# Systematic Assessment of Survey Scan and MS2-Based Abundance Strategies for Label-Free Quantitative Proteomics Using High-Resolution MS Data

Chengjian Tu,*[†,‡] Jun Li,[†,‡] Quanhu Sheng,[§] Ming Zhang,[†,‡] and Jun Qu*[†,‡]

[†]Department of Pharmaceutical Sciences, University at Buffalo, State University of New York, Buffalo, NY 14260, United States
[‡]New York State Center of Excellence in Bioinformatics and Life Sciences, 701 Ellicott Street, Buffalo, NY 14203, United States
[§]Center for Quantitative Sciences, Vanderbilt University, Nashville, TN 37232, United States

S *Supporting Information*

ABSTRACT: Survey-scan-based label-free method have shown no compelling benefit over fragment ion (MS2)-based approaches when low-resolution mass spectrometry (MS) was used, the growing prevalence of high-resolution analyzers may have changed the game. This necessitates an updated, comparative investigation of these approaches for data acquired by high-resolution MS. Here, we compared survey scan-based (ion current, IC) and MS2-based abundance features including spectral-count (SpC) and MS2 total-ion-current (MS2-TIC), for quantitative analysis using various high-resolution LC/MS data sets. Key discoveries include: (i) study with seven different biological data sets revealed only IC achieved high reproducibility for lower-abundance proteins; (ii) evaluation with 5-replicate analyses of a yeast sample showed IC provided much higher quantitative precision and lower missing data; (iii) IC, SpC, and MS2-TIC all showed good quantitative linearity ($R^2 > 0.99$) over a >1000-fold concentration range; (iv) both MS2-TIC and IC showed good linear response to various protein loading amounts but not SpC; (v) quantification using a well-characterized CPTAC data set showed that IC exhibited markedly higher quantitative accuracy, higher sensitivity, and lower false-positives/false-negatives than both SpC and MS2-TIC. Therefore, IC achieved an overall superior performance than the MS2-based strategies in terms of reproducibility, missing data, quantitative dynamic range, quantitative accuracy, and biomarker discovery.

KEYWORDS: ion current based quantification, label-free quantitative proteomics, missing data, MS/MS total ion current, spectral count
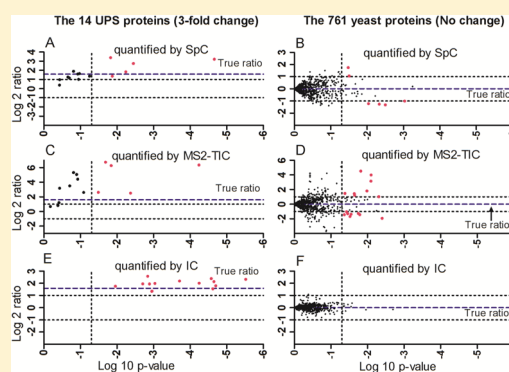
INTRODUCTION

Accurate and precise quantitative strategy is critical for reliable proteomic expression profiling and discovery of biomarker candidates. Roughly, LC/MS-based relative quantification methods can be divided into two main categories: isotope labeling and label-free methods. Stable isotope labeling approaches play a prominent role in quantitative proteomics. Since the introduction of the isotope-coded affinity tag (ICAT) in 1999,[1] a variety of chemical- or metabolic-labeling methods have been developed, such as the ^{18}O-labeling,[2,3] stable isotope labeling by amino acids in cell culture (SILAC),[4] isobaric tags for relative and absolute quantification (iTRAQ),[5] tandem mass tags (TMT),[6,7] super-SILAC,[8] and more recently, neutron-encoded mass signatures (NeuCode).[9] Although most of these strategies have been tremendously successful and widely applied in proteomics profiling, certain drawbacks do exist, such as the high expense of reagents that renders the techniques cost-prohibitive for large-scale studies and that the efficiency and consistency of labeling may not be perfect for some methods and in some cases, complex data interpretation.[10,11]

Label-free approaches have emerged as an attractive alternative to isotope-labeling methods, due to their simplicity, cost-effectiveness, and feasibility of quantifying multiple biological samples.[10,12] These approaches consist of two conceptually different types, which employ quantitative features either derived from MS2 product ion scans[13,14] or peptide precursor signals (MS1) obtained by the survey scan[15−17] to measure relative protein abundances in proteolytic digests. One of the classical MS2-based methods, termed spectral counting, estimates protein abundance by counting the total number of MS/MS spectra matched to all peptides from a given protein. This approach was recently improved by incorporating the MS2 fragment ion intensities and unique peptide number for quantitative analysis, for example, the normalized spectral index (SI_N);[18] other examples in this avenue include exponentially modified protein abundance index (emPAI)[19] and the Normalized spectral abundance factor (NSAF).[20] Nonetheless, the MS2-based approaches are challenged by the nature of

current MS/MS sampling techniques such as the data-dependent MS2 fragmentation. First, dynamic exclusion of the precursors fragmented in a previous scan, a widely practiced technique to improve the chance of detecting low-abundance peptides, significantly affects spectral acquisition; second, the MS2 acquisition for low-abundance peptides are often suppressed by coeluting peptides of higher abundance; finally, the accurate quantitative information for lower-abundance proteins/peptides, (e.g., these resulting in spectral counts of "1" and "0", a very common sight in LC-MS data), is often elusive.[10,21,22] By comparison, the survey scan-based (or ion current-based) approach quantifies proteins by measuring the extracted ion current peak areas of peptide precursors (MS1) from each protein. The calculation of peak areas is independent to MS2 acquisition, and consequently, the above-mentioned problems associated with the MS2 sampling processes are either avoided or greatly alleviated. An additional salient advantage of ion current-based method is that as long as well-defined ion current peaks are observed and aligned properly across all samples, the corresponding peptide can be quantified without missing data (missing abundance values in one or more replicates) even if it was only successfully identified for once in the entire sample set.[23,24] Nevertheless, carrying out ion current-based quantification is generally more technically demanding than MS2-based approaches, owing to the require-ment of accurate matching and quantification of precursor peaks among all samples, which in turn requires specific and accurate MS detection (i.e., the use of a high-resolution MS analyzer), as well as highly reproducible sample preparation and chromatographic separation.[10,16]

In the recent several years, the rapid-growing availability of high-resolution MS analyzers such as the new generation of time-of-flight, Fourier transform ion cyclotron resonance, and Orbitrap may have favored the application of ion current-based approaches in proteomic studies.[16,25−27] The use of high-resolution analyzers permits extraction of peptide ion currents within a very narrow m/z range (e.g., <0.02 mass unit) to substantially reduce chemical noises and interferences, and therefore, greatly improves sensitivity and specificity of ion current-based quantification.[28,29] Meanwhile, the MS2 total ion current (or MS/MS fragment ion intensities, MS2-TIC) approach was introduced, a technique which utilizes the sum of the product ion intensities in each MS2 spectrum assigned to a given protein as the quantitative feature.[18,30] More recently, on the basis of high-resolution MS data, researchers have described that using MS2 intensities resulted in protein abundance measurements nearly as accurately as with MS1-intensities,[31] and combing precursor intensities with spectral counts, these researchers identified more true positives.[32] However, our previous works found that when using data from high-resolution MS, the MS1-based approach dramatically improved the quantitative accuracy compared to MS2-based methods, especially for the quantification of low-abundance proteins.[24,33] Given the above-mentioned developments in both MS1-based and MS2-based approaches, it would be of high value to perform an updated, comprehensive comparison of the quantitative performance of the MS1-based method versus MS2-based methods using high-resolution MS data. Such a comparison would greatly help us to understand the limitation and capacity of each approach and is highly valuable for the development of label-free quantification strategy. However, to our knowledge, a systematic and comprehensive comparison has not been conducted before this study.

Here, we assessed the MS1-based ion current-based method (IC) along with two popular MS2-based methods, including spectral count (SpC) and MS2 total ion current (MS2-TIC),[18] using data sets generated by LTQ/Orbitrap MS. Quantitative metrics including reproducibility, precision, accuracy, missing data, dynamic linear range, and sensitivity/specificity for discovery of significantly altered proteins were thoroughly evaluated.

■ MATERIALS AND METHODS

Protein Sample Preparation

The human bronchoalveolar lavage fluids, rat brain, rat liver, and rat retina were from Buffalo General Medical Center (Buffalo, NY). The human skeletal muscle cells, *E. coli* cells, and yeast cells were from Kinex Pharmaceuticals (Buffalo, NY). Cell or tissue samples used in this study were homogenized in an ice-cold lysis buffer (50 mM Tris-formic acid, 150 mM NaCl, 0.5% sodium deoxycholate, 2% SDS, 2% NP-40, pH 8.0) using a Polytron homogenizer (Kinematica AG, Switzerland). Homogenization was performed for a 5−10 s burst at 15 000 rpm, followed by a 20 s cooling period until the foam settled. This procedure was repeated 10 times. The mixture was then sonicated in a cold room for ∼10 min with a low-power sonicator until the solution was clear, followed by centrifuga-tion at 140 000g for 1 h at 4 °C. The supernatant was carefully transferred to a fresh tube, and the protein concentrations were measured using BCA Protein Assay (Pierce, Rockford, IL). The resulted samples were stored at −80 °C until analysis. In order to remove undesirable components in the samples while maintaining high peptide recovery, a precipitation/on-pellet-digestion protocol was employed as previously de-scribed.[24,33−35] The precipitation/on-pellet-digestion procedure was directly performed without protein extraction when processing the human bronchoalveolar lavage fluid sample. Specimens (each containing 100 μg of total protein) were reduced with TCEP (3 mM) for 10 min and then alkylated with 20 mM IAM for 30 min in darkness. The mixture was precipitated by stepwise addition of 9 volumes of cold acetone with continuous vortexing and then incubated overnight at −20 °C. After centrifugation at 12 000g for 20 min at 4 °C, the supernatant was removed, and the pellet was allowed to air-dry. Two digestion phases were employed for the on-pellet digestion. In phase 1 (pellet-dissolving phase), 50 μL of trypsin solution at an enzyme/substrate ratio of 1:30 (w/w) was added and incubated at 37 °C for 6 h with agitation; then in phase 2 (complete-cleavage phase), another 50 μL of trypsin solution was added at an enzyme/substrate ratio of 1:25 (w/w), and the mixture was incubated overnight to achieve complete digestion.

NanoLC-MS/MS Analysis

The nano-RPLC (reverse-phase liquid chromatography) system consisted of a Spark Endurance autosampler (Emmen, Holland) and an ultrahigh pressure Eksigent (Dublin, CA) nano-2D Ultra capillary/nano-LC system. Mobile phase A and B were 0.1% formic acid in 2% acetonitrile and 0.1% formic acid in 88% acetonitrile, respectively. Four microliters of sample was loaded onto a reversed-phase trap (300 μm I.D. x1 cm) unless otherwise noted in the paper, with 1% mobile phase B at a flow rate of 10 μL/min, and the trap was washed for 3 min. A series of nanoflow gradients (flow rate, 250 nL/min) was used to back-flush the trapped samples onto the nano-LC column (75 μm i.d. × 75 cm) for separation. The nano-LC column was heated at 52 °C to greatly improve both chromatographic

resolution and reproducibility. An LTQ/Orbitrap XL hybrid mass spectrometer (Thermo Fisher Scientific, San Jose, CA) was used for protein identification. The parameters for MS are shown in our previous publications.[24,33−35]

In this study, two technical replicates of each of the seven biological samples (human bronchoalveolar lavage fluid, human skeletal muscle cells, rat brain, rat liver, rat retina, and *E. coli* cells) and five replicates of *S. cerevisae* (yeast) cell sample were analyzed in order to assess quantitative reproducibility and missing data by the three label-free approaches. To investigate the correlation between quantitative values given by the three abundance features and protein abundances in a complex proteome, an *E. coli* extract was spiked with bovine serum albumin (BSA) at six different levels (0.025, 0.1, 0.5, 2.5, 12.5, and 62.5% of total proteins) and analyzed in triplicate. The evaluation of the correlation among the quantitative values given by SpC, MS2-TIC, and IC with different amounts of sample loading, a pooled digest of a prostate cancer cell line (PC3-LN4) sample was loaded at protein amounts of 0.5, 1, 2, and 4 μg.

In addition, to assess the performance of biomarker discovery by SpC, MS2-TIC and IC, we employed the "Study 6 LTQ Orbitrap XL @P65" data set generated by the program of Clinical Proteomic Technology Assessment for Cancer (CPTAC).[34,36] According to the publicly available documentation associated with this study, the Universal Proteomics Standard set 1 (UPS1, a 48-protein equimolar standard) was spiked at amounts of 0.25, 0.74, and 2.2 fmol/μL into yeast lysate for sets A, B and C, the subset of studies investigated in the current work. Each sample was analyzed by nano-LC/MS with an Orbitrap XL analyzer in triplicate.

Database Search and Data Validation

The raw data files were searched against the Swiss-Prot protein database (version 06/13/2012) using the Sequest algorithm embedded in Proteome Discoverer 1.2 (Thermo-Scientific). A total of 7766 protein entries, 20 238 entries, 4431 entries, and 7801 entries were presented in respective rat, human, E.coli, and yeast protein database. The databases were augmented with the sequences of bovine serum albumin and the UPS1 48 proteins (Sigma-Aldrich) when appropriate. The search parameters used were as follows: 10 ppm tolerance for precursor ion masses and 1.0 Da for fragment ion masses. Two missed cleavages were permitted for fully tryptic peptides. Carbamidomethylation of cysteines was set as a fixed modification, and a variable modification of methionine oxidation was allowed. The false discovery rate (FDR) was determined by using a target-decoy search strategy.[37] The sequence database contains each sequence in both forward and reversed orientations, enabling FDR estimation. Scaffold 3.6[38] (Proteome Software, Portland, OR), which is capable of handling large-scale proteomic data sets, was used to validate MS2-based peptide and protein identification based on cutoffs of cross-correlation (Xcorr) and Delta Cn values. The peptide FDR was controlled at 0.1%. Validated peptides were grouped into individual protein clusters by Scaffold software.

Protein Quantification

The protein quantitative values based on SpC and MS2-TIC were obtained using Scaffold 3.6[38] (Proteome Software, Portland, OR) under the same protein/peptide identification criteria as described above. The quantitative analysis by IC was performed by two steps: procurement of area-under-the-curve (AUC) data for peptides using SIEVE v2.0 (Thermo Scientific,

San Jose, CA) and then a sum-intensity method to aggregate the quantitative data from peptide level to protein level as previously reported.[24] SIEVE is a label-free differential expression package that performs chromatographic alignment and global intensity-based MS1 feature extraction.[39] The package processes chromatographic alignment among sequential LC/MS runs using the ChromAlign algorithm.[40] Quantitative "frames" were defined based on m/z (width: 10 ppm) and retention time (width: 2.5 min) of peptide precursors in the aligned runs. Peptide ion current areas were calculated for individual replicates in each frame. Subsequent to ion current values extraction, MS2 fragmentation scans associated with each frame were identified by importing the msf files created by Proteome Discoverer (cf. the database search and data validation procedure). Peptides shared among different protein groups were excluded from quantitative analysis.

For SpC, MS2-TIC, and IC, relative quantification of protein levels were based on the sum of respective abundance values of all peptides assigned to each protein, without any statistical outlier analysis. Normalization was performed against total abundance values in individual runs. In case of missing data, baseline quantitative values were assigned (e.g., a 0.5 and 1000, respectively, for SpC and MS2-TIC at the protein level and 1000 for IC at peptide level). The value of 0.5 counts for spectrum counting, and 1000 for MS2-TIC and ion current (IC) were experimentally determined (Supplemental Figure 1). Statistical significance between groups (comparing case vs control samples) was evaluated using a Student's *t* test, with a *p*-value cutoff of 0.05. The relative protein ratio of a protein between the groups was calculated by comparing the average abundance values of the protein in each group. Abundance change > 2-fold and *p*-value < 0.05 were used as the thresholds to define altered proteins.

■ RESULTS AND DISSCUSION

Label-free approaches play a prominent role for relative proteomic quantification and biomarker discovery. To date, MS2-based methods have been the most common type of label-free approaches, especially for data generated by lower resolution MS.[10,31] This is largely due to the relatively poor specificity of low-resolution MS, which leads to difficulties in precise and selective measurement/match of peptide precursor ion currents among multiple runs of complex proteomic samples. Owing to the drastically increased availability of high-resolution instruments and the technical advances in hybrid instruments that have tremendously improved the robustness, throughput and sensitivity of high-resolution MS,[41−43] the application of the ion-current-based approach has been rapidly growing in the most recent years.[16,25,26] A comprehensive and *updated* comparison of these label-free methods based on the data generated by the high-resolution analyzers is highly valuable but remains to be conducted. To address this need, here we evaluated the ion current-based and several MS2-based approaches for quantitative reproducibility and accuracy, missing values, dynamic linear range, and performance for discovery of significantly altered proteins, using various high-resolution-MS data sets on complex proteomes that are either generated by our lab or publicly available.
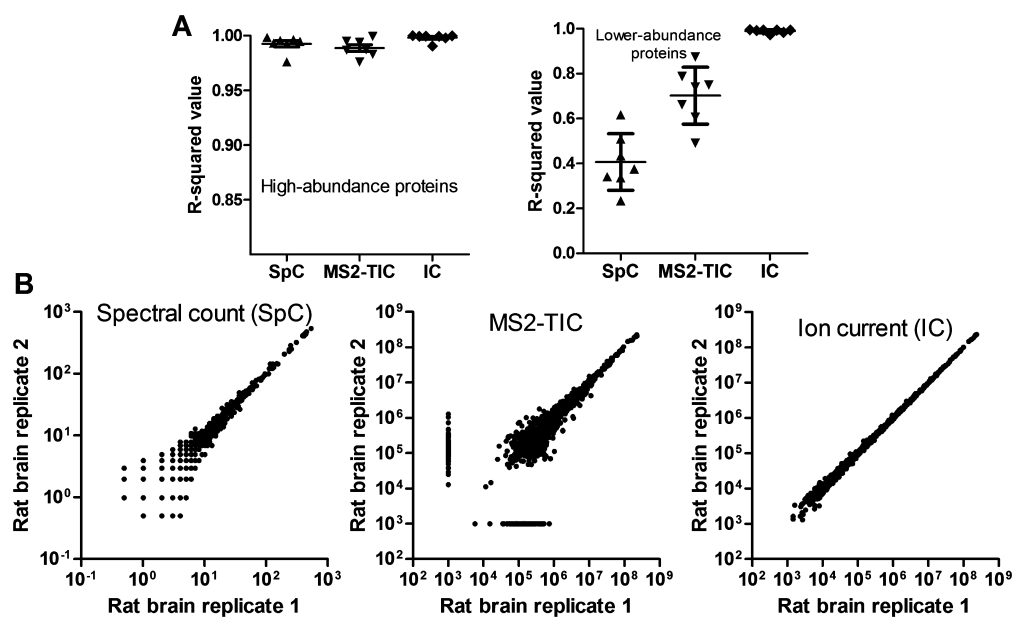
Figure 1. Quantitative reproducibility of the three label-free methods. (A) Comparison of the coefficient-of-determination (R^2) of the linear regression by three methods including spectral count (SpC), MS/MS total ion current (MS2-TIC), and ion current (IC). Data of duplicate LC/MS runs of seven types of proteomic samples (human bronchoalveolar lavage fluid, human skeletal muscle cells, rat brain, rat liver, rat retina, *E. coli* cells, and yeast cells) were analyzed, and each data point represents the R^2 of one of the proteome samples. The high-abundance proteins refer to the top 33% of all proteins ranked by spectral count, and the rest are designated as lower-abundance proteins. (B) Representative scatter plots of duplicate LC-MS/MS analyses by SpC, MS2-TIC, and IC. The two axes represent the quantitative abundance values of the same proteins, respectively, by the two duplicate runs.
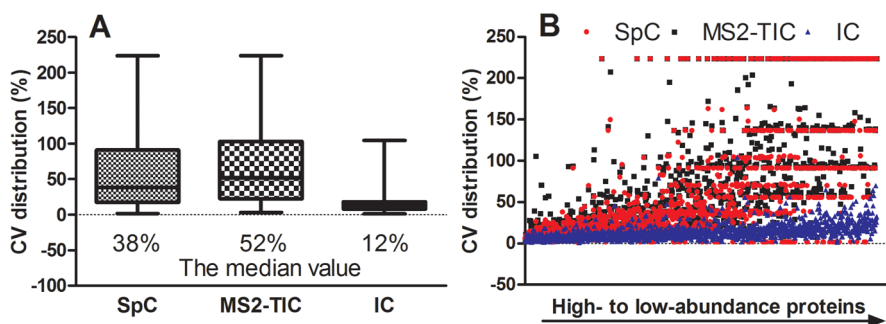

Figure 2. Coefficients of variation (CV) of the abundance values of the 1196 quantified yeast proteins by SpC, MS2-TIC, and IC ($N = 5$ LC-MS analyses). (A) Box-and-whisker plot analysis was employed to show the spread of protein CVs around the median value (the horizontal line inside the box); bottom and top of the boxes correspond to the top 25th and 75th percentile of the CV distribution and whiskers to the minimum and maximum values. (B) The distribution of CV vs protein abundance. Red circles indicate SpC, black squares indicate MS2-TIC, and blue triangles indicate IC data spot.

Quantitative Reproducibility by the Three Label-Free Approaches

Good quantitative reproducibility is indispensable for accurate and precise proteomic quantification and reliable biomarker discovery. Here we evaluated the reproducibility of the three approaches (SpC, MS2-TIC, and IC), by correlating the quantitative results between duplicate LC-MS analyses of proteomic samples from seven different biological sources, including human bronchoalveolar lavage fluid, human skeletal muscle cells, rat brain, rat liver, rat retina, *E. coli* cells, and yeast cells. These samples represent a wide variety of biological matrices seen in typical proteomic investigations. Among the three approaches that are based on different abundance features, only IC is based on survey-scan (MS1). In case of missing data in any of the methods (i.e., the quantitative value of a protein is not measured in one or more replicates) a zero value was assigned to the affected replicate. The normalization was performed against the sum of all individual abundance values in the same replicate. In order to obtain a reliable comparison, a set of strict cutoff criteria for protein identification and validation were employed, resulting in a peptide FDR of 0.1% in individual data set (as determined by the target-decoy database searching strategy, see Materials and Methods).

Linear regression of the correlation between the duplicate runs was performed for each of the seven types of proteomic samples. The R-squared values for SpC, MS2-TIC, and IC are 0.993 ± 0.007, 0.990 ± 0.007, and 0.998 ± 0.003, respectively. The good reproducibility achieved by these label-free approaches is in line with previous reports.[16,18,44,45] To further assess whether such correlations are abundance-dependent, we conducted comparison by separating the quantified proteins into two groups: high-abundance proteins (the top 33% abundant proteins as determined by spectral count) and

lower-abundance proteins (the remaining 67% proteins). For high-abundance proteins, the R^2 values for SpC, MS2-TIC, and IC are 0.992 ± 0.008, 0.989 ± 0.008, and 0.998 ± 0.003, respectively; for lower-abundance proteins, the R^2 values for SpC, MS2-TIC, and IC are 0.407 ± 0.126, 0.702 ± 0.127, and 0.990 ± 0.008, respectively (Figure 1A). For lower-abundance proteins, only IC achieved a high quantitative reproducibility, which can also be visualized in Figure 1B, which shows representative scatter plots between two replicate runs for SpC, MS2-TIC, and IC. A substantially higher degree of reproducibility for quantifying lower-abundance proteins was observed for IC over SpC and MS2-TIC. All protein abundance values by these three approaches in the paired LC-MS runs of seven different biological samples are in the Supplemental Table I.

Assessment of Quantitative Precision and the Level of Missing Data

The precision of SpC, MS2-TIC, and IC was evaluated by measuring coefficients of variation (CV) for the quantification of individual proteins, using five LC-MS runs (technical replicates, $N = 5$) of a yeast digest. It was observed that the distributions of CV for the 1196 quantified proteins are quite different by the three label-free methods. Figure 2A shows the box-and-whisker plot of these distributions, where the bottom and the top of the boxes, respectively, correspond to the top 25th and 75th percentile values of the CV distribution, the horizontal lines inside the box to the median CV values, and whiskers to the minimum and maximum values. The median CV values for quantification of individual proteins are 38%, 52%, and 12%, respectively for SpC, MS2-TIC, and IC. Figure 2B shows the distribution of CV versus relative protein abundance. While more than 99% of proteins have CV < 50% using IC approach, only 56.5% and 49.6% of the proteins are under this threshold, respectively, for SpC and MS2-TIC. Furthermore, for all three methods, lower quantitative precision (i.e., higher CV) was observed for low-abundance proteins compared to the high-abundance ones. Among them, IC achieved much lower CV for low-abundance proteins than the other two methods, which is in agreement with above observation that IC enabled more reproducible quantification for lower-abundance proteins (cf. Figure 1B). This result suggests that IC may be much more reliable than SpC and MS2-TIC for quantifying lower-abundance proteins. Supports of this notion can also be found in previous observations that spectral-count-based approaches yielded unreliable quantitative values for low-abundance peptides/proteins, by various laboratories including ours.[10,24,46,47]

Owing to the high complexity and wide dynamic ranges of typical proteomes and certain technical issues such as the sampling nature of data-dependent MS2 analysis, missing data (i.e., missing quantitative values in one or more replicates) is a prevalent challenge in quantitative proteomics, which may severely undermine the reliability of quantification and biomarker discovery.[48,49] Missing values may arise from technical and/or biological sources,[48–50] and here we evaluated the levels of missing values exclusively from technical aspects. The five replicate runs of yeast sample were utilized to assess the frequency of missing abundance values at protein level. For both SpC and MS2-TIC, the frequency of missing quantitative values equals to the frequency of missing identifications of proteins. The missing values by SpC/MS2-TIC are summarized in Table 1, which shows 13.2% of all analyzed proteins were identified/quantified only in one replicate and thus resulting

Table 1. Frequency of Missing Quantitative (Abundance) Values among the 1196 Quantified Proteins by Spectral Count (SpC), MS2-TIC, and Ion Current (IC) among Five Replicate LC-MS/MS Runs of a Yeast Sample ($N = 5$)

	no missing	1 missing	2 missing	3 missing	4 missing
SpC	58.4%	10.7%	10.4%	7.4%	13.2%
MS2-TIC	58.4%	10.7%	10.4%	7.4%	13.2%
IC	99.8%	0.2%	0.0%	0.0%	0.0%

missing values in 4 other replicates (denoted as "4 missing" in Table 1); only 58.4% of all proteins did not have missing value (i.e., identified/quantified in all five replicates). By comparison, the IC approach was able to quantify the vast majority of proteins (99.8%) in all five replicates, rendering these proteins free of missing quantitative value (Table 1); the missing values of the rest 0.2% of proteins only exist in one out of the five replicates. This result demonstrated the IC is considerably less prone to the problem of missing quantitative value than MS2-based approaches. The explanation for such a dramatic difference is that the IC approach does not rely on MS2 for calculation of peak areas and thus decoupling the missing quantitative values from missed MS2-based identification in a replicate; given that the high-resolution MS such as Orbitrap enables highly sensitive and specific MS1 detection and excellent matching of peptide precursor among different runs,[10,39,44] in many cases the IC approach is capable of quantifying a peptide in all replicates with sufficient sensitivity even if the peptide was only identified once in all LC-MS experiments. This also contributes to a more reliable relative quantification by IC (discussed below). Per contra, as discussed in the Introduction section, both SpC and MS2-TIC are liable to missing values because of the relatively lower sensitivity and reproducibility of MS2 spectra acquisition, rooting from the use of dynamic exclusion technique and the fact that MS2 spectral acquisitions of low-abundance peptides are often suppressed by coeluting peptides.[10,21,22]

The abundance values of each protein by the three methods and in each replicate are in the Supplemental Table 2.

Quantitative Responses of SpC, MS2-TIC, and IC to Levels of Protein Spiked in a Complex Proteome

For a relative quantification method, the capacity of obtaining linear, quantitative responses to protein abundances in a complex sample is important. A number of previous studies showed good linear correlations between the quantitative values by SpC or IC and protein abundances.[14,22,51,52] Here, to evaluate the correlation between the quantitative values obtained by different approaches and protein abundances in complex proteomes, we spiked *E. coli* extract with bovine serum albumin (BSA) at six different levels (0.025, 0.1, 0.5, 2.5, 12.5, and 62.5% BSA in the total protein). This series of mixtures were independently processed and digested using a precipitation/on-pellet digestion method[34] and then analyzed by LC-MS in triplicate. The relationships of the quantitative values of BSA given by of SpC, MS2-TIC, and IC versus the relative BSA abundances are shown in Figure 3A−C. In this study, no peptide was identified by MS2 from the mixture of 0.025% BSA in *E. coli*, thus only five different levels (0.1% to 62.5% BSA) were quantified for SpC and MS2-TIC methods. For the IC approach, the ion currents of BSA peptides from 0.025% BSA were quantified with well-defined peaks, and the derived quantitative values fit the trend line well (Figure 3C). This indicates that the IC method achieves a wider dynamic range of
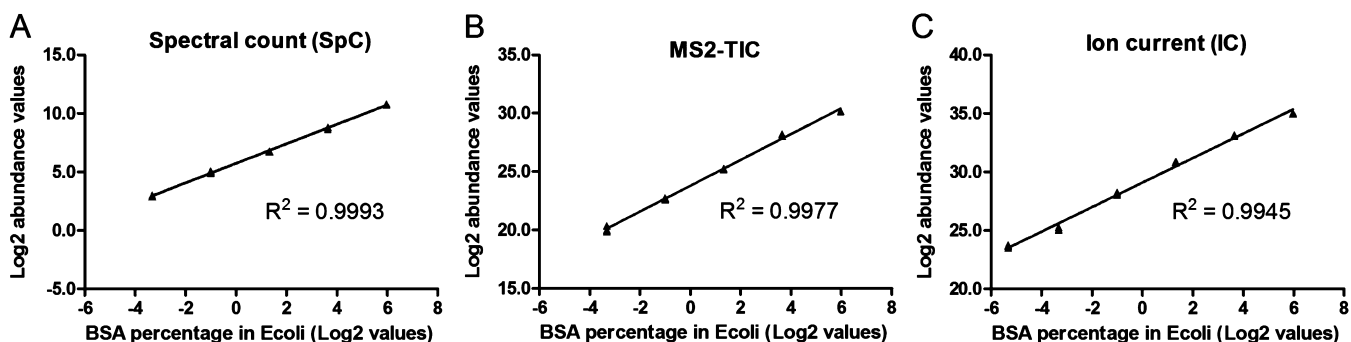
Figure 3. Quantitative responses of spectral count (SpC), MS2-TIC and ion current (IC) vs protein abundance levels. BSA was spiked into *E. coli* extract at six different levels spanning a concentration range >1000. Excellent linearity was observed for (A) spectral count (SpC), (B) MS2-TIC, and (C) ion current (IC). As no BSA-derived peptide was identified in the lowest level, the level was below the detection limits of SpC and MS2-TIC; by comparison, this level can be quantified by IC with sufficient S/N.
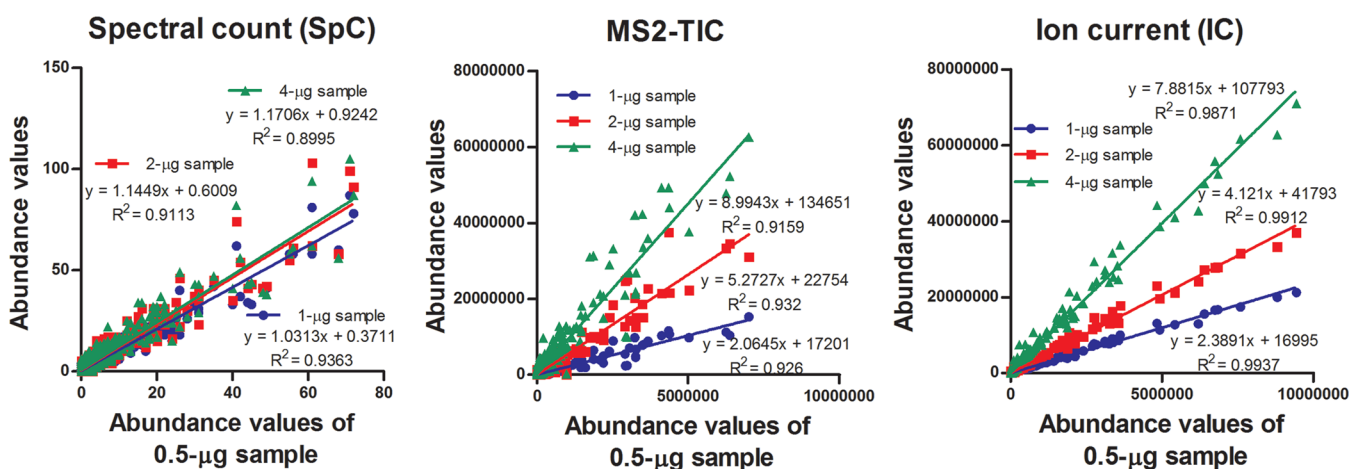

Figure 4. Linear regression analysis correlating the quantitative values with protein loading amounts, by spectral count (SpC), MS2-TIC, and ion current (IC). The quantitative values of individual proteins in 1, 2, and 4 μg loading were individually plotted against these with 0.5 μg. Slopes of trend lines and R^2 values are shown.

protein quantification than SpC or MS2-TIC. Excellent linearity was observed for all three methods ($R^2 \geq 0.99$) over the entire concentration range, which spun at least 3 orders of magnitude. This wide linear range suggests that the three methods may be able to accurately reveal large protein changes. Protein abundance data of BSA by the three methods is available in the Supplemental Table III.

Quantitative Responses to Protein Loading Amounts by SpC, MS2-TIC, and IC

Investigation of correlation between the quantitative values and the total amounts of proteins loaded per LC/MS analysis may reveal the effect of variations in sample preparation and loading and the capacity of each approach to detect or tolerate uneven loading, which would provide valuable information for method development and quality control of label-free quantification approaches. Nonetheless, such a study has been hardly conducted. Griffin et al. demonstrated that protein SI_N values (incorporating abundance features of unique peptide number, spectral count, and fragment ion intensity) of two LC-MS analyses with different protein loads exhibited a linear correlation ($R^2 > 0.94$), and the slope of the line corresponded to the ratio of the two loading amounts.[18]

Here the changes of quantitative values in response to varying total protein loading amounts by SpC, MS2-TIC, and IC were investigated. LC-MS analyses of a prostate cancer cell

(PC3-LN4) sample at the loading levels of 0.5, 1, 2, and 4 μg per injection were utilized for this evaluation. In total, 1122 common protein groups were identified in all four loading levels, and linear regression analysis correlating the quantitative values of these proteins to loading amount respectively at 1, 2, and 4 μg versus 0.5 μg was performed. In order to investigate the "native" quantitative responses by the three methods, normalization was not performed. The results are shown in Figure 4. The linearity of the correlations is good for all three methods, whereas the R^2 values of IC (ranged from 0.987 to 0.994) are better than either SpC (0.899–0.936) or MS2-TIC (0.916–0.932). Interestingly, not all methods showed a linear response to protein loading levels. If a method exhibited a "perfect" linear response to loading amounts, the true values of the slopes of trend lines of 1, 2, and 4 μg injections (all against 0.5 μg) would have been 2.00, 4.00, and 8.00, respectively. As shown in Figure 4, the three slopes by SpC were 1.03, 1.14 and 1.17, indicating no perceivable change in quantitative values responding to varying loading amounts. This is likely because the "dynamic exclusion" (a commonly used feature in data-dependent MS2 experiments) compensates the changes in total protein abundance. By comparison, both MS2-TIC and IC exhibited linear response to loading amounts: the slopes for 1, 2, and 4 μg vs 0.5 μg injections were respectively 2.06, 5.27, and 8.99 for MS-TIC and 2.39, 4.12, and 7.88 for IC. These results indicate both MS2-TIC and IC are capable of perceiving the
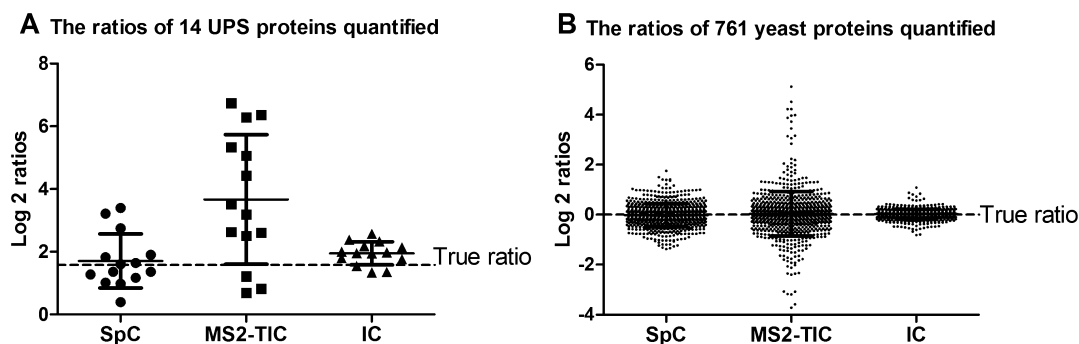
Figure 5. Distribution of the protein ratios in a CPTAC data set (Study 6B *over* 6A) for (A) the 14 UPS proteins and (B) 761 yeast proteins quantified by spectral count (SpC), MS2-TIC, and ion current (IC).
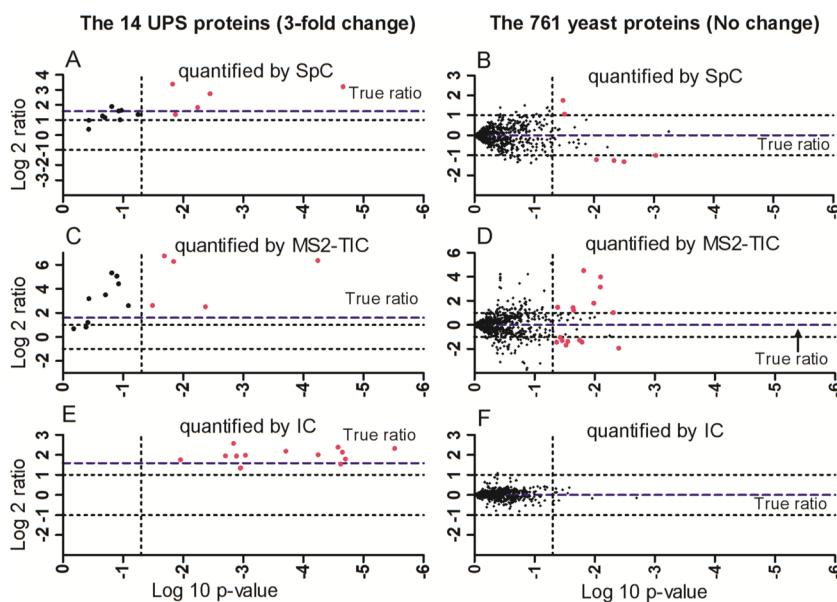

Figure 6. Volcano plots illustrating the discovery of altered proteins in CPTAC study 6B vs 6A set by spectral count (SpC, panels A and B), MS2-TIC (C and D), and ion current (IC, E and F) approaches. The levels of the 14 UPS proteins are different between the two groups (nominal 6B/6A ratio ≈ 3), whereas the levels of yeast proteins are the same. The *Y*-axis shows the log2 ratios of proteins quantified, and the *X*-axis shows the *p*-values (by Student's *t*-test) for the comparison. Each dot represents a unique protein group, and the dashed lines denote the cutoff thresholds ($p \leq 0.05$ and >2-fold change) that define significantly altered proteins, which are in turn shown as red dots.

differences in sample loading, which may be useful characteristics for assessment of the quality of sample preparation and for such cases as the comparison of protein levels relative to units other than the amount of total proteins (e.g., comparing protein levels per volume of a body fluid). On the other hand, these results also indicated SpC is more tolerant to uneven sample loading, which is a valuable feature when it is difficult to achieve a uniform loading across all samples. Finally, the results suggest that for both MS2-TIC and IC, it is necessary to achieve highly reproducible procedures for sample preparation and LC/MS analysis to minimize the effect of variations, and proper normalization approach needs to be in place. Protein abundance values determined by SpC, MS2-TIC, and IC for each protein group are shown in the Supplemental Table IV.

Investigation of the Performance in Discovery of Significantly Altered Proteins Using a Publicly Available Data Set (CPTAC)

One of the most common goals of proteomics is to discover differentially expressed proteins in two different states. In this study, we investigated the performances of SpC, MS2-TIC, and IC in discovering significantly altered proteins in a complex

proteome. To rule out the possibility that the findings to be obtained were associated only with the specific experimental procedures in our lab, a third-party data set was employed for this investigation. Here we compared the performances in discovering significantly altered proteins by the SpC, MS2-TIC, and IC, using one well-characterized, publicly available data set (CPTAC study 6[36,53]). Only the data sets generated by the high-resolution LTQ/Orbitrap were selected. In this subset of CPTAC experiments, the Universal Proteomics Standard set 1 (UPS1 from Sigma-Aldrich, containing 48 human proteins) protein mixture was spiked at different levels into yeast whole lysate, which represents an unchanged, complex proteomic background that is typical in routine biomarker discovery studies. More details of this study set are in previous publications.[36,53]

We first evaluate the three methods based on the relative quantification Study 6B (0.74 fmol/μL UPS1 spiked into yeast lysate) vs 6A (0.25 fmol/μL UPS1 spiked into yeast lysate) samples, which contain relatively low abundance of spiked UPS proteins. Each sample was analyzed in triplicate. Stringent cutoffs for protein identifications were employed to yield a

FDR of 0.1% at peptide level, and two unique peptides were required for each protein group. Quantitative values are normalized against the sum of total spectral counts (SpC), total product ion intensity (MS2-TIC), or total ion current peak area (IC). In the case of missing values, baseline values of 0.5, 1000, and 1000 were respectively assigned for SpC, MS2-TIC, and IC.

A total of 761 yeast proteins and 14 UPS proteins were identified and quantified in the 6B vs 6A set, and the distributions of the ratios of these proteins (6B over 6A) are illustrated in Figure 5. The theoretical 6B/6A ratios of UPS proteins and yeast proteins are respectively ~3 (1.57 on Log2 scale) and 1 (0 on Log2 scale). A previous study demonstrated that accurate quantification in CPTAC 6B vs 6A data set may be difficult due to their low concentrations in these samples.[32] In this study, the observed mean ratios of the 14 UPS proteins were 3.94 ± 2.84, 29.02 ± 35.01, and 3.97 ± 0.99 by SpC, MS2-TIC, and IC, respectively. As shown in Figure 5A, the ratios of the 14 UPS proteins determined by IC were much more tightly centered around the theoretical value than either SpC or MS2-TIC. This is in agreement with the observed good quantitative performance of IC for low-abundance proteins (cf. Figures 1B and 2B) but not SpC or MS2-TIC. The similar trend was also observed in the ratio distribution of the yeast proteins (Figure 5B), where the mean ratios of 761 yeast proteins were 1.04 ± 0.35, 1.35 ± 2.15, and 1.03 ± 0.13, respectively, for SpC, MS2-TIC, and IC, and the ratios of individual yeast proteins by IC were also more tightly centered around the true value compared with the other two approaches. Therefore, it is clearly evident that IC showed better accuracy and precision than SpC and MS2-TIC for relative protein quantification.

In this set of study, the levels of UPS proteins were significantly different between the two groups (mimicking the significantly altered proteins between two proteomic samples), whereas the levels of all the yeast proteins remain constant. Here we compared the three methods for their capacity of accurately discovering the changed UPS proteins, the specificity and sensitivity of discovery, and the levels of false-discoveries. The cutoff thresholds for significantly altered proteins were determined as at least 2-fold differences between the two groups and statistical p-values ≤0.05 (by student t-test) for all three methods. The volcano plots (log2 ratios vs p-values) of UPS and yeast proteins by SpC, MS2-TIC, and IC are shown in Figure 6. The black dashed lines denote the cutoff thresholds, and the altered proteins under these thresholds are indicated by red dots. As shown in Figure 6A,B and Table 2, SpC discovered 11 significantly altered proteins, among which 5 are UPS proteins (i.e., true positives) and 6 are yeast proteins (i.e., false-positives); as to MS2-TIC, 5 UPS proteins and 16 yeast proteins were determined as significantly altered (Figure 6C,D and Table 2). In contrast, all 14 UPS proteins were correctly discovered as significantly altered by IC with no false-positive, as demonstrated in Figure 6E,F and Table 2. For this CPTAC study 6B vs 6A data set, the sensitivity of altered protein discovery were 36%, 36%, and 100% by SpC, MS2-TIC, and IC, respectively, and the levels of false discovery rate by IC are far lower than the two other methods (Table 2).

We further investigated the performances of relative quantification and altered-protein discovery using the next tier of quantification data set, the study 6C (2.2 fmol/μL UPS1 spiked into yeast lysate) versus 6B (0.74 fmol/μL UPS1 spiked into yeast lysate), which also demonstrated the superior

Table 2. Sensitivity and Specificity for the Discovery of Altered Proteins (Biomarkers) by SpC, MS2-TIC, and IC Based on CPTAC Study 6 Data Sets[a]

	spectral count (SpC)		MS2-TIC		ion current (IC)	
	6B/6A	6C/6B	6B/6A	6C/6B	6B/6A	6C/6B
identified biomarkers[b]	11	28	21	36	14	39
true positives (TP)[c]	5	11	5	13	14	28
true negatives (TN)	755	712	745	706	761	718
false positives (FP)	6	17	16	23	0	11
false negatives (FN)	9	21	9	19	0	4
sensitivity, TP/(TP + FN)	36%	34%	36%	41%	100%	88%
specificity, TN/(TN + FP)	99%	98%	98%	97%	100%	98%
false discovery rate, FP/(TP + FP)	55%	61%	76%	64%	0%	28%

[a]The data set consists of the high-resolution MS data of CPTAC study 6A, 6B, and 6C sets, which are triplicate analyses of UPS1 protein mixture spiked, respectively, at 0.25, 0.74, and 2.2 fmol/μL into yeast proteins (representing an unchanged proteomic background). [b]The cutoff thresholds for biomarker discovery are >2-fold changes and p-value ≤0.05. [c]Definition of the terms: if a UPS1 protein were determined as a biomarker, it is a true positive (TP), otherwise a false negative (FN); if a yeast protein was NOT determined as a biomarker, it is a true negative (TN), otherwise a false negative (FN).

performance of the IC approach (Table 2). A total of 729 yeast proteins and 32 UPS proteins were identified and quantified in this data set. Consistent with the observations in the 6B vs 6A study, IC achieved the best sensitivity (88%) for altered-protein discovery, compared to 34% and 41%, respectively, for SpC and MS2-TIC (Table 2); moreover, the false-positive and false-negative levels of IC were far lower (Table 2). The quantitative values by SpC, MS2-TIC, and IC for each quantified protein group in Study 6B vs 6A and 6C vs 6B are shown in the Supplemental Tables V and VI, respectively.

■ CONCLUSIONS

A comprehensive comparison of two types of label-free quantification approaches, ion current-based (IC) and MS2-based (SpC and MS2-TIC) approaches, was conducted using various data sets acquired by high-resolution MS. To date, SpC and MS2-TIC remain powerful and play an important role in classic biomarker discovery studies with apposite statistical tools, especially when low-resolution MS data is used. Nonetheless, it is evident that when a high-resolution MS is used, the IC approach is considerably superior to SpC and MS2-TIC in terms of quantitative reproducibility and accuracy and is much less prone to the missing-data problem, and thereby enabling more reliable proteomics quantification. Furthermore, IC was proved to be a more sensitive, accurate and reliable tool for biomarker discovery than SpC or MS2-TIC, with markedly lower false-positive and false-negative rates. Though high sample-to-sample reproducibility is more crucial for IC approach, development of informatics tools such as good algorithms for LC-MS alignment, normalization of quantitative values in each replicate, and statistical outlier analysis, may significantly reduce this demand. Moreover, when coupled with extensive fractionation and separation approaches such as long-gradient nano-LC and SDS-PAGE fractionation,[16,24,33] IC-based strategy may provide a dependable means for in-depth

analysis and biomarker discovery in complex proteomes.[16,24] Given these favorable characteristics of IC approach, it is expected that further research of this technique on highly reproducible LC/MS analysis and statistics tools, and its clinical applications will emerge rapidly, and its popularity in users of high-resolution MS will continue to expand.

■ DATA SHARING

All raw files and data processing files associated with this paper will be available to public for download upon request.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Protein abundance values by spectral count, MS2-TIC, and ion current for each protein group identified from proteomics sampels of rat brain, rat liver, rat retina, yeast, *E. coli*, human muscle cell, and human bronchoalveolar lavage fluid. Protein abundance values by spectral count, MS2-TIC, and ion current for each protein group quantified in the five replicate LC-MS runs of a yeast sample. Protein abundance values of bovin serum albumin (BSA) by spectral count, MS2-TIC, and ion current approaches. The protein was spiked in *E. coli* extract at six different levels (0.025%, 0.1%, 0.5%, 2.5%, 12.5%, and 62.5% of total proteins). Protein abundance values by spectral count, MS2-TIC, and ion current for each protein group in the PC3-LN4 cell sample loaded at different amounts (0.5, 1, 2, and 4 μg of total proteins). Protein abundance values by spectral count, MS2-TIC, and ion current for each protein group identified in the CPTAC study 6A vs 6B data set. Protein abundance values by spectral count, MS2-TIC, and ion current for each protein group identified in the CPTAC study 6C vs 6B data set. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*Email: (J.Q.) junqu@buffalo.edu. Fax: (716) 645-3693. Tel.: (716) 645-2844, ext. 283.
*Email: (C.T.) ctu3@buffalo.edu. Fax: (716) 645-3693. Tel.: (716) 888-4729.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ ABBREVIATIONS

AUC, area under the curve; CPTAC, the clinical proteomic technology assessment for cancer; CV, coefficient of variation; FDR, false discovery rate; IC, ion current; LTQ/Orbitrap, a linear ion trap combined with an Orbitrap analyzer mass spectrometer; SpC, spectral count; TIC, total ion current; UPS, universal proteomics standard

■ REFERENCES

(1) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **1999**, *17*, 994−999.

(2) Qian, W. J.; Monroe, M. E.; Liu, T.; Jacobs, J. M.; Anderson, G. A.; Shen, Y.; Moore, R. J.; Anderson, D. J.; Zhang, R.; Calvano, S. E.; Lowry, S. F.; Xiao, W.; Moldawer, L. L.; Davis, R. W.; Tompkins, R. G.; Camp, D. G., II; Smith, R. D. Quantitative proteome analysis of human plasma following in vivo lipopolysaccharide administration using 16O/18O labeling and the accurate mass and time tag approach. *Mol. Cell. Proteomics* **2005**, *4*, 700−709.

(3) Yao, X.; Freas, A.; Ramirez, J.; Demirev, P. A.; Fenselau, C. Proteolytic 18O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Anal. Chem.* **2001**, *73*, 2836−2842.

(4) Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **2002**, *1*, 376−386.

(5) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; 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, D. J. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **2004**, *3*, 1154−1169.

(6) Thompson, A.; Schafer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Johnstone, R.; Mohammed, A. K.; Hamon, C. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* **2003**, *75*, 1895−1904.

(7) Dayon, L.; Hainard, A.; Licker, V.; Turck, N.; Kuhn, K.; Hochstrasser, D. F.; Burkhard, P. R.; Sanchez, J. C. Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags. *Anal. Chem.* **2008**, *80*, 2921−2931.

(8) Geiger, T.; Cox, J.; Ostasiewicz, P.; Wisniewski, J. R.; Mann, M. Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nat. Methods* **2010**, *7*, 383−385.

(9) Hebert, A. S.; Merrill, A. E.; Bailey, D. J.; Still, A. J.; Westphall, M. S.; Strieter, E. R.; Pagliarini, D. J.; Coon, J. J. Neutron-encoded mass signatures for multiplexed proteome quantification. *Nat. Methods* **2013**, *10*, 332−334.

(10) Mueller, L. N.; Brusniak, M. Y.; Mani, D. R.; Aebersold, R. An assessment of software solutions for the analysis of mass spectrometry based quantitative proteomics data. *J. Proteome Res.* **2008**, *7*, 51−61.

(11) Bantscheff, M.; Schirle, M.; Sweetman, G.; Rick, J.; Kuster, B. Quantitative mass spectrometry in proteomics: a critical review. *Anal. Bioanal. Chem.* **2007**, *389*, 1017−1031.

(12) Neilson, K. A.; Ali, N. A.; Muralidharan, S.; Mirzaei, M.; Mariani, M.; Assadourian, G.; Lee, A.; van Sluyter, S. C.; Haynes, P. A. Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics* **2011**, *11*, 535−553.

(13) Gao, J.; Friedrichs, M. S.; Dongre, A. R.; Opiteck, G. J. Guidelines for the routine application of the peptide hits technique. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1231−1238.

(14) Liu, H.; Sadygov, R. G.; Yates, J. R., III A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* **2004**, *76*, 4193−4201.

(15) Wiener, M. C.; Sachs, J. R.; Deyanova, E. G.; Yates, N. A. Differential mass spectrometry: a label-free LC-MS method for finding significant differences in complex peptide and protein mixtures. *Anal. Chem.* **2004**, *76*, 6085−6096.

(16) Gautier, V.; Mouton-Barbosa, E.; Bouyssie, D.; Delcourt, N.; Beau, M.; Girard, J. P.; Cayrol, C.; Burlet-Schiltz, O.; Monsarrat, B.; Gonzalez de Peredo, A. Label-free quantification and shotgun analysis of complex proteomes by one-dimensional SDS-PAGE/NanoLC-MS: evaluation for the large scale analysis of inflammatory human endothelial cells. *Mol. Cell. Proteomics* **2012**, *11*, 527−539.

(17) Bondarenko, P. V.; Chelius, D.; Shaler, T. A. Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography-tandem mass spectrometry. *Anal. Chem.* **2002**, *74*, 4741−4749.

(18) Griffin, N. M.; Yu, J.; Long, F.; Oh, P.; Shore, S.; Li, Y.; Koziol, J. A.; Schnitzer, J. E. Label-free, normalized quantification of complex

mass spectrometry data for proteomic analysis. *Nat. Biotechnol.* **2010**, *28*, 83−89.

(19) Ishihama, Y.; Oda, Y.; Tabata, T.; Sato, T.; Nagasu, T.; Rappsilber, J.; Mann, M. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol. Cell. Proteomics* **2005**, *4*, 1265−1272.

(20) Zhang, Y.; Wen, Z.; Washburn, M. P.; Florens, L. Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. *Anal. Chem.* **2010**, *82*, 2272−2281.

(21) Lai, X.; Wang, L.; Tang, H.; Witzmann, F. A. A novel alignment method and multiple filters for exclusion of unqualified peptides to enhance label-free quantification using peptide intensity in LC-MS/MS. *J. Proteome Res.* **2011**, *10*, 4799−4812.

(22) Old, W. M.; Meyer-Arendt, K.; Aveline-Wolf, L.; Pierce, K. G.; Mendoza, A.; Sevinsky, J. R.; Resing, K. A.; Ahn, N. G. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol. Cell. Proteomics* **2005**, *4*, 1487−1502.

(23) Matzke, M. M.; Brown, J. N.; Gritsenko, M. A.; Metz, T. O.; Pounds, J. G.; Rodland, K. D.; Shukla, A. K.; Smith, R. D.; Waters, K. M.; McDermott, J. E.; Webb-Robertson, B. J. A comparative analysis of computational approaches to relative protein quantification using peptide peak intensities in label-free LC-MS proteomics experiments. *Proteomics* **2013**, *13*, 493−503.

(24) Tu, C.; Li, J.; Jiang, X.; Sheflin, L.; Pfeffer, B.; Behringer, M.; Fliesler, S.; Qu, J. Ion current-based proteomic profiling of the retina in a rat model of Smith-Lemli-Opitz Syndrome. *Mol. Cell. Proteomics* **2013**, Epub ahead of print.

(25) Domon, B.; Aebersold, R. Mass spectrometry and protein analysis. *Science* **2006**, *312*, 212−217.

(26) Higgs, R. E.; Butler, J. P.; Han, B.; Knierman, M. D. Quantitative proteomics via high resolution MS quantification: capabilities and limitations. *Int. J. Proteomics* **2013**, *2013*, 674282.

(27) Perrin, R. J.; Payton, J. E.; Malone, J. P.; Gilmore, P.; Davis, A. E.; Xiong, C.; Fagan, A. M.; Townsend, R. R.; Holtzman, D. M. Quantitative label-free proteomics for discovery of biomarkers in cerebrospinal fluid: assessment of technical and inter-individual variation. *PLoS One* **2013**, *8*, e64314.

(28) Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008**, *26*, 1367−1372.

(29) Yates, J. R.; Ruse, C. I.; Nakorchevsky, A. Proteomics by mass spectrometry: approaches, advances, and applications. *Annu. Rev. Biomed. Eng.* **2009**, *11*, 49−79.

(30) Asara, J. M.; Christofk, H. R.; Freimark, L. M.; Cantley, L. C. A label-free quantification method by MS/MS TIC compared to SILAC and spectral counting in a proteomics screen. *Proteomics* **2008**, *8*, 994−999.

(31) Krey, J. F.; Wilmarth, P. A.; Shin, J. B.; Klimek, J.; Sherman, N. E.; Jeffery, E. D.; Choi, D.; David, L. L.; Barr-Gillespie, P. G. Accurate label-free protein quantitation with high- and low-resolution mass spectrometers. *J. Proteome Res.* **2013**, Epub ahead of print.

(32) Chen, Y. Y.; Chambers, M. C.; Li, M.; Ham, A. J.; Turner, J. L.; Zhang, B.; Tabb, D. L. IDPQuantify: combining precursor intensity with spectral counts for protein and peptide quantification. *J. Proteome Res.* **2013**, *12*, 4111−4121.

(33) Tu, C.; Li, J.; Bu, Y.; Hangauer, D.; Qu, J. An ion-current-based, comprehensive and reproducible proteomic strategy for comparative characterization of the cellular responses to novel anti-cancer agents in a prostate cell model. *J. Proteomics* **2012**, *77*, 187−201.

(34) Duan, X.; Young, R.; Straubinger, R. M.; Page, B.; Cao, J.; Wang, H.; Yu, H.; Canty, J. M.; Qu, J. A straightforward and highly efficient precipitation/on-pellet digestion procedure coupled with a long gradient nano-LC separation and Orbitrap mass spectrometry for label-free expression profiling of the swine heart mitochondrial proteome. *J. Proteome Res.* **2009**, *8*, 2838−2850.

(35) Tu, C.; Li, J.; Young, R.; Page, B. J.; Engler, F.; Halfon, M. S.; Canty, J. M., Jr.; Qu, J. Combinatorial peptide ligand library treatment followed by a dual-enzyme, dual-activation approach on a nanoflow liquid chromatography/orbitrap/electron transfer dissociation system for comprehensive analysis of swine plasma proteome. *Anal. Chem.* **2011**, *83*, 4802−4813.

(36) Tabb, D. L.; Vega-Montoto, L.; Rudnick, P. A.; Variyath, A. M.; Ham, A. J.; Bunk, D. M.; Kilpatrick, L. E.; Billheimer, D. D.; Blackman, R. K.; Cardasis, H. L.; Carr, S. A.; Clauser, K. R.; Jaffe, J. D.; Kowalski, K. A.; Neubert, T. A.; Regnier, F. E.; Schilling, B.; Tegeler, T. J.; Wang, M.; Wang, P.; Whiteaker, J. R.; Zimmerman, L. J.; Fisher, S. J.; Gibson, B. W.; Kinsinger, C. R.; Mesri, M.; Rodriguez, H.; Stein, S. E.; Tempst, P.; Paulovich, A. G.; Liebler, D. C.; Spiegelman, C. Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* **2010**, *9*, 761−776.

(37) Elias, J. E.; Haas, W.; Faherty, B. K.; Gygi, S. P. Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. *Nat. Methods* **2005**, *2*, 667−675.

(38) Searle, B. C. Scaffold: a bioinformatic tool for validating MS/MS-based proteomic studies. *Proteomics* **2010**, *10*, 1265−1269.

(39) Lopez, M. F.; Kuppusamy, R.; Sarracino, D. A.; Prakash, A.; Athanas, M.; Krastins, B.; Rezai, T.; Sutton, J. N.; Peterman, S.; Nicolaides, K. Mass spectrometric discovery and selective reaction monitoring (SRM) of putative protein biomarker candidates in first trimester Trisomy 21 maternal serum. *J. Proteome Res.* **2011**, *10*, 133−142.

(40) Sadygov, R. G.; Maroto, F. M.; Huhmer, A. F. ChromAlign: A two-step algorithmic procedure for time alignment of three-dimensional LC-MS chromatographic surfaces. *Anal. Chem.* **2006**, *78*, 8207−8217.

(41) Olsen, J. V.; de Godoy, L. M.; Li, G.; Macek, B.; Mortensen, P.; Pesch, R.; Makarov, A.; Lange, O.; Horning, S.; Mann, M. Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* **2005**, *4*, 2010−2021.

(42) Frese, C. K.; Altelaar, A. F.; Hennrich, M. L.; Nolting, D.; Zeller, M.; Griep-Raming, J.; Heck, A. J.; Mohammed, S. Improved peptide identification by targeted fragmentation using CID, HCD and ETD on an LTQ-Orbitrap Velos. *J. Proteome Res.* **2011**, *10*, 2377−2388.

(43) Makarov, A.; Denisov, E.; Lange, O.; Horning, S. Dynamic range of mass accuracy in LTQ Orbitrap hybrid mass spectrometer. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 977−982.

(44) Zhang, B.; VerBerkmoes, N. C.; Langston, M. A.; Uberbacher, E.; Hettich, R. L.; Samatova, N. F. Detecting differential and correlated protein expression in label-free shotgun proteomics. *J. Proteome Res.* **2006**, *5*, 2909−2918.

(45) Wang, W.; Zhou, H.; Lin, H.; Roy, S.; Shaler, T. A.; Hill, L. R.; Norton, S.; Kumar, P.; Anderle, M.; Becker, C. H. Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Anal. Chem.* **2003**, *75*, 4818−4826.

(46) Colinge, J.; Chiappe, D.; Lagache, S.; Moniatte, M.; Bougueleret, L. Differential proteomics via probabilistic peptide identification scores. *Anal. Chem.* **2005**, *77*, 596−606.

(47) Carvalho, P. C.; Han, X.; Xu, T.; Cociorva, D.; Carvalho Mda, G.; Barbosa, V. C.; Yates, J. R., III XDIA: improving on the label-free data-independent analysis. *Bioinformatics* **2010**, *26*, 847−848.

(48) Webb-Robertson, B. J.; McCue, L. A.; Waters, K. M.; Matzke, M. M.; Jacobs, J. M.; Metz, T. O.; Varnum, S. M.; Pounds, J. G. Combined statistical analyses of peptide intensities and peptide occurrences improves identification of significant peptides from MS-based proteomics data. *J. Proteome Res.* **2010**, *9*, 5748−5756.

(49) Karpievitch, Y. V.; Dabney, A. R.; Smith, R. D. Normalization and missing value imputation for label-free LC-MS analysis. *BMC Bioinf.* **2012**, *13*, S5.

(50) Polpitiya, A. D.; Qian, W. J.; Jaitly, N.; Petyuk, V. A.; Adkins, J. N.; Camp, D. G., II; Anderson, G. A.; Smith, R. D. DAnTE: a statistical tool for quantitative analysis of -omics data. *Bioinformatics* **2008**, *24*, 1556−1558.

(51) Bildl, W.; Haupt, A.; Muller, C. S.; Biniossek, M. L.; Thumfart, J. O.; Huber, B.; Fakler, B.; Schulte, U. Extending the dynamic range of label-free mass spectrometric quantification of affinity purifications. *Mol. Cell. Proteomics* **2012**, *11*, M111 007955.

(52) Wang, G.; Wu, W. W.; Zeng, W.; Chou, C. L.; Shen, R. F. Label-free protein quantification using LC-coupled ion trap or FT mass spectrometry: reproducibility, linearity, and application with complex proteomes. *J. Proteome Res.* **2006**, *5*, 1214−1223.

(53) Rudnick, P. A.; Clauser, K. R.; Kilpatrick, L. E.; Tchekhovskoi, D. V.; Neta, P.; Blonder, N.; Billheimer, D. D.; Blackman, R. K.; Bunk, D. M.; Cardasis, H. L.; Ham, A. J.; Jaffe, J. D.; Kinsinger, C. R.; Mesri, M.; Neubert, T. A.; Schilling, B.; Tabb, D. L.; Tegeler, T. J.; Vega-Montoto, L.; Variyath, A. M.; Wang, M.; Wang, P.; Whiteaker, J. R.; Zimmerman, L. J.; Carr, S. A.; Fisher, S. J.; Gibson, B. W.; Paulovich, A. G.; Regnier, F. E.; Rodriguez, H.; Spiegelman, C.; Tempst, P.; Liebler, D. C.; Stein, S. E. Performance metrics for liquid chromatography-tandem mass spectrometry systems in proteomics analyses. *Mol. Cell. Proteomics* **2010**, *9*, 225−241.