucidate PPIs on a structural and functional level (Mancone *et al.*, 2013).

PPIs can be studied using a variety of proteomic approaches ranging from biochemical and genetic to large-scale high-throughput technologies (Zheng *et al.*, 2011; Kshirsagar *et al.*, 2013). Here, we will review and evaluate techniques employed to elucidate PPIs with a



primary focus on Protein Microarrays (PMA) and MS as potential tools in the discovery of novel therapeutic strategies.

Traditional approaches

There are multiple traditional molecular biological and biochemical approaches that have been used for many decades to successfully detect PPIs. However, discussion of all of the methods is out of the scope of this review. In this review, we will include contextual information on the most popular classical methods employed for analyzing PPIs which include yeast two-hybrid systems (Y2H) and co-immunoprecipitations (co-IP) (Drewes & Bouwmeester, 2003; Miernyk & Thelen, 2008; White & Howley, 2013). Together, these two methods were responsible for deciphering over 70% of interactions in the IntAct database as of May 2008 (Aranda *et al.*, 2010; Chen *et al.*, 2010).

The Y2H system is popular for the detection of binary physical interactions (Drewes & Bouwmeester, 2003; Hsu & Spindler, 2012; White & Howley, 2013). This method relies on the interaction between a 'bait' fusion protein and a 'prey' fusion protein to activate reporter genes or selectable markers. The advantages to this system are that it is economical and can be used to process large sample sizes (Drewes & Bouwmeester, 2003; White & Howley, 2013). For example, a pooled library screen approach can be used where a library of known 'prey' clones are combined and tested as pools against 'bait' strains. An example of determining PPI by the Y2H system was demonstrated by the influenza virus NS1 protein and human Staufen proteins *in vivo* (Falcón *et al.*, 1999; Cho *et al.*, 2013), where coupling genome wide expression profiling with the Y2H system suggested a large number of host-viral protein interactions (Shapira *et al.*, 2009; Tafforeau *et al.*, 2011). Additionally, this system was used to identify: nine putative host cell proteins interacting with NSm of Rift Valley Fever Virus (RVFV); Vaccinia virus interactions with a variety of proteins (McCraith *et al.*, 2000); Severe Acute Respiratory Syndrome-CoronaVirus (SARS-CoV) proteins self-interacting and/or interacting with other viral proteins to form multimeric complexes (Von Brunn *et al.*, 2007); and numerous cellular proteins binding to Dengue virus (DENV) proteins (Khadka *et al.*, 2011). Unfortunately this procedure is also limited in that interactions between more than two proteins cannot be detected (Drewes & Bouwmeester, 2003; White & Howley, 2013) and the PPIs are observed outside of their natural context and will always require independent validation in relevant model systems.

The analysis of PPIs traditionally involves binding systems, where the protein of interest is tagged for recovery in methods such as co-immunoprecipitations (co-IP) (Miernyk & Thelen, 2008; Noisakran *et al.*, 2008). Co-IPs rely on the idea that a fairly strong interaction between two proteins is stable under stringent elution conditions (such as high salt). When one partner of such an interaction is precipitated, the interacting protein is also precipitated and can be detected by antibody-based methods. Co-IP usually uses a sepharose-protein A column. Briefly, an antibody specific to the

protein of interest is incubated with prepared protein samples to form an antibody-protein complex, which is then analyzed by SDS-PAGE and Western blotting, or further processed for more modern approaches including Matrix Assisted Laser Desorption/Ionization – Time of Flight Mass Spectrometry (MALDI-TOF MS) (Drewes & Bouwmeester, 2003; Miernyk & Thelen, 2008; Zheng *et al.*, 2011). As examples, the use of co-IP has allowed for the identification of host-viral PPIs such as: DENV nonstructural protein 1 and human heterogeneous nuclear ribonucleoprotein C1/C2 (Noisakran *et al.*, 2008); the methyltransferase domain of West Nile Virus NS5 protein and mammalian cGMP-dependent protein kinase, protein kinase G (Keating *et al.*, 2013); and Vaccinia virus p37 with the host Rab9 and tail interacting protein of 47 kDa (TIP47) (Chen *et al.*, 2009). Additional examples of co-IP studies include the Venezuelan Equine Encephalitis Virus, nonstructural protein 2 with the host's major ribosomal phosphoprotein RpS6 (Montgomery *et al.*, 2006) and Ebola virus VP35 protein with the host inhibitor of κ B kinase ϵ (IKK ϵ) and TANK-binding kinase 1 (TBK-1) (Prins *et al.*, 2009). While valuable information has been determined through co-IP studies, the major drawbacks to this technique are that it may not be an accurate representation of an *in vivo* scenario, and weak or transient PPIs may be overlooked (Hsu & Spindler, 2012).

Additional methods are emerging to analyze PPIs that utilize a new generation of aptamers that contain chemically modified nucleotides. Aptamers are short single-stranded oligonucleotides that bind with high affinity and specificity to proteins, peptides, and small molecules (Gold, 1995; Brody & Gold, 2000). One method uses Systematic Evolution of Ligands by EXponential enrichment (SELEX) to select aptamers from libraries of randomized sequences (Ellington & Szostak, 1990; Gold, 1995). Although this technology has high potential for high-throughput biomarker identification, there have been some difficulties creating high affinity aptamers for some protein targets (Gold, 1995). In addressing this problem, a new form of aptamers has been developed, called Slow Off-rate Modified Adapters (SOMAmers). The basis for SOMAmers is that the addition of functional groups can give aptamers protein-like properties that enable a wider variety of high affinity aptamers (Gold *et al.*, 2010). The aptamer-based technology wave has advantages, but requires further characterization, validation and standardization with standards such as affinity capture-MS (Gold *et al.*, 2010).

Due to the limitations of traditional methods, improved approaches for assessing PPIs are in demand, which can not only increase our capacity to analyze broad ranges of interactions in a systems biology approach, but also function synergistically with traditional methods. PPIs cannot be observed in isolation as binary interactions, which strengthen the need for new approaches that would encompass both the ability to identify PPIs but also relevant biological pathway information. More recently, large-scale 'precision proteomics' based on MS and microarrays have enabled the system-wide characterization of host-based events at the levels of post-translational modifications, PPIs and changes in protein expression. This advancement

delivers accurate and unbiased quantitative information regarding protein modifications in response to any perturbation. As large-scale proteomics-based signaling research continues to develop and integrate existing and novel technologies alongside improved databases, it is likely that our understanding of signaling networks will undergo significant change in the coming years.

Protein microarrays

PMA have been used to profile PPIs leading to the discovery of possible drug targets at the molecular level (Wulfkuhle *et al.*, 2004; Gulmann *et al.*, 2006; Wilson *et al.*, 2010; Mancone *et al.*, 2013). The basis of a protein microarray is an array of immobilized protein spots, arranged on a slide so that each spot contains either a homogenous or heterogeneous set of bait molecules (Liotta *et al.*, 2003). Currently there are two popular types of PMAs: traditional or forward phase microarrays (FPMA) and reverse phase protein microarrays (RPMA; Liotta *et al.*, 2003; Sheehan *et al.*, 2005; Wilson *et al.*, 2010; Zhu *et al.*, 2012; Sutandy *et al.*, 2013). In FPMA, bait molecules, usually antibodies are immobilized on a glass slide and a cell lysate sample (the antigens) is placed on the array. A layer of signal-generating antibodies is then added to allow for binding. A signal is emitted only at positive spots on the array (Fig. 1a). In contrast, RPMA methodology immobilizes the test sample analytes (the antigens) on the slide, and then antibodies are applied as the second mobile phase. The RPMA format immobilizes a different test sample in each spot, meaning that one array can be comprised of hundreds to thousands of different samples. The array is incubated with a signal-generating antibody to obtain positive signals (Fig. 1b). Each of these arrays will be described further below.

Forward phase protein microarrays

FPMA is a protein microarray that is designed by fixing and immobilizing multiple bait molecules on a glass slide surface, where the baits are usually antibodies (Fig. 1a), but can also be nucleic acids, peptides and phage lysates (Liotta *et al.*, 2003; Wulfkuhle *et al.*, 2004; Sheehan *et al.*, 2005; Gulmann *et al.*, 2006; Wilson *et al.*, 2010; Hu *et al.*, 2011). FPMA allows the simultaneous measurement of multiple proteins from a single sample (Liotta *et al.*, 2003; Sheehan *et al.*, 2005; Gulmann *et al.*, 2006; Wilson *et al.*, 2010; Hu *et al.*, 2011). The analytes from either serum or tissue samples are pre-labeled, often with fluorophores such as Cy3 or Cy5, is placed over the array. A layer of signal-generating antibodies is added to the analytes to allow binding such that a signal will be emitted only at positive spots on the array (Liotta *et al.*, 2003; Gulmann *et al.*, 2006; Wilson *et al.*, 2010; Hu *et al.*, 2011).

Studies using FPMA have been used in the study of a wide range of topics such as, determination of proteomic profile of endogenous proteins in LoVo colon carcinoma cells exposed to ionizing radiation to identify regulatory sites for radiation-induced apoptosis signaling (Sreekumar *et al.*, 2001). Additional applications of FPMA involved profiling protein tyrosine phosphorylation and characterizing changes in post-translational modifications, such as acetylation and ubiquitination in mammalian cells (Ivanov *et al.*, 2004). Other studies have developed a proteomic profile of cancer progression in oral cavity cancer tissue samples where a combination of FPMA was employed with Laser Capture Microdissection (LCM) (Knezevic *et al.*, 2001). LCM enables researchers to isolate cells of interest from heterogeneous tissue (with multiple kinds of cells) and arrive at a homogeneous cell population without contamination from surrounding nonrelevant cells (Espina *et al.*, 2006). For

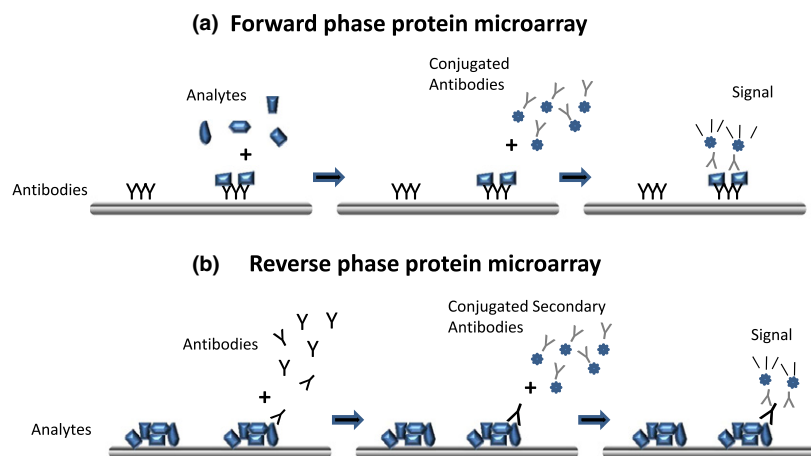


Fig. 1 Comparisons between Forward and Reverse Protein Microarrays. (a) The forward phase microarray format is based on immobilization of an analyte capture reagent, such as an antibody, onto a solid support which is then exposed to analytes. Immobilized analytes are then probed with a suitably conjugated antibody for visualization, utilizing a sandwich assay-based approach that requires two well-performing antibodies. (b) RPMA is characterized by immobilization of analytes onto the substrate, allowing direct comparison of hundreds of samples side by side, and requiring only one well-performing antibody, increasing chances of epitope recognition while decreasing the likelihood of non-specific interactions.

additional technical information on LCM, the reader is directed to Espina *et al.* (2006). The data obtained from these studies illustrate just a few examples of applying FPMA as a high-throughput tool to profile a variety of post-translational modifications in cells under different treatments.

The two major limitations for this procedure are as follows: fluorophore labels can potentially disrupt binding capabilities of the analyte, and labeling all of the proteins in the sample can result in a high background due to non-specific binding. This is because analytes in a complex mixture may be present in low abundance, alongside an excess of other proteins. Performing serial dilutions of the samples is often necessary to dilute protein concentrations to alleviate high background issues (Gulmann *et al.*, 2006). An additional limitation is that FPMA requires specific antibodies to targeted proteins, with differences in antibody design further introducing potential variability when comparing independent experiments (Gulmann *et al.*, 2006).

Reverse phase protein microarrays

RPMA is a protein microarray that immobilizes the test sample analytes on the slide, and then antibodies are applied on top of those analytes (Fig. 1b, compared to FPMA in Fig. 1a). RPMA is a newer iteration of traditional FPMA technology developed in 2001 for the cancer field by Drs. Lance Liotta and Emanuel Petricoin III, where cell populations from samples taken from cancer tissue were subjected to LCM and used to analyze the state of pro-survival checkpoints and growth regulation proteins (Wulfschuhle *et al.*, 2004; Pawletz *et al.* 2001). Since this publication, RPMA technology has been adopted by many research groups and applied not only to LCM tissue, but also to heterogenous tissue samples, cell culture lysates (Nishizuka *et al.*, 2003), and serum/plasma samples (Janzi *et al.*, 2005; Mueller *et al.*, 2010), ovarian effusions (Davidson *et al.*, 2006), vitreous (Davuluri *et al.*, 2009), fine needle aspirates (Rapkiewicz *et al.*, 2007), and peptides (MacBeath & Schreiber, 2000). Moreover, the use of RPMA has not been limited to pre-clinical research studies, but has been a critical component of several clinical trials [for review see Mueller *et al.* (2010)], cancers (VanMeter *et al.*, 2008; Wilson *et al.*, 2010; Einspahr *et al.*, 2012), cellular pathway characterizations, bacterial infection mechanisms (Popova *et al.*, 2009), immunological disorders and viral–host interactions (Wilson *et al.*, 2010; Einspahr *et al.*, 2012; Aguilar-Mahecha *et al.*, 2009; Narayanan *et al.*, 2012; Baer *et al.*, 2012; Austin *et al.*, 2012; Popova *et al.*, 2010). RPMA is capable of monitoring protein dynamics as a function of time, in diseased vs. nondiseased states before, during, and after treatments (Sheehan *et al.*, 2005; Spurrier *et al.*, 2008) and to quantitatively monitor protein expression levels of many samples simultaneously (Liotta *et al.*, 2003; Wulfschuhle *et al.*, 2004; Sheehan *et al.*, 2005; Gulmann *et al.*, 2006; Hultschig *et al.*, 2006; Spurrier *et al.*, 2008; Wilson *et al.*, 2010; Hu *et al.*, 2011; Sutandy *et al.*, 2013).

RPMA technology provides the necessary analytical sensitivity to measure even ultra-low abundant proteins.

For this reason, the combination of LCM technology with RPMA creates a powerful analytical tool based on its increased sensitivity, reduced sample volume/concentration input requirements and combination of hundreds to thousands of samples on a single array. Moreover, because the RPMA is not a two-site sandwich-type antibody array and requires only one well-performing primary antibody for detection, the number of analytes that can be measured is much greater than what could be performed on a forward-phase array. Typically, with the RPMA nanoliters (nl) of each sample are printed in duplicate or triplicate onto glass-backed nitrocellulose slides using either contact (Sheehan *et al.*, 2005; Gulmann *et al.*, 2006; Löbke *et al.*, 2008) or non-contact (Schena, 2000) arrays. The source of these analytes could be *in vitro* samples generated from cultured cells or *in vivo* samples generated from animal tissue or clinical samples. With a spot size of only tens to a few hundred microns in diameter, the RPMA platform allows for thousands of samples to fit on one slide and hundreds of slides can be analyzed in every array run giving this approach a high throughput flavor. Each array is then probed with a single primary antibody, in principle similar to other immunoassays.

The biggest challenge for RPMA is the same as for any immunoassay: the need and availability of high quality, specific antibodies. Prior to application to RPMA, antibodies have to be validated using Western blotting to demonstrate high target specificity. Currently, the development of validated antibody libraries is an individual effort by each lab. In George Mason University, we currently possess a large repertoire of more than 400 validated antibodies relevant to phosphorylated and unphosphorylated proteins which map to diverse nodes in many phospho-signaling cascades. It will be beneficial for the protein microarray field to combine such efforts in the future and assemble and maintain a central repository of validated and “RPMA-certified” antibodies.

Diagnostics and current applications of RPMA

To describe the workflow for RPMA (Fig. 2), we will follow an example of a RPMA study conducted by Popova *et al.* (2010). In this study, Popova *et al.* used RPMA to identify phospho-signaling protein pathways modulated in cultured human cells at very early (of the order of a few minutes) to late time points after RVFV infection. Human cells were infected with the wildtype strain of RVFV, ZH-501 and cells were lysed at various time points. For the RPMA procedure, c. 30 nL of each sample (equivalent to the amount of material from 40 lysed cells) was arrayed on nitrocellulose slides. An advantage of the array layout is that each sample is printed in a 4–8 point dilution series which ensures that the quantification of positive signals can be performed in the linear range. Each slide was then probed with one of 60 different antibodies specific against phosphorylated or total forms of signaling proteins. The antibodies were selected to monitor the molecular networks involved in host responses most likely affected by virus exposure such as apoptotic and cell survival pathways. To validate the RPMA data, the levels of multiple selected proteins were analyzed by

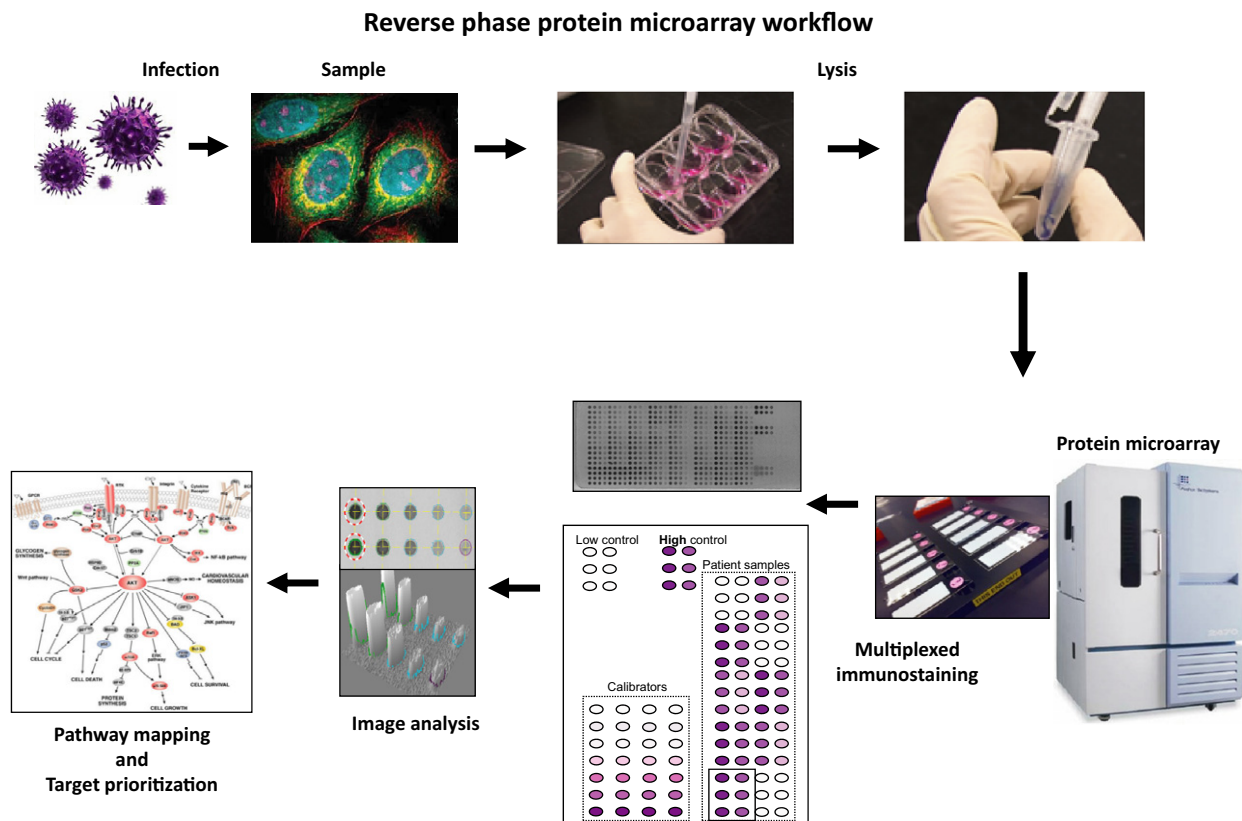


Fig. 2 Protein Microarrays. Representative RPMA workflow schematic. The source of analytes may be infected cells as shown here or *in vivo* samples. The lysed samples are then arrayed on a nitrocellulose glass slide in a multiplexed manner. This allows for many hundreds of slides to be imprinted with sample at the same time. The size of the pins that imprint the samples on the slides determine how many hundreds to thousands of analytes can be printed on every slide. Each slide also is imprinted with positive controls – known analytes of predetermined concentration (high and low controls shown on the slide). Finally, each slide also contains calibrator spots in a dilution series. The high and low control spots and the calibrators not only permit quantitative interpretation of data within a slide, but permit comparisons between slides and between multiple experiments. Each slide in the array is queried with a single predetermined antibody. The total number of slides in each experiment is determined by the total number of antibodies (in other words, total number of desired targets). Following an antigen : antibody interaction, the slides are stained and the intensities of the spots on each slide are quantified. Relative differences in signal intensities between biologically distinct analytes can then be plotted in a graphical format.

Western blots using antibodies against phosphorylated and total forms of proteins; both procedures were in agreement with each other. This demonstrated that while RPMA has enormous technical utility independently, it can also function in a synergistic capacity with traditional methods.

Following imaging of arrays, the software technology used to capture and quantify the analyte spots is similar to software used for DNA microarray analysis [i.e. IMAGEQUANT (GE Healthcare Life Sciences, Pittsburgh, PA) or MICROVIGENE™ (VigeneTech, North Billerica, MA)]. Some software packages, such as MICROVIGENE, include several features to facilitate quantification of array data: automated spot-boundary detection with image contrast enhancement, dust/scratch removal, outlier flagging, regional background correction and total protein, negative controls and internal reference standard normalization options (Gulmann *et al.*, 2006). Data normalization algorithms range from calibration curve normalization (Sevecka & MacBeath, 2006), spiked-in internal standard normalization (Korf *et al.*, 2008), Robust-Linear-Model normalization (Sboner *et al.*, 2009),

to normalization of respective sample total protein. In the example of Popova *et al.* (2010), the average total level of cellular protein at every time point was determined by staining with Sypro Ruby Protein Blot Stain of four randomly selected slides throughout the print run (Popova *et al.*, 2010). Spot finding software programs convert pixel density for each spot into numerical values (Mueller *et al.*, 2010). The MICROVIGENE software offers automated curve-fitting approaches to quantify the analyte across a five-point dilution curve, including replicates (Gulmann *et al.*, 2006). For additional technical information about the software the reader is directed to Gulmann *et al.* (2006).

The downstream biostatistical analysis of the raw, normalized array data depends highly on the type of samples printed on the array and the scientific questions that are being addressed. In general, many of the statistical methods employed in RPMA data analysis are traditional methods used to evaluate confidence intervals. Methods such as unsupervised hierarchical clustering, k-means clustering, self-organizing maps and principal component analysis are

tools used to determine a relationship between the data points generated from the microarray (Gulmann *et al.*, 2006). Clustering of closely related data points can be determined with a cut-off point that makes the data biologically sensible; however, statistically significant differences between data points cannot be provided with hierarchical clustering methods (Gulmann *et al.*, 2006). Other common statistical methods such as, ANOVA, *t*-tests and Mann–Whitney *U*-test can be employed to determine differences between data sets provided prior knowledge of the classification of the samples is known (Gulmann *et al.*, 2006).

Several technologies exist to graphically map protein signaling pathway activation data from RPMA. One approach is to overlay the RPMA data over a static image of PPIs that are generally accepted by the larger scientific community, such as the well-known “Pathways in Human Cancer” cell signaling map by Dr. Robert Weinberg. CScape uses the Google Maps API to overlay this cell signaling map with relative protein/phosphoprotein abundances, enabling the visual identification of “hot spots” within the cell signaling architecture or whole activated/nonactivated signaling pathways (Fig. 3). Another approach is to visualize protein abundance and protein–protein correlations (i.e. as calculated by Spearman’s Rho analysis) in a Bayesian network like computational analysis using general network analysis systems such as Gephi (Bastian *et al.*, 2009). The graphs generated with this application graphically illustrate influences of pathway components on each other in the form of nodes and arcs, where nodes represent the variables and the arcs statistically significant relationships between the variables.

One of the strongest advantages of RPMA is its ability to generate multiple hypotheses for subsequent testing and development of the field. The outcome of each RPMA run is a virtual treasure trove of information on multiple signaling events analyzed as a single snapshot in time. Each one of those altered signal events will form a basis of an individual hypothesis. If this involved specific viral strains or individual

viral proteins, for example, we can go to great lengths in deciphering functional consequences of PPIs. This was exemplified by a RPMA study conducted by Popova *et al.* (2012) that revealed phosphorylation and activation of the NF- κ B signaling cascade. Narayanan *et al.* (2012) demonstrated in a follow up study that phosphorylation of p65 (serine 536) involved phosphorylation of I κ B α through the classical NF κ B cascade and that RVFV utilized the host NF- κ B signaling cascade to establish a robust infection. In addition, this group demonstrated that inhibition of the NF- κ B cascade was able to inhibit viral replication (Narayanan *et al.*, 2012). The more important outcome of that study was that the viral protein NSs may have an influence on the activation status of the cascade. The same RPMA study also revealed that the tumor suppressor protein, p53 was phosphorylated in the event of ZH501 infection (Austin *et al.*, 2012). A study by Austin *et al.* (2012) expanded on this data to show that in RVFV infections, p53 was activated, phosphorylated and localized in the nucleus following transient interaction with the RVFV virulence factor NSs. Further characterization of p53 phosphorylation prompted investigation into the DNA damage signaling pathway during RVFV infection. Baer *et al.* (2012) demonstrated that RVFV infection induced an NSs-dependent DNA damage signaling response and concurrent S-phase arrest, providing evidence for a novel function of NSs that was shown to directly impact viral replication. Collectively, modifications in phospho-protein signaling in RVFV examined by a single RPMA study resulted in the elucidation of numerous cellular mechanisms that played important roles during viral infection and that were subsequently exploited in targeting viral replication. The results drawn from these studies illustrate how the high throughput identification of PPIs between the host and the virus can lead to the identification of potential therapeutic targets during viral infection (Baer *et al.*, 2012).

In 2009 Popova *et al.*, utilized RPMA to characterize and measure the innate cell signaling responses of lung epithelial cells to *Bacillus anthracis* infection (Popova *et al.*, 2009). The authors demonstrated RPMA that *B. anthracis* infection

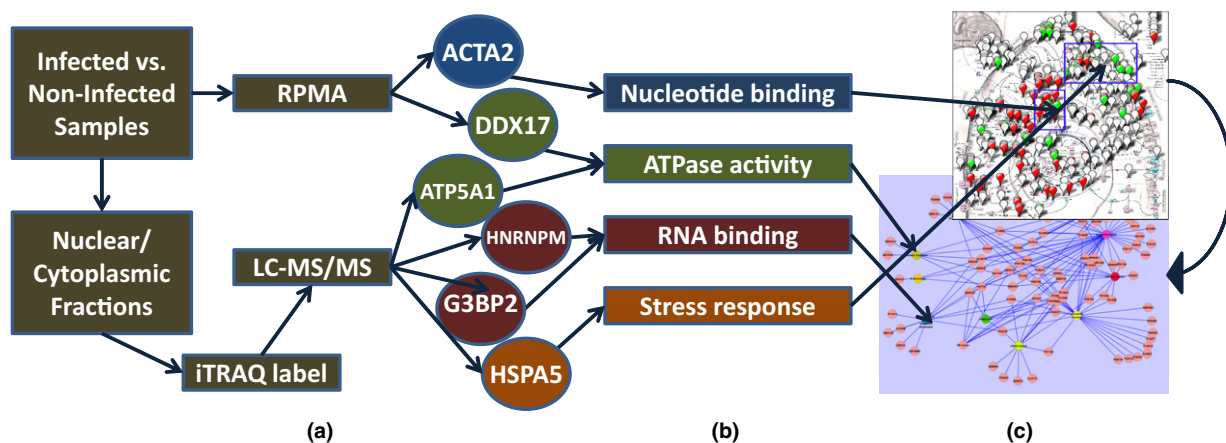


Fig. 3 Development of Protein Interaction Networks. Differentially expressed proteins identified by proteomics experiments (a, which) are then converted into commonly shared functional groupings (b) and used to develop network visualizations using tools such as CScape (c. top) and Cytoscape (c. bottom). These networks can be further analyzed and cross-compared to identify key pathways and host–pathogen interactions.

inhibited MAPK and PI3K/AKT signaling pathways and characterized host signaling from nonlethal and lethal strains of *B. anthracis* (Popova *et al.*, 2009). By virtue of the comparative nature of RPMA, this study opened avenues for detecting previously unrecognized host responses silenced upon infection as part of the pathogenic process (Popova *et al.*, 2009).

Another example of diagnostic and current applications of RPMA technology involves an effort to create a rapid, accurate and sensitive diagnostic tool to detect biomarkers for SARS-CoV. To address this problem, Zhu *et al.* (2006) developed the first coronavirus protein RPMA. The entire proteome of the human SARS-CoV, HCoV-229E virus and partial proteomes of other coronaviruses were imprinted on the microarray and probed with symptomatic patient sera (Zhu *et al.*, 2006). The presence of viral-specific antibodies was detected using Cy3-labeled goat anti-human IgG antibodies (Zhu *et al.*, 2006). Analysis of the microarray revealed increased sensitivity, reactivity and accuracy for biomarker detection in SARS infected individuals (Zhu *et al.*, 2006). This example demonstrates the feasibility of using RPMA technology in a clinical environment to detect biomarkers.

Mass spectrometry

While the advent of novel and high throughput techniques such as RPMA have greatly expanded on the ability to generate informative datasets, complementary approaches are necessary to account for the dependence on antibodies. Additionally, antibodies have to be selected by the researcher and are inherently directed to a small number of known signaling events and modifications, giving rise for the need to complement antibody validation using a technique that is independent of these limitations. To get around the inherent limitations and biases of antibody-based research, modern quantitative MS is increasingly being used to balance and direct antibody-based approaches (De Chasse *et al.*, 2012; Zhou *et al.*, 2011; Patwa *et al.*, 2010; Patwa *et al.*, 2009). MS-based proteomics is not limited to specific sites or proteins of interest which makes it suitable for use in an unbiased (hypothesis free) and systems-wide manner, representing a fundamentally different approach to studying cell signaling (Go *et al.*, 2006; Zheng *et al.*, 2011). One commonly used MS approach depends on peptide separation normally performed by liquid chromatography (LC) tandem MS or LC-MS-MS (Gonzalez-Galarza *et al.*, 2012).

When looking at host cell signaling responses, complex protein samples may be derived from whole cells or from biological fluids, and while all of the proteins in an entire sample may be processed as is, it is often preferable to perform separation techniques to better characterize information on proteins of particular biological interest (Brewis & Brennan, 2010). Protein separation is initially performed prior to this step, by methods such as one dimensional (1D) or two dimensional (2D) gel electrophoresis, column chromatography, or affinity purification, followed by enzymatic digestion (often using trypsin). Sample preparation is critical

to subsequent data interpretation, particularly when dealing with large datasets such as those generated by LC-MS. Techniques such as subcellular fractionation can provide cleaner samples while also giving valuable protein localization information compared with that from whole cell lysates alone (Brewis & Brennan, 2010; Gonzalez-Galarza *et al.*, 2012). With biological fluids, it is also possible to remove the most abundant proteins by immunodepletion or enrichment.

An alternative or complimentary method for global protein separation is to resolve proteins by 1D or 2D electrophoresis before subjecting individual protein bands to digestion and LC-MS (Brewis & Brennan, 2010). In both methods, many gel segment/protein band cuts can be processed to identify numerous separated proteins. In a typical workflow, an individually separated protein is physically isolated and removed, trypsin digested and the resulting peptides are separated on the basis of relative hydrophobicity through LC before tandem MS (MS/MS Brewis & Brennan, 2010). The tandem MS data are then used to search existing protein databases to achieve a matching protein spectra based on amino acid sequences typically derived from MS/MS spectra. Low abundance, very large, or very small proteins have, however, proved difficult to resolve using 2D gels, and for global analysis it is now much more commonplace to trypsin digest the entire solubilized protein mixture to produce a peptide "soup" of all the proteins in the sample (gel-free LC-MS proteomics) (Petricoin *et al.*, 2002). Peptides can then be separated by LC on the basis of relative hydrophobicity and charge as a multidimensional separation (Petricoin *et al.*, 2002). Extensive MS/MS and database searches can then be performed to identify many of the proteins in the original sample.

One of the advantages of this protein interactome mapping workflow is that it is possible to achieve quantitative data at the same time by introducing a peptide labeling step, such as isobaric Tags for Relative and Absolute Quantification (iTRAQ) method, using labeled peptides as internal standards, or to perform protein labeling, such as stable isotope labeling by amino acids in cell culture (SILAC), without the need for additional LC-MS (Brewis & Brennan, 2010). iTRAQ represents a method for differentiating between multiple conditions (up to eight samples) simultaneously (Wiese *et al.*, 2007; Brewis & Gadella, 2010). This approach lends itself to the labeling of *in vivo* samples and/or primary cells as such labels are applied postsample preparation. As an example, iTRAQ has been used to investigate influenza virus infection in primary human macrophages (Lietzén *et al.*, 2011). Typically, two sets of samples are differentially labeled with a stable isotope, one sample with the light form, the other with the heavy form of the label. Depending on the type of the chemical labeling reagent used, samples are labeled before, during, or after enzymatic digestion, combined then subjected to a separation/enrichment technique followed by analysis with MS for a quantitative comparison. Changes in protein expression levels are then quantified by determining the ratio of the peak intensities of the light and heavy forms of the generated peptides (Brewis & Brennan, 2010). The pairwise comparison of the peak intensities of peptides

labeled with the heavy and light form of the label serves as the basis for quantitative protein analysis (Brewis & Brennan, 2010). It should be noted that the accuracy of the quantitative measurement depends on when the label was incorporated, with the highest level of accuracy being achieved when the label is incorporated early on in the process due to sample loss and processing variability.

SILAC has also been applied to study viral infections in cell culture to provide quantitative information regarding pathogen–host cell interactions (Brewis & Gadella, 2010; Zhou *et al.*, 2011; Toss *et al.*, 2013). During SILAC, cells and viruses are differentiated by growing each cell population (or the virus within a cell population) in media containing unique stable isotope labeled amino acids that become incorporated into newly synthesized proteins, eventually supplanting their equivalent nonlabeled homologs (Brewis & Gadella, 2010). This labeling technique greatly reduces sample complexity and increases labeling efficiency prior to sample preparation, while allowing the relative quantification of proteins in samples by MS (Brewis & Gadella, 2010). When analyzed on a mass spectrometer, pairs of chemically identical peptides of different isotope composition can then be differentiated, as the labeled amino acids induce a shift in the mass/charge ratio (m/z) in comparison to the unlabeled peptides (Brewis & Brennan, 2010; Brewis & Gadella, 2010). By comparing the intensities of the labeled and unlabeled m/z peaks, it is possible to obtain accurate quantitative data on the relative abundance of labeled and unlabeled peptides present in the sample (Brewis & Brennan, 2010; Brewis & Gadella, 2010). In this way, proteins that are increased or decreased in abundance in virus-infected, compared with mock or drug-treated cells, can be simultaneously identified and quantified (Brewis & Brennan, 2010; Brewis & Gadella, 2010).

The relative quantification of peptides using iTRAQ or SILAC labeling, coupled to LC-MS/MS and bioinformatic analysis, is one of the most popular and powerful options for global proteomic quantification and has provided an excellent resource for studying host cell proteomes and is readily being applied in the study of host–pathogen infections (Wiese *et al.*, 2007; Brewis & Brennan, 2010; Brewis & Gadella, 2010; Munday *et al.*, 2012).

Biomarker identification

In the proteome biomarker discovery pathway, the aim is to combine multidimensional fractionation and labeling methods along with MS/MS analysis to identify proteins that are unique or highly abundant in complex samples obtained from specific disease states and comparing those profiles to healthy matched controls. Through the use of peptide labeling such as the previously mentioned iTRAQ or SILAC, samples from both disease-affected and healthy controls samples can be run side by side through an MS workflow with the data obtained then processed using bioinformatic algorithms and programs such as iTRACKER for iTRAQ that search for differences in peak intensities between the sample sets or SILACAnalyzer for SILAC which is an open source tool for the fully automated analysis of quantitative

proteomics data (Petricoin *et al.*, 2002; Wiese *et al.*, 2007; Munday *et al.*, 2012). SILACAnalyzer identifies pairs of isotopic envelopes with fixed m/z separation and requires no prior sequence identification of the peptides. The discriminating pathogenic pattern formed by the key subset of proteins or peptides buried among the entire repertoire of thousands of proteins represented in the sample spectrum, can then be compared to its control group and potentially identified.

A recent study of the Human respiratory syncytial virus (HRSV) using SILAC in conjunction with LC MS/MS allowed the direct and simultaneous identification and quantification of both cellular and viral proteins (Munday *et al.*, 2010). To reduce sample complexity and increase data return on potential protein localization, cells were further fractionated into nuclear and cytoplasmic extracts (Munday *et al.*, 2010). Novel HRSV–host cell interactions, including those associated with the antiviral response and alterations in subnuclear structures such as the nucleolus and ND10 (promyelocytic leukemia bodies) were identified (Munday *et al.*, 2010). In addition, novel changes in mitochondrial proteins and functions, cell cycle regulatory molecules, nuclear pore complex proteins and nucleocytoplasmic trafficking proteins were observed in infected A549 cells (Munday *et al.*, 2010). Commonly available bioinformatics programs such as Ingenuity Pathway Analysis were used in the organization, expansion and interrogation of the derived data sets from virus-infected cells which were then validated using tradition orthogonal assays (Munday *et al.*, 2010; Lietzén *et al.*, 2011). In this study, the use of SILAC in conjunction with LC MS/MS resulted in the identification and potential interaction of 1140 cellular proteins along with six viral proteins (Munday *et al.*, 2010). In another study using Flock House Virus (FHV)-infected *Drosophila* cells, profile changes in the protein expression levels were found based on the direct analysis of intensities using labeling in combination with LC-MS/MS (Go *et al.*, 2006). Overall a total of 1500 host proteins were identified and quantified, of which 150 were up-regulated while 66 were down-regulated in response to viral infection (Go *et al.*, 2006).

While stable isotope labeling has traditionally been used in comparative host–viral proteomics, a label-free approach is becoming a viable and attractive alternative, through the use of chromatographic separation and high mass accuracy measurements along with data normalization methods. Moreover, informatics algorithms are employed to facilitate the data analysis due to the large datasets that can be obtained from LC-MS experiments. Overall, this is highly effective in comparative proteomics due to its comprehensiveness and high throughput nature. To date, label-free quantitative proteomics approaches have been successfully applied in the analysis of numerous samples and organisms such as: human serum, yeast, and *Shewanella oneidensis* (Fang *et al.*, 2006; Brewis & Brennan, 2010). Despite these significant advances and efforts, challenges remain with current label-free methods such as those associated with the dynamic range of measurements and the extent of proteome coverage, confidence in peptide/protein identifications, quantization accuracy, analysis throughput, and the

robustness of present instrumentation are all issues that still need to be addressed and improved on before these technologies can be reliably used in a clinical setting (Qian *et al.*, 2006; Findeisen & Neumaier, 2009).

There are many options available for quantitative, unbiased proteome studies, only some of which have been discussed in this review. New techniques utilizing traditional and novel next-generation targeted proteomics have started to shift the traditional paradigm from using discovery-based identification to targeted quantification, such as in the utilization of triple quadrupole mass spectrometers (QQQ; Boja & Rodriguez, 2012; Shi *et al.*, 2012; Fung *et al.*, 2013). With a discovery-based strategy, the goal is usually to identify as many proteins as possible, while the goal of a targeted proteomics is to monitor a select few proteins of interest with high sensitivity, reproducibility and quantitative accuracy. In the traditional discovery-based approach peptide ions are automatically selected in the mass spectrometer for fragmentation on the basis of their signal intensities, generating rich but complex tandem mass spectra for each peptide sequence, requiring careful analysis and complex bioinformatics. In contrast when using a targeted workflow, the QQQ mass spectrometer can be programmed to detect specific peptide ions derived from proteins of interest, and can select specific 'precursor' ions (on the basis of their *m/z* ratio) for fragmentation. In the second mass filter, target ions are selected and guided to the detector for quantification, resulting in a trace of signal intensity vs. retention time for each precursor ion-product ion pair. This process is called selected reaction monitoring (SRM) or multiple reaction monitoring (MRM; Lange *et al.*, 2008). While this up-front approach is much more labor intensive to develop than a traditional discovery-based pipeline, once a reliable assay is generated for a specific protein, analysis of the MS data is relatively straightforward and uncomplicated in comparison. While SRM is the most mature MS-based technology for targeted proteome analysis, new methodologies that obviate the need for laborious SRM assay optimization are currently being developed. One example of several novel methodologies is SWATH; complex mass spectra generated by data-independent acquisition (in which peptides are selected for fragmentation without regard to signal intensity) are queried for the presence of specific peptides using libraries of qualified peptide fragment spectra (Gillet *et al.*, 2012).

For MS, novel approaches and developments continue to improve and push the field rapidly forward towards targeted approaches and provide many viable options when looking at proteome interactomes from a host-pathogen perspective. While every toolset discussed in this review has its strengths and weakness, combining MS along with other high throughput methods of protein detection such as RPMA, in conjunction with the use of traditional tools and means of validation such as Y2H screens or immunoblot analysis, have demonstrated the ability to greatly strengthen the inherent deficiencies of any individual approach. The use of these orthogonal methods in conjunction is thus able to arrive at a much more well-rounded and comprehensive picture of the proteome and its mechanisms when used in combination (Patwa *et al.*, 2009; Mancone *et al.*, 2013).

Informatics approaches for integrating proteomics data

As previously discussed, cellular function is driven by a complex network of interacting proteins that are controlled by various signaling cascades among other mechanisms. The proteomics experiments described here aim to paint a molecular portrait of the essential proteins activated in specific cellular states, such as comparing infected vs. uninfected states (Fig. 3). Following statistical analysis to identify proteins that are significantly changed between desired analytical states (described above), protein lists are generated that must be transformed into informative networks that are representative of the functional changes observed in the disease state.

There are several key analytical steps that must be undertaken to develop these comprehensive networks, including mapping protein IDs to function, cross-comparing protein function, and identifying, visualizing and analyzing protein interaction networks (Sanz-Pamplona *et al.* 2012). Additionally, protein function, structure and interaction databases are sparse and represent only a small fraction of known human proteins, with host-pathogen interaction databases even more sparse (Kshirsagar *et al.*, 2013; Kshirsagar, Carbonell, and Klein-Seetharaman 2012). Trying to find commonalities among functions is therefore essential for generating the most complete networks possible. Tools such as Ingenuity Pathway Analysis (IPA) (Goulet *et al.* 2013), National Institute of Allergy and Infectious Diseases' Database for Annotation, Visualization and Integrated Discovery (DAVID; Huang, Sherman, and Lempicki 2009), and Cytoscape (Sanz-Pamplona *et al.* 2012) among others can be used to identify common and statistically overrepresented pathways and ontological terms (functions) from protein lists which can then be used to interrogate broader protein interaction networks. These tools draw from large public databases that describe protein functions—such as the Gene Ontology (GO), pathways—such as KEGG, Reactome and Biocarta, and protein interaction databases—such as DIP, BIND, and IntAct among others. These tools can also be used to identify important pathways and interactions between host and pathogen proteins (Kshirsagar *et al.* 2012; Goulet *et al.* 2013).

Although MS experiments are comprehensive in their measurements of cellular proteins, issues described above such as variability in protein abundance, can confound experiments and result in results that cannot reproducibly identify individual proteins. This can confound the types of basic network analyses described in the previous paragraph. Several existing tools can help with assessing the similarities of identified gene functions and merging similar proteins into like functional groups, including SORA and simGIC (Teng *et al.* 2013). These tools can be used to merge protein expression information by protein function across different methods such as RPMA and tandem MS to help cluster proteomics data into functional bins which can then be used to identify pathways and interaction networks. SORA and simGIC, among others, measure the functional similarity among proteins in terms of information content in

the context of GO terms, which allows for more precise binning of proteins with similar functions and roles (Teng *et al.*, 2013).

Networks generated from proteomics experiments are often highly complex and benefit greatly from further analysis with interactive visualization tools that can help to tease out key subnetworks for further analysis or laboratory validation experiments. Tools such as IPA (Goulet *et al.* 2013), Cytoscape (Kacprowski, Doncheva, and Albrecht 2013; Sanz-Pamplona *et al.* 2012) and CScape (Einspahr *et al.*, 2012) allow for insightful visualizations and analysis of these complex cellular networks analysis that allow scientists to directly interact with their data and help to turn information into knowledge of key cellular functions affected by disease. For example, nodes and edges within these networks may be weighted based on common functionality, and/or directionality or magnitude of expression to identify key connections among experimentally identified proteins.

Conclusion

Proteins are currently the major drug targets of choice and play a critical role in the process of modern drug design. Host proteins are particularly attractive targets as pathogens are less likely to develop an obstructive mutation to the therapeutic and develop resistance. Therapeutic development typically involves numerous steps: the construction of drug compounds based on the structure of a specific drug target, validation for therapeutic efficacy of the drug compound, evaluation of drug toxicity, and finally clinical trial. RPMA and MS along with other functional proteomic approaches can be employed at all steps in the investigative process. The use of MS in combination with the use of high throughput antibody platforms will allow a much more global, comprehensive, and directed approach in the study of viral infection or any pathogenic state for that matter, revealing the complexity of the events within an infected cell while compensating for the limitations of a technique in isolation.

Post-translational modifications—including phosphorylation, acetylation, and ubiquitination—are specific modifications that can alter the activity of an individual protein target. The cumulative effect of these small modifications is the regulation of large signaling pathways and networks within cells. A comprehensive understanding of the molecular mechanisms underlying viral infection remains a major challenge in the discovery of new antiviral drugs and host susceptibility factors. New advances in the field are expected to arise from systems-level modeling and the integration of proteomic and genomic disciplines, with current wet lab techniques. Here, we attempted to briefly explore the importance and benefits of using two powerful proteomic techniques in a combinatorial capacity, to better understand the molecular relationships between viruses and their host, chiefly looking at cellular signaling pathways from a global perspective. Past studies have already demonstrated that viral proteomes target a wide range of functional and inter-connected modules of proteins within the human interactome and that using a global and systems-level approach can provide relevant, accurate and valuable

information (Go *et al.*, 2006; Brewis & Gadella, 2010; Munday *et al.*, 2010; Lietzén *et al.*, 2011; Zhou *et al.*, 2011; De Chassesey *et al.*, 2012).

To summarize, protein interaction studies have come a long way in the past decade in terms of technology/platform development and protein chemistry. Together with the use of microarrays and bioinformatics tool sets, along with other “Omics”; the identification of the molecular signatures of diseases based on protein pathways and signaling cascades, holds great promise and utility for disease diagnosis and therapeutic development.

Acknowledgements

The authors would like to thank members of the Narayanan and Kehn-Hall laboratories and Center for Applied Proteomics and Molecular Medicine at GMU for help in manuscript preparation and revision. The authors declare that there are no conflicts of interest relevant to this manuscript.

Authors' contribution

M.A. and A.B. contributed equally to this work.

References

- Aguilar-Mahecha A, Cantin C, O'Connor-McCourt M, Nantel A & Basik M (2009) Development of reverse phase protein microarrays for the validation of clusterin, a mid-abundant blood biomarker. *Proteome Sci* 7: 15.
- Aranda B, Achuthan P, Alam-Faruque Y, *et al.* (2010) The IntAct molecular interaction database in 2010. *Nucleic Acids Res* 38: D525–D531.
- Austin D, Alan B, Lundberg L, *et al.* (2012) p53 Activation following Rift Valley fever virus infection contributes to cell death and viral production. *PLoS ONE* 7: e36327.
- Baer A, Austin D, Narayanan A, Popova T, Kainulainen M, Bailey C, Kashanchi F, Weber F & Kehn-Hall K (2012) Induction of DNA damage signaling upon Rift Valley fever virus infection results in cell cycle arrest and increased viral replication. *J Biol Chem* 287: 7399–7410.
- Bastian M, Heymann S & Jacomy M (2009) Gephi: an open source software for exploring and manipulating networks. *Third International AAAI Conference on Weblogs and Social Media*. <http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154>.
- Boja ES & Rodriguez H (2012) Mass spectrometry-based targeted quantitative proteomics: achieving sensitive and reproducible detection of proteins. *Proteomics* 12: 1093–1110.
- Braun P & Gingras A-C (2012) History of protein-protein interactions: from egg-white to complex networks. *Proteomics* 12: 1478–1498.
- Brewis IA & Brennan P (2010) Proteomics technologies for the global identification and quantification of proteins. *Adv Protein Chem Struct Biol* 80: 1–44.
- Brewis IA & Gadella BM (2010) Sperm surface proteomics: from protein lists to biological function. *Mol Hum Reprod* 16: 68–79.
- Brody EN & Gold L (2000) Aptamers as therapeutic and diagnostic agents. *J Biotechnol* 74: 5–13.
- Chen Y, Honeychurch KM, Yang G, Byrd CM, Harver C, Hruby DE & Jordan R (2009) Vaccinia virus p37 interacts with host proteins associated with le-derived transport vesicle biogenesis. *Virol J* 6: 44.

- Chen YC, Rajagopala SV, Stellberger T & Uetz P (2010) Exhaustive benchmarking of the yeast two-hybrid system. *Nat Methods* 7: 667–668.
- Cho H, Ahn SH, Kim KM & Kim YK (2013) Non-structural protein 1 of influenza viruses inhibits rapid mRNA degradation mediated by double-stranded Rna-binding protein, stau1. *FEBS Lett* 587: 2118–2124.
- Davidson B, Espina V, Seth SM, Flørenes VA, Lance LA, Gunnar KB, Tropé CG, Berner A & Elise KC (2006) Proteomic analysis of malignant ovarian cancer effusions as a tool for biologic and prognostic profiling. *Clin Cancer Res* 12(3 Pt 1): 791–799.
- Davuluri G, Espina V, Petricoin EF 3rd, Ross M, Deng J, Liotta LA & Glaser BM (2009) Activated VEGF receptor shed into the vitreous in eyes with wet AMD: a new class of biomarkers in the vitreous with potential for predicting the treatment timing and monitoring response. *Arch Ophthalmol* 127: 613–621.
- De Chasse B, Meyniel-Schicklin L, Aublin-Gex A, André P & Lotteau V (2012) New horizons for antiviral drug discovery from virus-host protein interaction networks. *Curr Opin Virol* 2: 606–613.
- Drewes G & Bouwmeester T (2003) Global approaches to protein-protein interactions. *Curr Opin Cell Biol* 15: 199–205.
- Einspahr JG, Calvert V, Alberts DS, *et al.* (2012) Functional protein pathway activation mapping of the progression of normal skin to squamous cell carcinoma. *Cancer Prev Res (Phila)* 5: 403–413.
- Ellington AD & Szostak JW (1990) *In vitro* selection of RNA molecules that bind specific ligands. *Nature* 346: 818–822.
- Espina V, Wulfkuhle JD, Calvert VS, VanMeter A, Zhou W, Coukos G, Geho DH, Petricoin EF & Liotta LA (2006) Laser-capture microdissection. *Nat Protoc* 1: 586–603.
- Falcón AM, Fortes P, Marión RM, Beloso A & Ortín J (1999) Interaction of influenza virus NS1 protein and the human homologue of Staufen *in vivo* and *in vitro*. *Nucleic Acids Res* 27: 2241–2247.
- Fang R, Elias DA, Monroe ME, *et al.* (2006) Differential label-free quantitative proteomic analysis of *Shewanella oneidensis* cultured under aerobic and suboxic conditions by accurate mass and time tag approach. *Mol Cell Proteomics* 5: 714–725.
- Findeisen P & Neumaier M (2009) Mass spectrometry based proteomics profiling as diagnostic tool in oncology: current status and future perspective. *Clin Chem Lab Med* 47: 666–684.
- Fung EN, Jemal M & Aubry A (2013) High-resolution MS in regulated bioanalysis: where are we now and where do we go from here? *Bioanalysis* 5: 1277–1284.
- Gillet LC, Navarro P, Tate S, Rost H, Selevsek N, Reiter L, Bonner R & Aebersold R (2012) Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics* 11: doi:10.1074/mcp.O111.016717.
- Go EP, Wikoff WR, Shen Z, *et al.* (2006) Mass spectrometry reveals specific and global molecular transformations during viral infection. *J Proteome Res* 5: 2405–2416.
- Gold L (1995) Oligonucleotides as research, diagnostic, and therapeutic agents. *J Biol Chem* 270: 13581–13584.
- Gold L, Ayers D, Bertino J, *et al.* (2010) Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS ONE* 5: e15004.
- Gonzalez-Galarza FF, Lawless C, Hubbard SJ, Fan J, Bessant C, Hermjakob H & Jones AR (2012) A critical appraisal of techniques, software packages, and standards for quantitative proteomic analysis. *OMICS* 16: 431–442.
- Goulet M-L, Olgner D, Xu Z *et al.* (2013) Systems analysis of a RIG-I agonist inducing broad spectrum inhibition of virus infectivity. *PLoS Pathog* 9: e1003298.
- Graves PR & Haystead TA (2002) Molecular biologist's guide to proteomics. *Microbiol Mol Biol Rev* 66: 39–63; table of contents.
- Gulmann C, Sheehan KM, Kay EW, Liotta LA & Petricoin EF 3rd (2006) Array-based proteomics: mapping of protein circuitries for diagnostics, prognostics, and therapy guidance in cancer. *J Pathol* 208: 595–606.
- Hsu TH, Spindler KR (2012) Identifying host factors that regulate viral infection. *PLoS Pathog* 8: e1002772.
- Hu S, Xie Z, Qian J, Blackshaw S & Zhu H (2011) Functional protein microarray technology. *Wiley Interdiscip Rev Syst Biol Med* 3: 255–268.
- da Huang W, Sherman BT & Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4: 44–57.
- Hultschig C, Kreutzberger J, Seitz H, Konthur Z, Büsow K & Lehrach H (2006) Recent advances of protein microarrays. *Curr Opin Chem Biol* 10: 4–10.
- Ivanov SS, Chung AS, Yuan ZL, Guan YJ, Sachs KV, Reichner JS & Chin YE (2004) Antibodies immobilized as arrays to profile protein post-translational modifications in mammalian cells. *Mol Cell Proteomics* 3: 788–795.
- Janzi M, Odling J, Pan-Hammarström Q, Sundberg M, Lundeberg J, Uhlén M, Hammarström L & Nilsson P (2005) Serum microarrays for large scale screening of protein levels. *Mol Cell Proteomics* 4: 1942–1947.
- Kacprowski T, Doncheva NT & Albrecht M (2013) NetworkPrioritizer: a versatile tool for network-based prioritization of candidate disease genes or other molecules. *Bioinformatics (Oxford, England)* 29: 1471–1473.
- Keating JA, Bhattacharya D, Lim PY, Falk S, Weisblum B, Bernard KA, Sharma M, Kuhn RJ & Striker R (2013) West Nile virus methyltransferase domain interacts with protein kinase G. *Virol J* 10: 242.
- Khadka S, Vangeloff AD, Zhang C, *et al.* (2011) A physical interaction network of dengue virus and human proteins. *Mol Cell Proteomics* 10: M111.012187.
- Knezevic V, Leethanakul C, Bichsel VE, Worth JM, Prabhu VV, Gutkind JS, Liotta LA, Munson PJ, Petricoin EF 3rd & Krizman DB (2001) Proteomic profiling of the cancer microenvironment by antibody arrays. *Proteomics* 1: 1271–1278.
- Korf U, Derdak S, Tresch A, *et al.* (2008) Quantitative protein microarrays for time-resolved measurements of protein phosphorylation. *Proteomics* 8: 4603–4612.
- Kshirsagar M, Carbonell J & Klein-Seetharaman J (2012) Techniques to cope with missing data in host-pathogen protein interaction prediction. *Bioinformatics (Oxford, England)* 28: i466–i472.
- Kshirsagar M, Carbonell J & Klein-Seetharaman J (2013) Multitask learning for host-pathogen protein interactions. *Bioinformatics* 29: i217–i226.
- Lange V, Picotti P, Domon B & Aebersold R (2008) Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol Syst Biol* 4: 222.
- Lietzén N, Öhman T, Rintahaka J, Julkunen I, Aittokallio T, Matikainen S & Nyman TA (2011) Quantitative subcellular proteome and secretome profiling of influenza A virus-infected human primary macrophages. *PLoS Pathog* 7: e1001340.
- Liotta LA, Espina V, Mehta AI, Calvert V, Rosenblatt K, Geho D, Munson PJ, Young L, Wulfkuhle J & Petricoin EF 3rd (2003) Protein microarrays: meeting analytical challenges for clinical applications. *Cancer Cell* 3: 317–325.
- Löbke C, Laible M, Rappl C, Ruschhaupt M, Sahin O, Arlt D, Wiemann S, Poustka A, Sültmann H & Korf U (2008) Contact spotting of protein microarrays coupled with spike-in of normalizer protein permits time-resolved analysis of ERBB receptor signaling. *Proteomics* 8: 1586–1594.

- MacBeath G & Schreiber SL (2000) Printing proteins as microarrays for high-throughput function determination. *Science* 289: 1760–1763.
- Mancone C, Ciccocanti F, Montaldo C, Perdomo AB, Piacentini M, Alonzi T, Fimia GM & Tripodi M (2013) Applying proteomic technology to clinical virology. *Clin Microbiol Infect* 19: 23–28.
- McCraith S, Holtzman T, Moss B & Fields S (2000) Genome-wide analysis of vaccinia virus protein-protein interactions. *P Natl Acad Sci USA* 97: 4879–4884.
- Miernyk JA & Thelen JJ (2008) Biochemical approaches for discovering protein-protein interactions. *Plant J* 53: 597–609.
- Montgomery SA, Berglund P, Beard CW & Johnston RE (2006) Ribosomal protein S6 associates with *Alphavirus* nonstructural protein 2 and mediates expression from *Alphavirus* messages. *J Virol* 80: 7729–7739.
- Mueller C, Liotta LA & Espina V (2010) Reverse phase protein microarrays advance to use in clinical trials. *Mol Oncol* 4: 461–481.
- Munday DC, Emmott E, Surtees R, Lardeau CH, Wu W, Duprex WP, Dove BK, Barr JN & Hiscox JA (2010) Quantitative proteomic analysis of A549 cells infected with human respiratory syncytial virus. *Mol Cell Proteomics* 9: 2438–2459.
- Munday DC, Surtees R, Emmott E, Dove BK, Digard P, Barr JN, Whitehouse A, Matthews D & Hiscox JA (2012) Using SILAC and quantitative proteomics to investigate the interactions between viral and host proteomes. *Proteomics* 12: 666–672.
- Narayanan A, Kehn-Hall K, Senina S, Lundberg L, Van Duyne R, Guendel I, Das R *et al.* (2012) Curcumin inhibits Rift Valley fever virus replication in human cells. *J Biol Chem* 287: 33198–33214.
- Nishizuka S, Charboneau L, Young L, *et al.* (2003) Proteomic profiling of the NCI-60 cancer cell lines using new high-density reverse-phase lysate microarrays. *P Natl Acad Sci USA* 100: 14229–14234.
- Noisakran S, Sengsai S, Thongboonkerd V, *et al.* (2008) Identification of human hnRNP C1/C2 as a dengue virus NS1-interacting protein. *Biochem Biophys Res Commun* 372: 67–72.
- Patwa TH, Qiu Y, Zhao J, Simeone DM & Lubman DM (2009) All-liquid separations, protein microarrays, and mass spectrometry to interrogate serum proteomes: an application to serum glycoproteomics. *Methods Mol Biol* 520: 75–87.
- Patwa T, Li C, Simeone DM & Lubman DM (2010) Glycoprotein analysis using protein microarrays and mass spectrometry. *Mass Spectrom Rev* 29: 830–844.
- Pawelz CP, Charboneau L, Bichsel VE, Simone NL, Chen T, Gillespie JW, Emmert-Buck MR, Roth MJ, Petricoin EF III & Liotta LA (2001) Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. *Oncogene* 20: 1981–1989.
- Petricoin EF, Ardekani AM, Hitt BA, *et al.* (2002) Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 359: 572–577.
- Popova T, Espina V, Bailey C, Liotta L, Petricoin E & Popov S (2009) Anthrax infection inhibits the AKT signaling involved in the E-cadherin-mediated adhesion of lung epithelial cells. *FEMS Immunol Med Microbiol* 56: 129–142.
- Popova TG, Turell MJ, Espina V *et al.* (2010) Reverse-phase phosphoproteome analysis of signaling pathways induced by Rift Valley fever virus in human small airway epithelial cells. *PLoS ONE* 5: e13805.
- Prins KC, Cárdenas WB & Basler CF (2009) *Ebola virus* protein VP35 impairs the function of interferon regulatory factor-activating kinases IKKepsilon and TBK-1. *J Virol* 83: 3069–3077.
- Qian WJ, Jacobs JM, Liu T, Camp DG & Smith RD (2006) Advances and challenges in liquid chromatography-mass spectrometry based proteomic profiling for clinical applications. *Mol Cell Proteomics* 5: 1727–1744.
- Rapkiewicz A, Espina V, Zujewski JA, Lebowitz PF, Filie A, Wulfkuhle J, Camphausen K, Petricoin EF 3rd, Liotta LA & Abati A (2007) The needle in the haystack: application of breast fine-needle aspirate samples to quantitative protein microarray technology. *Cancer* 111: 173–184.
- Sanz-Pamplona R, Berenguer A, Cordero D, Riccadonna S, Solé X, Crous-Bou M, Guinó E *et al.* (2012) Clinical value of prognosis gene expression signatures in colorectal cancer: a systematic review. *PLoS One* 7: e48877.
- Sboner A, Karpikov A, Chen G, Smith M, Mattoon D, Dawn M, Freeman-Cook L, Schweitzer B & Gerstein MB (2009) Robust-linear-model normalization to reduce technical variability in functional protein microarrays. *J Proteome Res* 8: 5451–5464.
- Schena M (2000) *Microarray Biochip Technology*. Eaton Publishers, Natick, MA.
- Sevecka M & MacBeath G (2006) State-based discovery: a multidimensional screen for small-molecule modulators of EGF signaling. *Nat Methods* 3: 825–831.
- Shapira SD, Gat-Viks I, Shum BO, *et al.* (2009) A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. *Cell* 139: 1255–1267.
- Sheehan KM, Calvert VS, Kay EW, *et al.* (2005) Use of reverse phase protein microarrays and reference standard development for molecular network analysis of metastatic ovarian carcinoma. *Mol Cell Proteomics* 4: 346–355.
- Shi T, Su D, Liu T, Tang K, Camp DG 2nd, Qian WJ & Smith RD (2012) Advancing the sensitivity of selected reaction monitoring-based targeted quantitative proteomics. *Proteomics* 12: 1074–1092.
- Spurrer B, Ramalingam S & Nishizuka S (2008) Reverse-phase protein lysate microarrays for cell signaling analysis. *Nat Protoc* 3: 1796–1808.
- Sreekumar A, Nyati MK, Varambally S, Barrette TR, Ghosh D, Lawrence TS & Chinnaiyan AM (2001) Profiling of cancer cells using protein microarrays: discovery of novel radiation-regulated proteins. *Cancer Res* 61: 7585–7593.
- Sutandy FX, Qian J, Chen CS & Zhu H (2013) Unit 27. Overview of protein microarrays. *Curr Protoc Protein Sci* 72: 27.1.1–27.1.16.
- Tafforeau L, Chantier T, Pradezynski F, Pellet J, Mangeot PE, Vidalain PO, Andre P, Combe CR & Lotteau V (2011) Generation and comprehensive analysis of an influenza virus polymerase cellular interaction network. *J Virol* 85: 13010–13018.
- Teng Z, Guo M, Liu X, Dai Q, Wang C & Xuan P (2013) Measuring gene functional similarity based on group-wise comparison of GO terms. *Bioinformatics* btt160. doi:10.1093/bioinformatics/btt160.
- Toss A, De ME, Rossi E, Casa LD, Iannone A, Federico M & Cortesi L (2013) Ovarian cancer: can proteomics give new insights for therapy and diagnosis? *Int J Mol Sci* 14: 8271–8290.
- VanMeter AJ, Rodriguez AS, Bowman ED, *et al.* (2008) Laser capture microdissection and protein microarray analysis of human non-small cell lung cancer: differential epidermal growth factor receptor (EGFR) phosphorylation events associated with mutated EGFR compared with wild type. *Mol Cell Proteomics* 7: 1902–1924.
- Von Brunn A, Teepe C, Simpson JC, Pepperkok R, Friedel CC, Zimmer R, Roberts R, Baric R & Haas J (2007) Analysis of intraviral protein-protein interactions of the SARS coronavirus ORF6. *PLoS ONE* 2: e459.
- White EA & Howley PM (2013) Proteomic approaches to the study of papillomavirus-host interactions. *Virology* 435: 57–69.
- Wiese S, Reidegeld KA, Meyer HE & Warscheid B (2007) Protein labeling by iTRAQ: a new tool for quantitative mass spectrometry in proteome research. *Proteomics* 7: 340–350.

- Wilson B, Liotta LA & Petricoin E 3rd (2010) Monitoring proteins and protein networks using reverse phase protein arrays. *Dis Markers* 28: 225–232.
- Wulfkuhle J, Espina V, Liotta L & Petricoin E (2004) Genomic and proteomic technologies for individualisation and improvement of cancer treatment. *Eur J Cancer* 40: 2623–2632.
- Zheng J, Sugrue RJ & Tang K (2011) Mass spectrometry based proteomic studies on viruses and hosts – a review. *Anal Chim Acta* 702: 149–159.
- Zhou ST, Liu R, Zhao X, Huang C & Wei Y (2011) Viral proteomics: the emerging cutting-edge of virus research. *Sci China Life Sci* 54: 502–512.
- Zhu H, Hu S, Jona G, *et al.* (2006) Severe acute respiratory syndrome diagnostics using a coronavirus protein microarray. *P Natl Acad Sci USA* 103: 4011–4016.
- Zhu H, Cox E & Qian J (2012) Functional protein microarray as molecular decathlete: a versatile player in clinical proteomics. *Proteomics Clin Appl* 6: 548–562.