Genome **Medicine**

Mass spectrometry for translational proteomics: progress and clinical implications

Erin Shammel Baker, Tao Liu, Vladislav A Petyuk, Kristin E Burnum-Johnson, Yehia M Ibrahim, Gordon A Anderson and Richard D Smith*

Abstract

The utility of mass spectrometry (MS)-based proteomic analyses and their clinical applications have been increasingly recognized over the past decade due to their high sensitivity, specificity and throughput. MS-based proteomic measurements have been used in a wide range of biological and biomedical investigations, including analysis of cellular responses and disease-specific post-translational modifications. These studies greatly enhance our understanding of the complex and dynamic nature of the proteome in biology and disease. Some MS techniques, such as those for targeted analysis, are being successfully applied for biomarker verification, whereas others, including global quantitative analysis (for example, for biomarker discovery), are more challenging and require further development. However, recent technological improvements in sample processing, instrumental platforms, data acquisition approaches and informatics capabilities continue to advance MS-based applications. Improving the detection of significant changes in proteins through these advances shows great promise for the discovery of improved biomarker candidates that can be verified pre-clinically using targeted measurements, and ultimately used in clinical studies - for example, for early disease diagnosis or as targets for drug development and therapeutic intervention. Here, we review the current state of MS-based proteomics with regard to its advantages and current limitations, and we highlight its translational applications in studies of protein biomarkers.

Keywords biomarker, clinical proteomics, ion mobility separations, mass spectrometry, multiple reaction monitoring, selected reaction monitoring, shotgun proteomics, targeted proteomics, translational proteomics

*Correspondence: rds@pnnl.gov
Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA 99352, USA

Translational proteomics: the importance of mass spectrometry-based approaches

Interest in using mass spectrometry (MS) for clinical analyses has grown significantly in the past few years due to its success in studies of human specimens, such as its recent applications for single cell analysis of bone marrow [1], direct blood sampling in multiple disease states, such as cardiac injury [2] and breast cancer [3], and the identification of Gram-negative bacilli in multiple clinical samples, including blood, tissue and urine [4,5]. MS analyses are utilized to obtain highly accurate mass measurements of molecules in a sample, and can sensitively detect and identify molecules and subtle changes in their composition and abundance. In particular, MS-based proteomic applications have received considerable attention. Proteomics is the study of the entire complement of proteins in an organism, tissue or cell and their changes under different conditions, from disease states to environmental variations. It has been estimated that the human proteome contains more than 2 million different protein products or 'proteoforms' [6-8].

Since human proteins perform cellular functions essential to health and/or disease, obtaining knowledge of their presence and variance is of great importance in understanding disease states and for advancing translational studies, especially those related to personalized medicine [9,10]. Human blood contains combinations of potentially detectable proteins from different parts of the body, and may be the single most informative sample for characterizing an individual's health [11]. From a clinical perspective, finding specific disease markers or biomarkers in such fluids represents an attractive alternative to tissue samples, due to the relative ease and less invasive nature of collection, and the large volumes that are normally obtainable. Proteomic studies promise to provide insights into the dynamic nature of biological systems through analysis of the proteins in biofluid and tissue samples, thereby defining the state of the organism at the molecular level. This approach not only incorporates the complexity of gene expression, but importantly also allows characterization of proteoforms generated by post-translational processes. Proteome measurements

therefore have great potential for translational application, since both normal and altered cellular functions of the human body are ultimately dependent on the expression and regulation of proteins. Moreover, disruptions in protein expression are likely to serve as early indicators of disease (that is, biomarkers) or targets for drug development and therapeutic intervention. These promising clinical applications have driven the development of MS-based approaches for proteomics, as well as other -omic level analyses, for studying human biofluid and tissue samples (Figure 1).

Over the past decade, studies of protein biomarkers have allowed advances in early, non-invasive diagnosis of significant diseases, such as the identification of C-reactive protein and troponin I as biomarkers for myocardial infarction, and prostate specific antigen for prostate cancer [12-14]. Despite these successes, efforts to identify biomarkers have not been nearly as successful as originally anticipated since proteomic analyses of blood and other biological fluids have proven to be immensely challenging because of the enormous complexity of the samples, the vast dynamic range of protein concentrations of potential interest (for example, greater than ten orders of magnitude in blood plasma), and the fact that analytes of clinical interest are often present at the low end of this concentration range [15-17]. To further exacerbate these challenges, verification and population-scale validation of biomarkers requires the analysis of hundreds or even thousands of high-quality clinical samples. The collection and storage of these samples must be done carefully and monitored using standardized protocols to reduce variations due to endogenous enzyme activities or sample contamination. These studies also require multiple control groups and diagnostic sub-categories of patients that are ideally gathered longitudinally over the course of disease progression. The analysis of many patient samples is required to characterize normal human genetic heterogeneity and disease heterogeneity [18,19]. High throughput measurements are therefore essential to achieve biostatistical significance (Figure 2).

While current MS-based proteomic measurements are capable of providing great depth of coverage through the use of extensive fractionation and analysis, this generally precludes the throughput required and the levels of sensitivity and specificity necessary for the rapid identification of clinically useful biomarkers. However, recent technological advances in automated parallel sample processing methods [20], multidimensional separations prior to MS [21,22], instrumentation components and approaches [23-27], and high-performance informatics tools [28-30] have facilitated measurements with both increased sensitivity and higher throughput for translational applicatio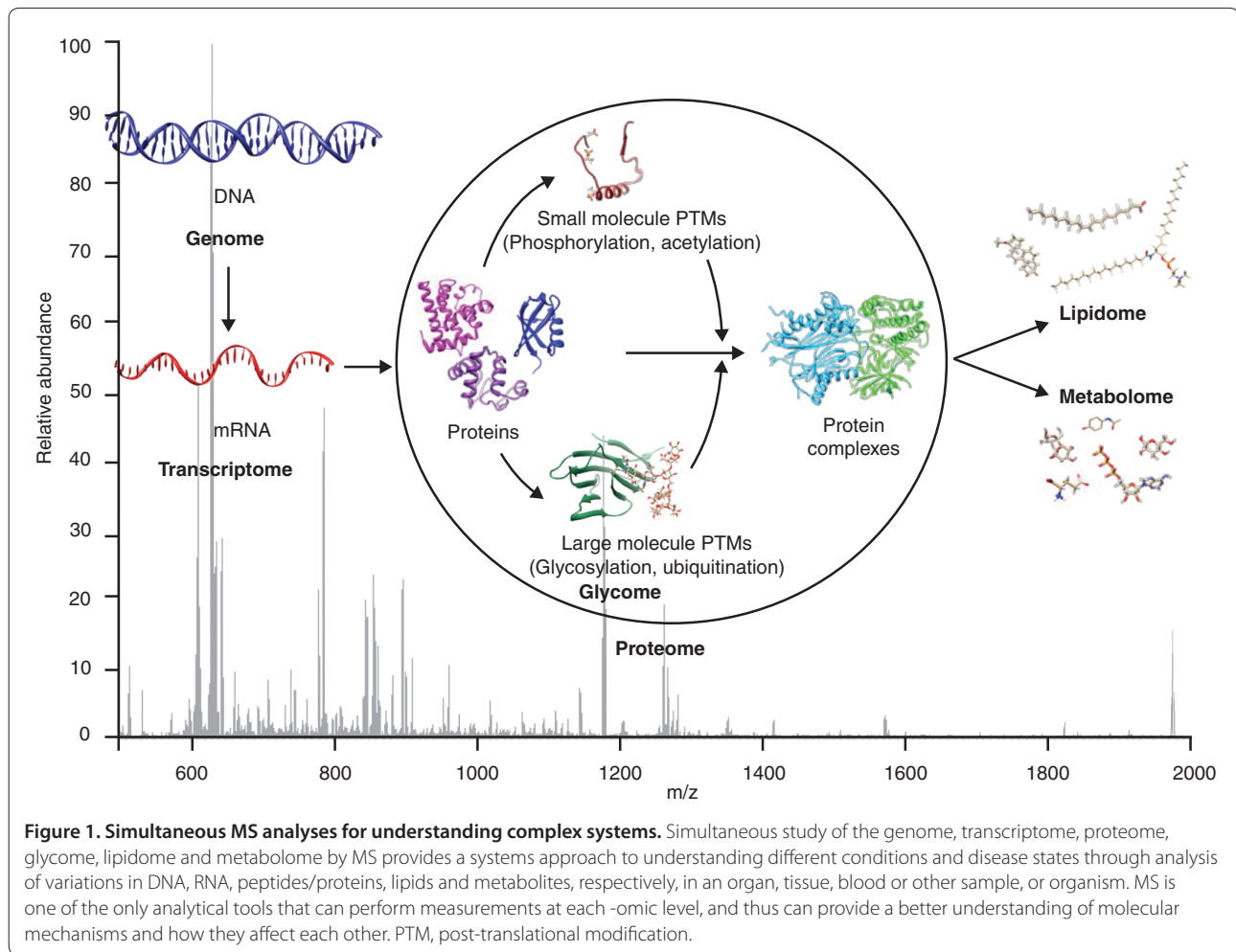ns. In this review, we discuss the current state of MS-based proteomics with regard to its advantages and current limitations, and we highlight translational applications that are being enabled by these recent technological advances.

Advances in MS-based translational proteomics

The primary translational application of MS-based proteomics is biomarker development. However, as already mentioned, its success has so far been quite modest and has been mainly limited to preclinical studies. Biomarker development is a multi-stage process that consists of discovery, verification, validation and commercialization [15]. For MS, the measurements fall into two categories, where the first utilizes a discovery approach to identify potential protein biomarkers and the second involves verification to further assess and initially validate these biomarkers using a larger population. Performing high-quality measurements and rigorous statistical analyses are essential in both steps as valuable patient samples are used. Currently, both MS-based proteomic discovery and verification approaches use bottom-up methods (Figure 3) in which proteins are digested into smaller peptides before analysis [31]. However, the two approaches aim to obtain different types of information.

Discovery approaches

In the discovery phase, broad quantitative MS measurements often aim to identify peptides and proteins that differ significantly in abundances between patient and control groups. The main advantage of this approach is its largely unbiased ability to characterize a whole proteome or enriched sub-proteome in a single measurement, so that the protein alterations corresponding to a pathological or biochemical condition at a given time can be investigated. However, performing discovery-based proteomic analysis has proven to be quite difficult using plasma and serum samples. In plasma, proteins have concentrations ranging from approximately 3×10^{10} pg/ml for albumin to the low pg/ml range for some cytokines and proteins, such as those potentially secreted or leaking into blood, for example from tumors (Figure 4a). Because of this huge dynamic range and the fact that the proteome in human biofluid samples is mainly represented by only a few high abundance proteins - the 22 most abundant proteins represent approximately 99% of the total protein mass (Figure 4b) - analyzing all plasma proteins simultaneously is enormously challenging [11,32], even after depletion of the most abundant proteins, as this exceeds the dynamic range of mass spectrometers that are typically used for discovery efforts (often approximately 1×10^3 to 1×10^4 for a single spectrum). To provide an extended dynamic range for increased protein coverage it is necessary to couple front-end separations such as liquid chromatography (LC), multi-stage immunoaffinity

Figure 1. Simultaneous MS analyses for understanding complex systems. Simultaneous study of the genome, transcriptome, proteome, glycome, lipidome and metabolome by MS provides a systems approach to understanding different conditions and disease states through analysis of variations in DNA, RNA, peptides/proteins, lipids and metabolites, respectively, in an organ, tissue, blood or other sample, or organism. MS is one of the only analytical tools that can perform measurements at each -omic level, and thus can provide a better understanding of molecular mechanisms and how they affect each other. PTM, post-translational modification.

depletion [33-35], fractionation [36], or a combination of all three with MS analyses. While advanced LC separations have already provided improvement in the depth of coverage for proteins detected in MS studies [37], a major problem is their concomitant reduction in throughput, as bottom-up LC-MS analyses typically require in the order of 1 h. The detection of more proteins (for example, thousands) from plasma is possible with extensive off-line fractionation prior to on-line LC-MS analyses [38], but days or weeks of LC-MS measurements are then necessary for analysis of the multiple fractions. While this approach is highly attractive for the detection and discovery of potential biomarkers, the inherently low throughput largely precludes population studies to enable investigation of human and disease heterogeneity, and also limits the possibility of performing personal profiling. Thus, technological advances that greatly decrease LC separation times or eliminate them entirely while still maintaining a high depth of coverage are crucial for future clinical applications.

To attain further information and identify unknown peptides with high accuracy in bottom-up MS studies, tandem MS (MS/MS) measurements, involving multiple steps of MS analysis and peptide fragmentation, are essential. Currently, many immunoassays used in translational studies measure analytes indirectly by detecting them through their interaction with other molecules, such as antibodies. MS provides an advantageous alternative to immunoassays as it involves direct measurements and allows the acquisition of exact peptide sequence information through high mass accuracy MS/MS measurements, thereby allowing unknown peptides to be identified with a great degree of confidence. The simultaneous collection of MS and MS/MS measurements involves the acquisition of a preliminary mass spectrum of intact peptides, followed by dissociation or fragmentation of a peptide(s) of interest, and acquisition of the fragmentation mass spectrum. This process is repeated for the duration of the entire LC separation, resulting in thousands of MS and MS/MS spectra. To
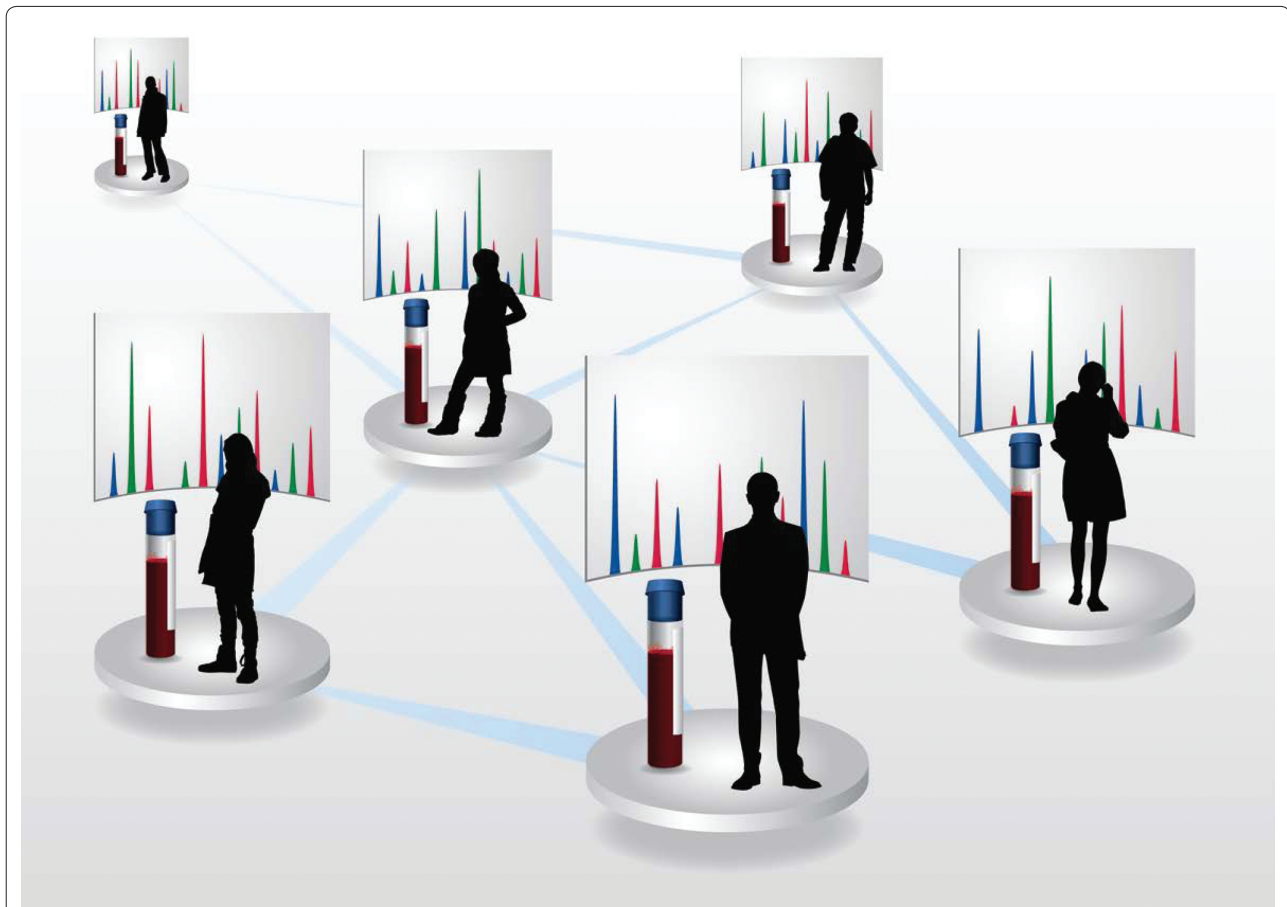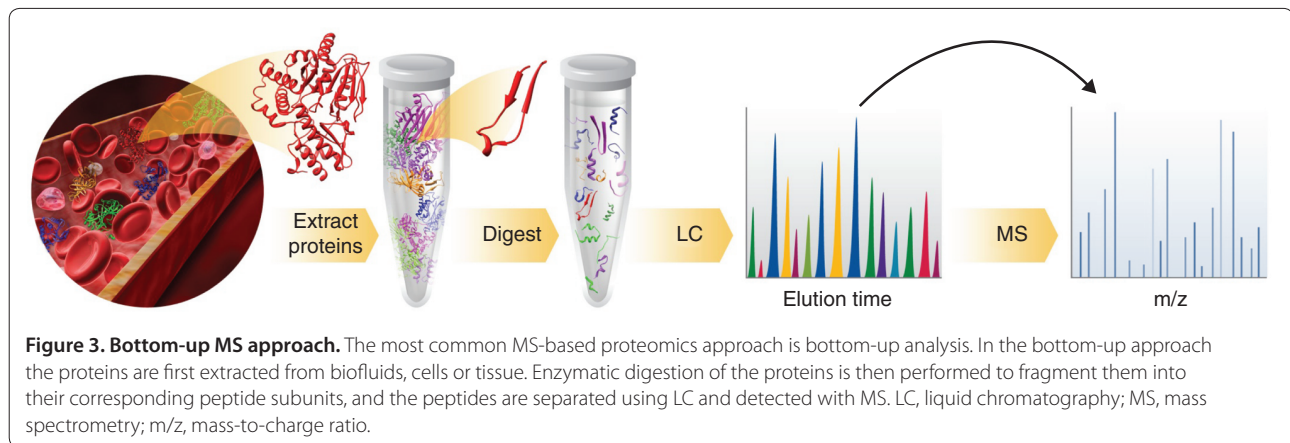
Figure 2. Biodiversity in population proteomic studies. Population proteomics allows the analysis of protein biodiversity within a population. Because it is known that individual variation, such as the presence of point mutations and varying protein abundances, will be present in all human studies (as depicted by the different chromatograms), it has become essential to develop high throughput, sensitive analytical applications to enable measurements necessary for personalized medicine.

identify the peptides with MS/MS, genomic data are frequently used to generate theoretical sequences for bioinformatics tools such as Mascot [39], Sequest [40] and X! Tandem [41]. By evaluating all of the matched MS/MS spectra, the false discovery rate of the peptide identifications can be estimated [42-44], and improved informatics tools are increasingly allowing identifications from spectra that were previously unattributed due to unexpected sequences or modification states.

Another important step in bottom-up measurements is quantification of observed peptides to determine if any significant changes are occurring between samples. Quantitative measurements of peptide abundance can be performed with or without stable isotope labeling (SIL) of peptides (or proteins) using peptide ion peak intensities or spectral counting (that is, 'label-free' quantification) [45]. Several *in vitro* and *in vivo* labeling techniques, such as stable isotope labeling of amino acids by cell culture (SILAC) [46,47], isobaric tags for relative and absolute quantification (iTRAQ) [48,49] and ^{18}O-labeling

[17] have been developed for MS-based quantification, and have been shown to provide lower standard deviations for peptide ion peak intensity measurements compared with the label-free methods [50]. When combined with off-line fractionation, these SIL methods provide broad coverage for comprehensive proteome characterization. However, label-free measurements using normalization of LC-MS analyses can also be quite effective and can avoid complications introduced by labeling approaches [51].

At present, data-dependent MS/MS analysis of selected peptides relies on an initial MS scan, and although it is widely used in proteomic discovery studies, it has inherent limitations that are associated with MS/MS undersampling in complex samples. To overcome these limitations and improve quantification, the accurate mass and time (AMT) tag strategy was developed for use on either labeled or label-free samples [52]. In a typical AMT tag study, a database is created and populated with peptide masses and LC elution times from many LC-MS/MS

Figure 3. Bottom-up MS approach. The most common MS-based proteomics approach is bottom-up analysis. In the bottom-up approach the proteins are first extracted from biofluids, cells or tissue. Enzymatic digestion of the proteins is then performed to fragment them into their corresponding peptide subunits, and the peptides are separated using LC and detected with MS. LC, liquid chromatography; MS, mass spectrometry; m/z, mass-to-charge ratio.

measurements using representative samples from experimental and control groups. High throughput LC-MS analyses are then performed for a large number of biological replicates and the acquired datasets are compared with the database to identify the peptides that are actually present. This approach allows the comparison of large numbers of peptide species that may not be identified in normal data-dependent MS/MS studies for reasons that include poor peptide fragmentation, co-elution of highly abundant species and/or informatics limitations - presumably similar factors that leave a significant number of detected species in LC-MS/MS analyses unidentified. Other approaches, such as data-independent MS/MS strategies [25-27] (discussed later), have also been developed recently to significantly enhance unbiased discovery studies.
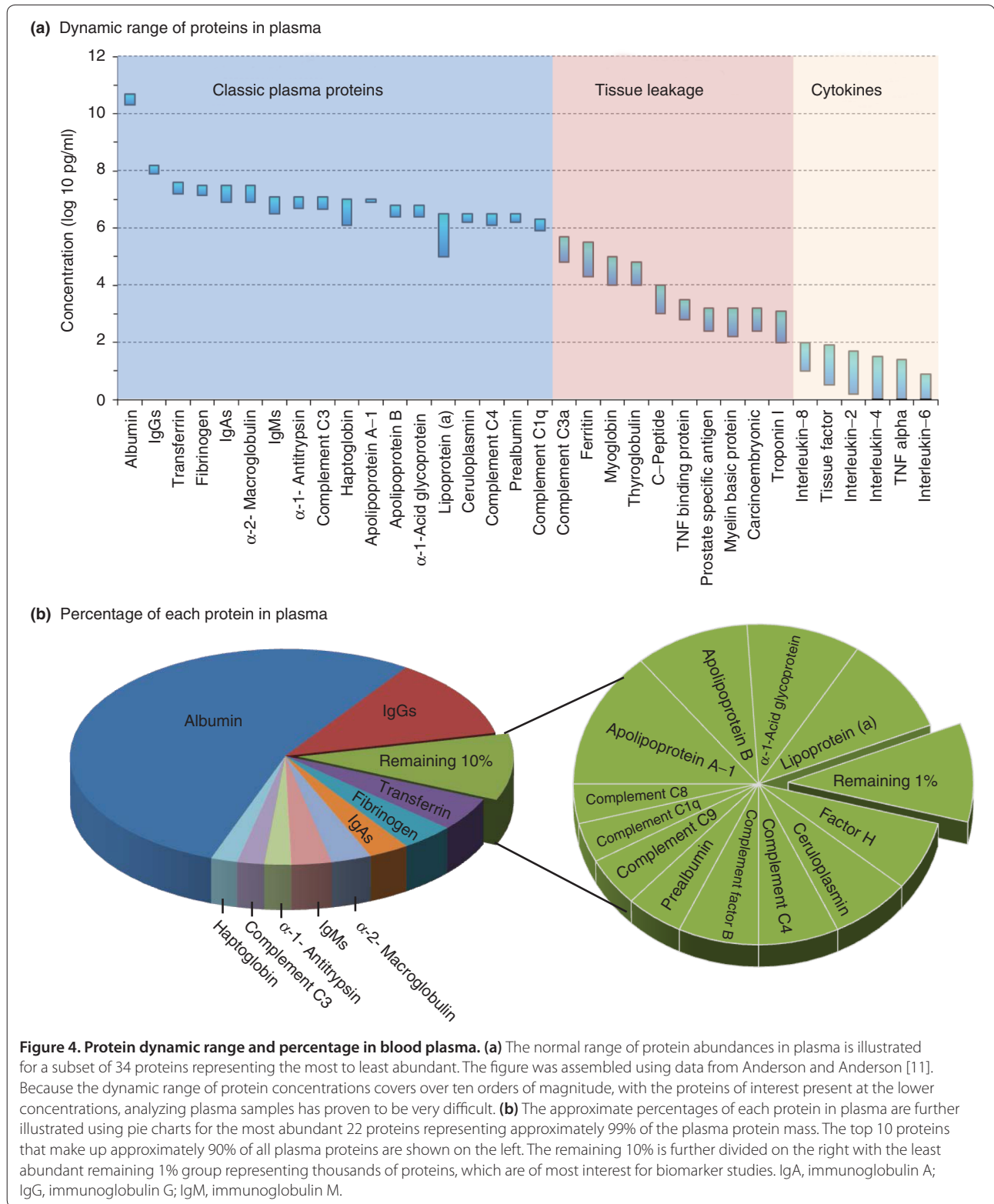
While these new approaches promise improved discovery of biomarkers, analysis of plasma samples still remains challenging for MS-based approaches. Discovery efforts increasingly use proximal fluids or tissues that are expected to be rich in biomarker candidates and present less of a challenge in terms of the dynamic range of proteins [15]. Various methods, such as optimal cutting temperature compound-embedded tissues, and formalin-fixation and paraffin-embedded tissues (with or without laser capture microdissection), have been developed for preparing clinical tissue samples for proteomic studies [53]. The results from these advanced preparation methods have been promising [54,55], and serve as a prelude to targeted discovery or verification efforts for measurements of candidate biomarker proteins at presumably much lower levels in blood samples.

Verification approaches
The verification phase typically uses a much larger number of samples, and focuses on a limited set of candidate peptides or proteins identified in the discovery approach. This approach can provide highly sensitive

quantification of protein abundances and aims to identify a set of biomarker candidates with greater confidence. There have been significant developments in MS-based methods for the verification approach, providing much greater sensitivity, specificity and throughput, and more accurate quantification than broad discovery-based measurements.

Targeted quantitative MS-based measurements typically employ selected reaction monitoring (SRM), using triple quadrupole mass spectrometers. In SRM measurements, the triple quadrupole MS allows rapid detection of a series of targeted peptide ions and their corresponding fragments (that is, transitions) with multiplexing and 'scheduling' capabilities (to perform pre-defined analyses during specific LC elution times) along with SIL internal peptide standards [56,57] to provide highly accurate quantification for up to hundreds of peptides during a single LC separation. The two-stage mass filtering in SRM (that is, for both peptide ions and their corresponding fragments) provides great sensitivity and specificity for detection of the targeted peptides. This capability often leads to observed limits of detection and limits of quantification (LOQ) of about 10 to 100 ng/ml in plasma - several orders of magnitude lower than presently feasible with discovery-based platforms. Moreover, recent advances such as the use of protein depletion, limited fractionation, and targeted peptide enrichment methodologies, such as peptide isolation with stable isotope standard capture with anti-peptide antibodies (SISCAPA) [58], extend practical LOQ values to low ng/ml (or even low pg/ml) levels in blood samples [33,59]. The implementation of other instrumental modifications such as multi-inlet capillaries and dual-stage ion funnels has led to further enhanced sensitivity [60]. While selection of the correct proteotypic or targeted peptides with good digestion and ionization efficiency requires some effort, this has increasingly been addressed using public repositories, including SRMAtlas

(a) Dynamic range of proteins in plasma

(b) Percentage of each protein in plasma

Figure 4. Protein dynamic range and percentage in blood plasma. (a) The normal range of protein abundances in plasma is illustrated for a subset of 34 proteins representing the most to least abundant. The figure was assembled using data from Anderson and Anderson [11]. Because the dynamic range of protein concentrations covers over ten orders of magnitude, with the proteins of interest present at the lower concentrations, analyzing plasma samples has proven to be very difficult. **(b)** The approximate percentages of each protein in plasma are further illustrated using pie charts for the most abundant 22 proteins representing approximately 99% of the plasma protein mass. The top 10 proteins that make up approximately 90% of all plasma proteins are shown on the left. The remaining 10% is further divided on the right with the least abundant remaining 1% group representing thousands of proteins, which are of most interest for biomarker studies. IgA, immunoglobulin A; IgG, immunoglobulin G; IgM, immunoglobulin M.

[61,62], PeptideAtlas [63] and the Global Proteome Machine [64]. Recent computational developments have also allowed the creation of programs that effectively predict proteotypic peptides given a protein amino acid sequence [65,56], allowing the list of targeted peptides to be derived without the need to rely on discovery-based

proteomics data. Moreover, MS targeted measurements have also proved reproducible in assays across many different proteomics laboratories [66]. These various features have made SRM the current method of choice for ultra-sensitive MS-based biomarker verification (or pre-clinical validation).

Addressing the challenges of translational proteomics

Despite significant advances in MS-based targeted analyses, several performance metrics, including measurement throughput and detection sensitivity, still require compromises to biomarker discovery and verification approaches for the translational application of MS-based proteomic analyses. In particular, these deficiencies result in low sampling numbers and measurement quality that prevents detection of proteins present at low concentrations. To achieve further progress in translational proteomics, technological developments in MS such as faster separations, more effective ion sources, higher instrumental resolution/mass accuracy, detectors with greater dynamic range, and advanced data acquisition approaches are expected to increasingly allow broad non-targeted measurements that retain the benefits of targeted approaches.
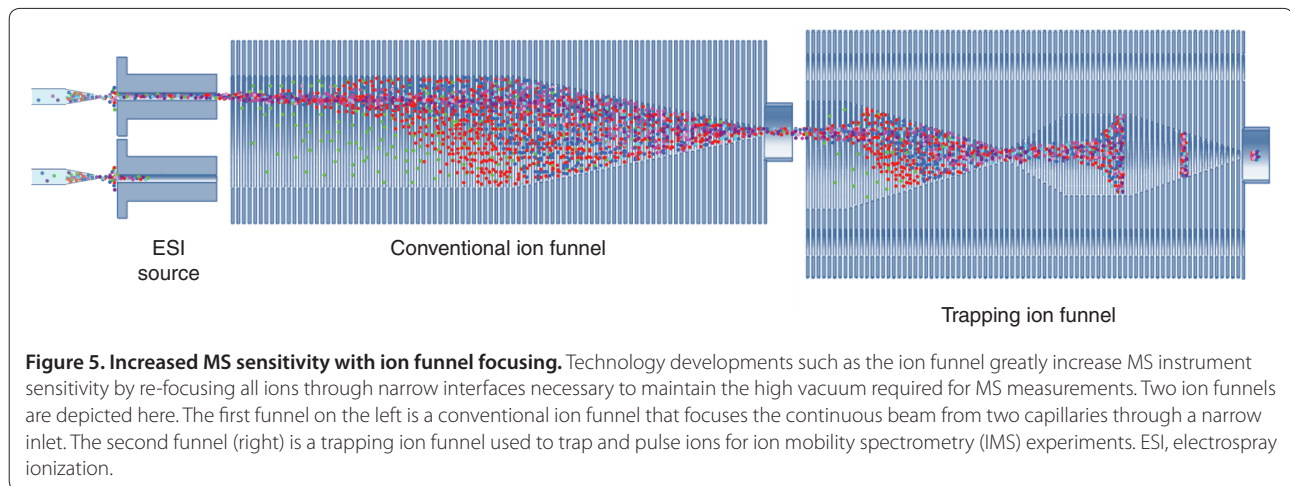
Data-independent MS/MS acquisition has shown promise for improving the consistency of peptide identifications, as well as for increasing protein sequence coverage in complex samples and creating broad untargeted measurements that are more similar to possible targeted measurements [25-27]. Data-independent acquisition is a strategy that systematically queries sample sets for the presence and quantity of essentially any protein of interest, using the information available in fragment ion spectral libraries to mine complete fragment ion maps. One way of performing data-independent acquisitions is by using sequential window acquisition of all theoretical fragment-ion spectra (SWATH™) MS, in which repeated cycling of a 25 Da precursor isolation window is used in a single analysis to obtain time-resolved fragment ion spectra for all analytes detectable within a user defined mass-to-charge ratio (m/z) precursor range. Initial results have been very promising, with queried peptides quantified with a consistency and accuracy apparently approaching that for SRM [25].Another approach to exploit data-independent acquisitions involves using an additional separation technique prior to fragmentation to increase measurement sensitivity and the ability to associate simultaneously fragmented precursor ions with their corresponding fragment ions. Fast gas-phase ion-mobility spectrometry (IMS), taking place in a timescale of tens of milliseconds, offers an attractive ion separation approach for data-independent acquisitions. IMS was introduced in the 1970s [67] and utilizes the fact that ions subject to an electric field in a buffer gas quickly reach a steady velocity dependent on the ion shape: compact species drift faster than those with extended structures [68,69]. IMS can be easily coupled to quadrupole time-of-flight MS, allowing placement of IMS between the LC and MS stages. The resulting IMS-MS instrument produces high-resolution spectra containing both the m/z and IMS drift time information concurrently. To perform data-independent acquisitions, all precursor ions are fragmented after the IMS separation but prior to MS detection to completely eliminate MS/MS undersampling. Because fragmentation occurs after the IMS separation, all fragment ions have the same drift time as the precursors [70-72], allowing simplification of spectral deconvolution, which adds a great benefit to this technique.

The increased sensitivity and reduced spectral congestion in the IMS separation also has another advantage of reducing or completely avoiding the LC time in complex samples [73]. When IMS is coupled with MS, ions are separated prior to detection, reducing detector suppression while supplying an additional dimension for peptide identification. Practical use of IMS-MS was initially impeded by low sensitivity due to significant ion losses at the IMS terminus and during transfer to MS. However, this problem was solved by re-focusing ions with an ion funnel at the IMS-MS interface [74], essentially preventing ion losses during the operation. The introduction of ion funnels in 1997 [75] provided a huge improvement in sensitivity of MS instruments as it allowed ions to be focused through the small interface orifices necessary for ultralow MS vacuum pressures (1×10^{-7} to 1×10^{-8} Torr), as shown in Figure 5. The ion funnel is most often implemented in the source region of mass spectrometers to greatly increase the sensitivity of measurements, and has gained importance with its recent inclusion in commercially developed instruments. While these developments are just a first step in the convergence of discovery and verification platforms, further progress will be facilitated by emerging approaches for faster and higher resolution separations, improved MS resolution and extended detector dynamic range.

Clinical implications

The potential to use MS-based proteomics in clinical settings is largely judged by their ability to make robust, sensitive, quantitative, specific and high-throughput measurements for highly complex biospecimens. Clinical questions and the corresponding requirements for biospecimen detection determine the ability of MS to find and routinely measure high-quality biomarkers that have sufficient sensitivity and specificity to be clinically useful in screening large populations - for example, for diagnostic tests or early disease detection. For instance, SRM has already been widely used for measuring metabolites

Figure 5. Increased MS sensitivity with ion funnel focusing. Technology developments such as the ion funnel greatly increase MS instrument sensitivity by re-focusing all ions through narrow interfaces necessary to maintain the high vacuum required for MS measurements. Two ion funnels are depicted here. The first funnel on the left is a conventional ion funnel that focuses the continuous beam from two capillaries through a narrow inlet. The second funnel (right) is a trapping ion funnel used to trap and pulse ions for ion mobility spectrometry (IMS) experiments. ESI, electrospray ionization.

for newborn screening in clinical laboratories [76]. However, the detection of low abundance proteins in complex samples requires measurements of high sensitivity and large dynamic range, aspects of MS performance that are currently achievable and presently being greatly improved due to developments in MS-based instruments and new separation technologies, as mentioned earlier. Narrowing large lists of differentially abundant proteins into defined patterns of biologically important variations, to reveal a much smaller set of candidate proteins that can be detected with the high sensitivity and specificity that is needed for clinical utility, requires verification studies typically involving many hundreds of samples at a minimum. Currently, the low throughput of conventional MS platforms severely inhibits biomarker verification. However, recent improvements in MS-based proteomic approaches, ranging from sample processing to data acquisition as outlined earlier, are resulting in rapid and highly sensitive MS analyses that are now providing a viable future for MS-based measurements in clinical laboratories.

When comparing MS-based proteomics with immunoassays, which are the current gold standard in clinical detection of protein biomarkers, MS-based proteomics offer a significantly shorter lead-time and cost for assay development, high capability for multiplexed analysis, and the ability to be highly configurable or flexible for measuring different clinical analytes. With these advances and unique features, MS-based translational proteomics have the potential to become powerful tools for decision making in the clinic, alongside other approaches such as physical examination, *in vivo* imaging, histology, biochemical assays and assessment of demographic risk factors. Their potential applications for discovering and measuring protein biomarkers could include routine screening, staging of disease progression, prediction of the course of disease, assessment of disease outcome, monitoring disease recurrence, and personalized assessment of drug response and toxicity, to name a few.

Outlook and perspectives

The future of MS-based translational proteomics can be categorized by what is currently practical, and what is being enabled by recent technological developments. In the short term, proteomic measurements using targeted approaches are effective for high sensitivity and high throughput analysis of a limited set of biomarker candidates, whereas unbiased broad measurements are effective for the detection of a much larger universe of biomarker candidates but with less sensitivity [27]. Improvements to the sensitivity of broad measurements and the scope of targeted measurements are ultimately driving a convergence of these platforms and are expected to increase the ability to gain a predictive understanding of molecular processes in complex biological systems [77]. While MS-based proteomics offers valuable information for understanding complex biological systems, systems-level quantitative analyses using a combination of broad proteomic, metabolomic, lipidomic and glycomic MS analyses (termed pan-omics) will be increasingly important (Figure 1). These approaches largely benefit from the same MS-based platform developments that are allowing advances in translational proteomics. The importance of each -omic measurement technology has already become evident - for example, through the success of targeted measurements [76,78] - and their combination into transformative pan-omics measurement capabilities would likely be crucial for understanding the complexity of biological systems. Thus, broad pan-omic discovery methods, if sufficiently sensitive and effective, would be expected to provide a much more informative clinical toolset of biomarkers for accurate prediction of disease onset, and for disease monitoring and prognosis.

Abbreviations
AMT, accurate mass and time; IMS, ion mobility spectrometry; LC, liquid chromatography; LOQ, limits of quantification; m/z, mass-to-charge ratio; MS, mass spectrometry; MS/MS, tandem mass spectrometry; SIL, stable isotope labeling; SRM, selected reaction monitoring.

Competing interests
The authors declare that they have no competing interests.

Published: 31 August 2012

References
1. Bendall SC, Simonds EF, Qiu P, Amir el-AD, Krutzik PO, Finck R, Bruggner RV, Melamed R, Trejo A, Ornatsky OI, Balderas RS, Plevritis SK, Sachs K, Pe'er D, Tanner SD, Nolan GP: **Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum.** *Science* 2011, **332**:687-696.
2. Addona TA, Shi X, Keshishian H, Mani DR, Burgess M, Gillette MA, Clauser KR, Shen D, Lewis GD, Farrell LA, Fifer MA, Sabatine MS, Gerszten RE, Carr SA: **A pipeline that integrates the discovery and verification of plasma protein biomarkers reveals candidate markers for cardiovascular disease.** *Nat Biotechnol* 2011, **29**:635-643.
3. Whiteaker JR, Lin C, Kennedy J, Hou L, Trute M, Sokal I, Yan P, Schoenherr RM, Zhao L, Voytovich UJ, Kelly-Spratt KS, Krasnoselsky A, Gafken PR, Hogan JM, Jones LA, Wang P, Amon L, Chodosh LA, Nelson PS, McIntosh MW, Kemp CJ, Paulovich AG: **A targeted proteomics-based pipeline for verification of biomarkers in plasma.** *Nat Biotechnol* 2011, **29**:625-634.
4. Saffert RT, Cunningham SA, Ihde SM, Jobe KE, Mandrekar J, Patel R: **Comparison of Bruker Biotyper matrix-assisted laser desorption ionization-time of flight mass spectrometer to BD Phoenix automated microbiology system for identification of gram-negative bacilli.** *J Clin Microbiol* 2011, **49**:887-892.
5. Bizzini A, Durussel C, Bille J, Greub G, Prod'hom G: **Performance of matrix-assisted laser desorption ionization-time of flight mass spectrometry for identification of bacterial strains routinely isolated in a clinical microbiology laboratory.** *J Clin Microbiol* 2010, **48**:1549-1554.
6. American Medical Association. **Proteomics** [http://www.ama-assn.org/ama/pub/physician-resources/medical-science/genetics-molecular-medicine/current-topics/proteomics.page]
7. Pruitt K, Brown G, Tatusova T, Maglott D: **The Reference Sequence (RefSeq) Database.** The NCBI Handbook [Internet]. 9 October 2002 (updated 6 April 2012) [http://www.ncbi.nlm.nih.gov/books/NBK21091]
8. Pruitt KD, Tatusova T, Maglott DR: **NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins.** *Nucleic Acids Res* 2005, **33**:D501-504.
9. Chan IS, Ginsburg GS: **Personalized medicine: progress and promise.** *Annu Rev Genomics Hum Genet* 2011, **12**:217-44.
10. Hutchinson L: **Personalized cancer medicine: era of promise and progress.** *Nat Rev Clin Oncol* 2011, **8**:121.
11. Anderson NL, Anderson NG: **The human plasma proteome: history, character, and diagnostic prospects.** *Mol Cell Proteomics* 2002, **1**:845-867.
12. Catalona WJ, Partin AW, Slawin KM, Brawer MK, Flanigan RC, Patel A, Richie JP, deKernion JB, Walsh PC, Scardino PT, Lange PH, Subong EN, Parson RE, Gasior GH, Loveland KG, Southwick PC: **Use of the percentage of free prostate-specific antigen to enhance differentiation of prostate cancer from benign prostatic disease: a prospective multicenter clinical trial.** *JAMA* 1998, **279**:1542-1547.
13. Antman EM, Tanasijevic MJ, Thompson B, Schactman M, McCabe CH, Cannon CP, Fischer GA, Fung AY, Thompson C, Wybenga D, Braunwald E: **Cardiac-specific troponin I levels to predict the risk of mortality in patients with acute coronary syndromes.** *N Engl J Med* 1996, **335**:1342-1349.
14. Danesh J, Wheeler JG, Hirschfield GM, Eda S, Eiriksdottir G, Rumley A, Lowe GD, Pepys MB, Gudnason V: **C-reactive protein and other circulating markers of inflammation in the prediction of coronary heart disease.** *N Engl J Med* 2004, **350**:1387-1397.
15. Rifai, N, Gillette MA, Carr SA: **Protein biomarker discovery and validation: the long and uncertain path to clinical utility.** *Nat Biotechnol* 2006, **24**:971-983.
16. Anderson NL: **The roles of multiple proteomic platforms in a pipeline for new diagnostics.** *Mol Cell Proteomics* 2005, **4**:1441-1444.
17. Jacobs JM, Adkins JN, Qian WJ, Liu T, Shen Y, Camp DG 2nd, Smith RD: **Utilizing human blood plasma for proteomic biomarker discovery.** *J Proteome Res* 2005, **4**:1073-1085.
18. Ricós C, Iglesias N, García-Lario JV, Simón M, Cava F, Hernández A, Perich C, Minchinela J, Alvarez V, Doménech MV, Jiménez CV, Biosca C, Tena R: **Within-subject biological variation in disease: collated data and clinical consequences.** *Ann Clin Biochem* 2007, **44**:343-352.
19. Nedelkov D, Kiernan UA, Niederkofler EE, Tubbs KA, Nelson RW: **Population proteomics: the concept, attributes, and potential for cancer biomarker research.** *Mol Cell Proteomics* 2006, **5**:1811-1818.
20. Mirsky P, Chatterjee A, Sauer-Budge AF, Sharon A: **An automated, parallel processing approach to biomolecular sample preparation.** *J Lab Autom* 2012, **17**:116-124.
21. Unwin RD, Griffiths JR, Whetton AD: **Simultaneous analysis of relative protein expression levels across multiple samples using iTRAQ isobaric tags with 2D nano LC-MS/MS.** *Nat Protoc* 2010, **5**:1574-1582.
22. Gilar M, Greibrokk T: **2D LC - embraced by life-science research and drug development.** *J Sep Sci* 2012, **35**:NA.
23. Belov ME, Prasad S, Prior DC, Danielson WF 3rd, Weitz K, Ibrahim YM, Smith RD: **Pulsed multiple reaction monitoring approach to enhancing sensitivity of a tandem quadrupole mass spectrometer.** *Anal Chem* 2011, **83**:2162-2171.
24. Ibrahim Y, Belov ME, Tolmachev AV, Prior DC, Smith RD: **Ion funnel trap interface for orthogonal time-of-flight mass spectrometry.** *Anal Chem* 2007, **79**:7845-7852.
25. Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, Bonner R, Aebersold R: **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* 2012, **11**:O111.016717.
26. Panchaud A, Scherl A, Shaffer SA, von Haller PD, Kulasekara HD, Miller SI, Goodlett DR: **Precursor acquisition independent from ion count: how to dive deeper into the proteomics ocean.** *Anal Chem* 2009, **81**:6481-6488.
27. Röst H, Malmström L, Aebersold R: **A computational tool to detect and avoid redundancy in selected reaction monitoring.** *Mol Cell Proteomics* 2012, **11**:540-549.
28. Schilling B, Rardin MJ, MacLean BX, Zawadzka AM, Frewen BE, Cusack MP, Sorensen DJ, Bereman MS, Jing E, Wu CC, Verdin E, Kahn CR, Maccoss MJ, Gibson BW: **Platform-independent and label-free quantitation of proteomic data using MS1 extracted ion chromatograms in skyline: application to protein acetylation and phosphorylation.** *Mol Cell Proteomics* 2012, **11**:202-214.
29. Farrah T, Deutsch EW, Kreisberg R, Sun Z, Campbell DS, Mendoza L, Kusebauch U, Brusniak MY, Hüttenhain R, Schiess R, Selevsek N, Aebersold R, Moritz RL: **PASSEL: The PeptideAtlas SRMexperiment library.** *Proteomics* 2012, **12**:1170-1175.
30. Swaney, DL, McAlister GC, Coon JJ: **Decision tree-driven tandem mass spectrometry for shotgun proteomics.** *Nat Methods* 2008, **5**:959-964.
31. Washburn MP, Wolters D, Yates JR 3rd: **Large-scale analysis of the yeast proteome by multidimensional protein identification technology.** *Nat Biotechnol* 2001, **19**:242-247.
32. Schuchard M, Melm C, Crawford H, Chapman S, Cockrill S, Ray K, Mehigh R, Chen D, Scott G: **Specific depletion of twenty high abundance proteins from human plasma.** Poster presented at NCI Proteomic Technologies Reagents Resource Workshop, 12-13 December 2005. [http://www.scribd.com/doc/12287730/Specific-Depletion-of-Twenty-High-Abundance-Proteins]
33. Shi T, Zhou JY, Gritsenko MA, Hossain M, Camp DG 2nd, Smith RD, Qian WJ: **IgY14 and SuperMix immunoaffinity separations coupled with liquid chromatography-mass spectrometry for human plasma proteomics**

biomarker discovery. *Methods* 2012, **56**:246-253.

34. Qian WJ, Kaleta DT, Petritis BO, Jiang H, Liu T, Zhang X, Mottaz HM, Varnum SM, Camp DG 2nd, Huang L, Fang X, Zhang WW, Smith RD: **Enhanced detection of low abundance human plasma proteins using a tandem IgY12-SuperMix immunoaffinity separation strategy.** *Mol Cell Proteomics* 2008, **7**:1963-1973.

35. Zhou JY, Petritis BO, Petritis K, Norbeck AD, Weitz KK, Moore RJ, Camp DG, Kulkarni RN, Smith RD, Qian WJ: **Mouse-specific tandem IgY7-SuperMix immunoaffinity separations for improved LC-MS/MS coverage of the plasma proteome.** *J Proteome Res* 2009, **8**:5387-5395.

36. Faca V, Pitteri SJ, Newcomb L, Glukhova V, Phanstiel D, Krasnoselsky A, Zhang Q, Struthers J, Wang H, Eng J, Fitzgibbon M, McIntosh M, Hanash S: **Contribution of protein fractionation to depth of analysis of the serum and plasma proteomes.** *J Proteome Res* 2007, **6**:3558-3565.

37. Shen Y, Zhao R, Berger SJ, Anderson GA, Rodriguez N, Smith RD: **High-efficiency nanoscale liquid chromatography coupled on-line with mass spectrometry using nanoelectrospray ionization for proteomics.** *Anal Chem* 2002 **74**:4235-4249.

38. Dowell JA, Frost DC, Zhang J, Li L: **Comparison of two-dimensional fractionation techniques for shotgun proteomics.** *Anal Chem* 2008 **80**:6715-6723.

39. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS: **Probability-based protein identification by searching sequence databases using mass spectrometry data.** *Electrophoresis* 1999, **20**:3551-3567.

40. Yates JR 3rd, Eng JK, McCormack AL, Schieltz D: **Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database.** *Anal Chem* 1995, **67**:1426-1436.

41. Beavis RC, Fenyo D: **A method for assessing the statistical significance of mass spectrometry-based protein identifications using general scoring schemes.** *Anal Chem* 2003, **75**:768-774.

42. Nesvizhskii AI: **A survey of computational methods and error rate estimation procedures for peptide and protein identification in shotgun proteomics.** *J Proteomics* 2010, **73**:2092-2123.

43. Gupta N, Bandeira N, Keich U, Pevzner PA: **Target-decoy approach and false discovery rate: when things may go wrong.** *J Am Soc Mass Spectrom* 2011, **22**:1111-1120.

44. Kim S, Gupta N, Pevzner PA: **Spectral probabilities and generating functions of tandem mass spectra: a strike against decoy databases.** *J Proteome Res* 2008, **7**:3354-3363.

45. Xie F, Liu T, Qian WJ, Petyuk VA, Smith RD: **Liquid chromatography-mass spectrometry-based quantitative proteomics.** *J Biol Chem* 2011, **286**:25443-25449.

46. Goshe MB, Smith RD: **Stable isotope-coded proteomic mass spectrometry.** *Curr Opin Biotechnol* 2003, **14**:101-109.

47. Geiger T, Cox J, Ostasiewicz P, Wisniewski JR, Mann M: **Super-SILAC mix for quantitative proteomics of human tumor tissue.** *Nat Methods* 2010, **7**:383-385.

48. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ: **Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents.** *Mol Cell Proteomics* 2004, **3**:1154-1169.

49. Zieske LR: **A perspective on the use of iTRAQ reagent technology for protein complex and profiling studies.** *J Exp Bot* 2006, **57**:1501-1508.

50. Qian WJ, Liu T, Petyuk VA, Gritsenko MA, Petritis BO, Polpitiya AD, Kaushal A, Xiao W, Finnerty CC, Jeschke MG, Jaitly N, Monroe ME, Moore RJ, Moldawer LL, Davis RW, Tompkins RG, Herndon DN, Camp DG, Smith RD; Inflammation and the Host Response to Injury Large Scale Collaborative Research Program: **Large-scale multiplexed quantitative discovery proteomics enabled by the use of an (18)O-labeled "universal" reference sample.** *J Proteome Res* 2009, **8**:290-299.

51. Nielsen ML, Vermeulen M, Bonaldi T, Cox J, Moroder L, Mann M: **Iodoacetamide-induced artifact mimics ubiquitination in mass spectrometry.** *Nat Methods* 2008, **5**:459-460.

52. Zimmer JS, Monroe ME, Qian WJ, Smith RD: **Advances in proteomics data analysis and display using an accurate mass and tim e tag approach.** *Mass Spectrom Rev* 2006, **25**:450-482.

53. Scicchitano MS, Dalmas DA, Boyce RW, Thomas HC, Frazier KS: **Protein extraction of formalin-fixed, paraffin-embedded tissue enables robust proteomic profiles by mass spectrometry.** *J Histochem Cytochem* 2009, **57**:849-860.

54. Waldemarson S, Krogh M, Alaiya A, Kirik U, Schedvins K, Auer G, Hansson KM, Ossola R, Aebersold R, Lee H, Malmström J, James P: **Protein expression changes in ovarian cancer during the transition from benign to malignant.** *J Proteome Res* 2012, **11**:2876-2889.

55. Angel TE, Jacobs JM, Spudich SS, Gritsenko MA, Fuchs D, Liegler T, Zetterberg H, Camp DG 2nd, Price RW, Smith RD: **The cerebrospinal fluid proteome in HIV infection: change associated with disease severity.** *Clin Proteomics* 2012, **9**:3.

56. Gallien S, Duriez E, Domon B: **Selected reaction monitoring applied to proteomics.** *J Mass Spectrom* 2011, **46**:298-312.

57. Stahl-Zeng J, Lange V, Ossola R, Eckhardt K, Krek W, Aebersold R, Domon B: **High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites.** *Mol Cell Proteomics* 2007, **6**:1809-1817.

58. Anderson NL, Anderson NG, Haines LR, Hardie DB, Olafson RW, Pearson TW: **Mass spectrometric quantitation of peptides and proteins using stable isotope standards and capture by anti-peptide antibodies (SISCAPA).** *J Proteome Res* 2004 **3**:235-244.

59. Whiteaker JR, Zhao L, Anderson L, Paulovich AG: **An automated and multiplexed method for high throughput peptide immunoaffinity enrichment and multiple reaction monitoring mass spectrometry-based quantification of protein biomarkers.** *Mol Cell Proteomics* 2010, **9**:184-196.

60. Hossain M, Kaleta DT, Robinson EW, Liu T, Zhao R, Page JS, Kelly RT, Moore RJ, Tang K, Camp DG 2nd, Qian WJ, Smith RD: **Enhanced sensitivity for selected reaction monitoring mass spectrometry-based targeted proteomics using a dual stage electrodynamic ion funnel interface.** *Mol Cell Proteomics* 2011, **10**:M000062-MCP201.

61. Picotti P, Lam H, Campbell D, Deutsch EW, Mirzaei H, Ranish J, Domon B, Aebersold R: **A database of mass spectrometric assays for the yeast proteome.** *Nat Methods* 2008, **5**:913-914.

62. SRM Atlas [http://www.srmatlas.org/]

63. Lange V, Picotti P, Domon B, Aebersold R: **Selected reaction monitoring for quantitative proteomics: a tutorial.** *Mol Syst Biol* 2008, **4**:222.

64. Beavis RC, Craig R, Cortens JP: **Open source system for analyzing, validating, and storing protein identification data.** *J Proteome Res* 2004, **3**:1234-1242.

65. Fusaro VA, Mani DR, Mesirov JP, Carr SA: **Prediction of high-responding peptides for targeted protein assays by mass spectrometry.** *Nat Biotechnol* 2009, **27**:190-198.

66. Addona TA, Abbatiello SE, Schilling B, Skates SJ, Mani DR, Bunk DM, Spiegelman CH, Zimmerman LJ, Ham AJ, Keshishian H, Hall SC, Allen S, Blackman RK, Borchers CH, Buck C, Cardasis HL, Cusack MP, Dodder NG, Gibson BW, Held JM, Hiltke T, Jackson A, Johansen EB, Kinsinger CR, Li J, Mesri M, Neubert TA, Niles RK, Pulsipher TC, Ransohoff D, *et al.*: **Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma.** *Nat Biotechnol* 2009, **27**:633-641.

67. Suhr H: *Plasma Chromatography*. New York: Plenum Press; 1984.

68. Mason E, McDaniel E: *Transport Properites of Ions in Gases.* New York: Wiley; 1988.

69. Guevremont R, Siu KW, Wang J, Ding L: **Combined ion mobility/time-of-flight mass spectrometry study of electrospray-generated ions.** *Anal Chem* 1997, **69**:3959-3965.

70. Baker ES, Tang K, Danielson WF 3rd, Prior DC, Smith RD: S**imultaneous fragmentation of multiple ions using IMS drift time dependent collision energies.** *J Am Soc Mass Spectrom* 2008, **19**:411-419.

71. Merenbloom SI, Koeniger SL, Valentine SJ, Plasencia MD, Clemmer DE: **IMS-IMS and IMS-IMS-IMS/MS for separating peptide and protein fragment ions.** *Anal Chem* 2006, **78**:2802-2809.

72. Pringle SD, Giles K, Wildgoose JL, Williams JP, Slade SE, Thalassinos K, Bateman RH, Bowers MT, Scrivens JH: **An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument.** *Int J Mass Spectrom* 2007, **261**:1-12.

73. Baker ES, Livesay EA, Orton DJ, Moore RJ, Danielson WF 3rd, Prior DC, Ibrahim YM, LaMarche BL, Mayampurath AM, Schepmoes AA, Hopkins DF, Tang K, Smith RD, Belov ME: **An LC-IMS-MS platform providing increased dynamic range for high-throughput proteomic studies.** *J Proteome Res* 2010, **9**:997-1006.

74. Tang K, Shvartsburg AA, Lee HN, Prior DC, Buschbach MA, Li F, Tolmachev AV, Anderson GA, Smith RD: **High-sensitivity ion mobility spectrometry/mass spectrometry using electrodynamic ion funnel interfaces.** *Anal Chem* 2005, **77**:3330-3339.

75. Shaffer SA, Tang K, Anderson GA, Prior DC, Udseth HR, Smith RD: **A novel ion**

funnel for focusing ions at elevated pressure using electrospray ionization mass spectrometry. *Rapid Commun Mass Spectrom* 1997, **11**:1813-1817.

76. Lindner M, Hoffmann GF, Matern D: **Newborn screening for disorders of fatty-acid oxidation: experience and recommendations from an expert meeting.** *J Inherit Metab Dis* 2010, **33**:521-526.

77. Smith RD: **Mass spectrometry in biomarker applications: from untargeted discovery to targeted verification, and implications for platform convergence and clinical application.** *Clin Chem* 2012, **58**:528-530.

78. Surinova S, Schiess R, Hüttenhain R, Cerciello F, Wollscheid B, Aebersold R: **On the development of plasma protein biomarkers.** *J Proteome Res* 2011, **10**:5-16.