The Development of a Proteomic Wellness Assay Using Dried Blood Spots: Moving Clinical Protein Diagnostics Towards Personalized Reference Ranges

James G. Bollinger[1], Clark M. Henderson[2], Christine C. Wu[3], Andrew N. Hoofnagle[2], Michael J. MacCoss[1]

[1]Department of Genome Sciences, University of Washington, Seattle, WA 98195

[2]Departments of Laboratory Medicine and Medicine, University of Washington, Seattle, WA 98195

[3]Stratus Biosciences, Seattle, WA 98117

Introduction

In clinical diagnostics, dried blood spot (DBS) samples have several advantages over their serum and plasma counterparts. These include a simpler and less invasive sampling procedure, lower biohazard risk associated with specimen handling, and improved analyte stability under ambient storage conditions. Despite the success of DBS/tandem MS protocols in newborn screening labs and in small-molecule drug development, there are relatively few reports of their use in protein diagnostics. Here we report a method for the systematic development of targeted proteomic assay within the DBS sample matrix. The proteins targeted in the current study are each of moderate-to-high abundance in the DBS matrix and are comprised of lipoproteins, acute phase response proteins, members of the complement cascade, liver enzymes, proteins involved in lipid metabolism, among others. Most of the proteins studied already have diagnostic assays used in the clinical laboratory as individual immunoassays or enzyme activity assays. For each assay target, we use an analytical standard to empirically derive an optimal set of peptide proxies. We then go on to demonstrate the value of our approach by creating and validating a selected reaction monitoring (SRM) assay for the multiplexed quantitation of a set of 50 proteins within DBS samples using a single protein as an internal standard.

Methods

We used a human cDNA library obtained from DNASU to express full-length proteins via *in vitro* translation (IVT). For those protein targets not in our library, we either cloned commercial cDNA constructs into an expression vector compatible with the IVT technology or, when possible, purchased the recombinant protein. Each analytical standard was subjected to a standard proteomic

work-up and submitted individually for LC-MS/MS analysis on a TSQ Quantiva triple-quadrupole instrument coupled with an Easy-nLC 1000 nano-flow chromatography system (Thermo). We then used the open-source application Skyline to extract fragment ion chromatograms for each candidate peptide of each target protein. All tryptic peptides from 7 to 25 amino acids in length are considered in their monoisotopic (+2) charge state with all cysteines monitored as carbamidomethyl cysteines. For tandem MS analysis, we measured each y_3 to y_{n-1} fragment ion. Relative peptide intensities were catalogued along with each peptide's MS/MS fragmentation fingerprint (relative distribution of y-ion intensities). Each digest is then spiked with a set of retention time standards, incubated in the autosampler at 4^{o} C for 48-72 hours, and then submitted for second round of LC-MS/MS analysis. This process enabled us to further characterize each candidate peptide's post-digestion stability and relative hydrophobicity. The use of full-length protein analytical standards permits the assembly of high-quality chromatogram libraries that can be subsequently used to screen the DBS matrix for peptides and MS/MS transitions with optimal signal-to-noise ratios. Using our empirically derived set of observable peptides and their respective MS/MS transitions, we then designed and validated an SRM assay for the multiplexed analysis of clinical protein targets in reconstituted DBS digests. Importantly, our methodology is devoid of any antibody-based enrichment steps, makes use of a single-point calibration method to minimize day-to-day variation, and uses a single protein as a global internal standard (15N-Apolipoprotein A1, APOA1).

Preliminary Results

We expressed or purchased 85 target proteins to be used as analytical standards. These standards were applied in the assembly of a high-quality spectral library and for the empirical derivation of an optimal set of peptides and fragments for use in our protein wellness assay. After screening a standard set of DBS samples, our protein target list was trimmed to 50. The filtered library of optimal peptides observed in the DBS matrix was subsequently used to design a multiplexed SRM assay on a TSQ Quantiva triple quadrupole mass spectrometer. (Thermo). The DBS utilized were obtained from a coordinated donor blood draw that provided matching whole blood and plasma samples. This dataset enabled cross-validation for each assay target via nephelometric immunoassay. We assessed the variability associated with each step of our sample preparation and derived coefficients of variation for a set biological replicates. Furthermore, assay linearity

was evaluated via a variation of the method of standard addition and lower limits of quantitation were estimated by diluting human DBS digests into whole blood from *Gallus gallus*.

Novel Aspect

The identification of an optimal set of peptides for each protein of interest is a crucial step in the development of targeted proteomic assays. Considerable amounts of time and resources are often spent to produce quantitative standards such as synthetic peptides, recombinant proteins from concatenated peptide sequences, or developing immunoaffinity reagents for the enrichment of low abundance tryptic peptides We define here an optimal or 'proteotypic' peptide for a targeted proteomic experiment is a peptide that is; unique to a given gene product, lacks common single nucleotide polymorphisms that result in an amino acid substitutions, is devoid of known post-translational modification sites, has physiochemical properties amenable to a robust detection in the mass spectrometer, and has salient features that generate characteristic MS/MS fragmentation patterns via collision-induced dissociation (CID).Traditional approaches for selecting peptides have relied either on the mining of DDA spectral libraries or the use of prediction algorithms trained on previous DDA experimental results. Both approaches are predicated on the assumption that the peptides most frequently identified in DDA experiments will produce peptides with the optimal signal-to-noise ratios for a targeted proteomic experiment. Unfortunately, this is not the case. There are numerous reasons why a peptide may not be selected for MS/MS during a DDA experiment. Thus, a peptide that is not observed in this type of experiment should not be excluded in a targeted experiment. Conversely, a peptide that is routinely sampled in a DDA style experiment might not necessarily be a suitable peptide for an SRM experiment. For these reasons and more, we advocate that an important initial step in the development of targeted proteomic assays should be the application of analytical standards and the empirical measurement of all peptides from a target protein that fulfill some minimal criteria directly within one's sample matrix of interest.