

ICan: An Optimized Ion-Current-Based Quantification Procedure with Enhanced Quantitative Accuracy and Sensitivity in Biomarker Discovery

Chengjian Tu,^{*,†,‡} Quanhu Sheng,[§] Jun Li,^{†,‡} Xiaomeng Shen,^{†,‡} Ming Zhang,^{†,‡} Yu Shyr,[§] and Jun Qu^{*,†,‡}

[†]Department of Pharmaceutical Sciences, University at Buffalo, State University of New York, Kapoor 318, North Campus, Buffalo, New York 14260, United States

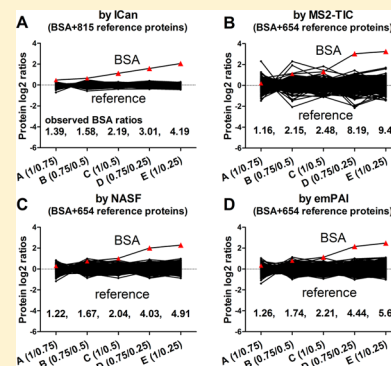
[‡]New York State Center of Excellence in Bioinformatics and Life Sciences, 701 Ellicott Street, Buffalo, New York 14203, United States

[§]Center for Quantitative Sciences, Vanderbilt University School of Medicine, 2220 Pierce Avenue, Nashville, Tennessee 37232, United States

Supporting Information

ABSTRACT: The rapidly expanding availability of high-resolution mass spectrometry has substantially enhanced the ion-current-based relative quantification techniques. Despite the increasing interest in ion-current-based methods, quantitative sensitivity, accuracy, and false discovery rate remain the major concerns; consequently, comprehensive evaluation and development in these regards are urgently needed. Here we describe an integrated, new procedure for data normalization and protein ratio estimation, termed ICan, for improved ion-current-based analysis of data generated by high-resolution mass spectrometry (MS). ICan achieved significantly better accuracy and precision, and lower false-positive rate for discovering altered proteins, over current popular pipelines. A spiked-in experiment was used to evaluate the performance of ICan to detect small changes. In this study *E. coli* extracts were spiked with moderate-abundance proteins from human plasma (MAP, enriched by IgY14-SuperMix procedure) at two different levels to set a small change of 1.5-fold. Forty-five (92%, with an average ratio of 1.71 ± 0.13) of 49 identified MAP protein (i.e., the true positives) and none of the reference proteins (1.0-fold) were determined as significantly altered proteins, with cutoff thresholds of ≥ 1.3 -fold change and $p \leq 0.05$. This is the first study to evaluate and prove competitive performance of the ion-current-based approach for assigning significance to proteins with small changes. By comparison, other methods showed remarkably inferior performance. ICan can be broadly applicable to reliable and sensitive proteomic survey of multiple biological samples with the use of high-resolution MS. Moreover, many key features evaluated and optimized here such as normalization, protein ratio determination, and statistical analyses are also valuable for data analysis by isotope-labeling methods.

KEYWORDS: ion current, label-free, normalization, protein ratio determination, quantitative proteomics



INTRODUCTION

Liquid chromatography–mass spectrometry (LC–MS) techniques have been prevalently employed for the identification and relative/absolute quantification of proteins. LC–MS-based quantification approaches can be roughly divided into two main categories: (i) labeling techniques such as isobaric tags for relative and absolute quantification (iTRAQ),¹ tandem mass tags (TMTs),² stable isotope labeling by amino acids in cell culture (SILAC),³ and neutron-encoded mass signatures (NeuCode)⁴ and (ii) label-free methods such as spectral counting^{5,6} and peptide ion-current-based^{7–9} approaches. Recently, because of their simplicity, cost-effectiveness, and feasibility of multiple biological samples analyses,^{10–12} ion-current-based approaches have emerged as an attractive tool in quantitative proteomics. This trend has been also boosted by

the dramatically increasing availability of high-resolution MS instrumentations in the past few years.¹³

Besides the well-controlled sample preparation and LC/MS procedures, an appropriate method for data analysis is also essential to achieve confident and accurate ion-current-based quantification. For instance, normalization is often applied in label-free quantitative proteomics to reduce the effect of the complicated analytical variability and systematic bias.^{14,15} Many normalization methods such as central tendency, lowess regression, and quantile normalization were first used in the analysis of microarray data^{16,17} and have been recently adapted for analyzing proteomics data.^{18,19} The evaluation of different

Received: August 5, 2014

Published: October 6, 2014

normalization approaches has been widely performed based on high-abundance peptides (common to all or the majority of LC–MS runs) in label-free quantitative proteomics.^{14,19} Kultima et al. demonstrated that the RegrRun (linear regression followed by analysis order normalization) effectively decreased the median SD by 43% on average compared with raw data in peaks that successfully matched across more than 50% LC–MS analyses.¹⁹ In addition, many factors involved in the normalization procedure such as imputation (for missing values), retention time, precursor m/z , and prefractionation of sample also have been studied in label-free quantification.^{15,18,20}

Another important issue for data analysis is choice of methods to compute protein ratios based on peptide quantitative information, which has been widely studied for labeling techniques.²¹ It has been demonstrated that a simple sum-of-intensities algorithm achieved superior performance over other algorithms such as average of the ratios, libra ratio, linear regression, and total least-squares for estimation of true protein ratios.²¹ A systematic evaluation in this regard has not been conducted for ion-current-based label-free method, although various methods were applied in popular packages and procedures. The sum or average intensity method has been employed in packages such as the intensity-based absolute quantification (iBAQ, though the intensity is divided by the number of theoretically observable peptides),²² Progenesis LC–MS software (Nonlinear Dynamics Limited, Newcastle upon Tyne, U.K.),²³ and the ion-current-based method we developed previously.^{9,13} Packages such as Census²⁴ and SIEVE (Thermo Fisher Scientific, San Jose, CA)^{12,25} have applied a variance-weighted method (based on standard deviation or coefficient variation of peaks/peptides) to calculate protein quantitation ratios. Other protein ratio estimation methods such as TOP3 (using the sum intensities of the top-three unique peptides)²⁶ and average ratios²⁷ are also employed in quantitative proteomics.

In this study, we developed and optimized a new label-free quantitative procedure for ion-current-based quantification, ICan (ion-current-based analysis), and evaluated its capacity for proteomic quantification and the discovery of significantly different proteins, even for these with small-fold changes (1.5-fold). Key quantitative features such as frame filtering, normalization, protein ratio determination, and statistical analysis were comprehensively evaluated and optimized. With these optimizations, ICan significantly improved the quantitative accuracy and sensitivity and performance in discovering altered proteins over existing methods.

MATERIALS AND METHODS

Sample Preparation

The PC3-LN4 cells and *E. coli* cells were from Kinex Pharmaceuticals (Buffalo, NY). The rat brain samples were from Buffalo General Medical Center (Buffalo, NY). Cell or tissue samples 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). After homogenization performed for a 5–10 s burst at 15 000 rpm for 10 times, the mixture was then sonicated in a cold room for ~10 min with a low-power sonicator until the solution was clear. Lysates were centrifuged at 140 000g for 1 h at 4 °C. The supernatant was collected and stored at –80 °C until analysis. For preparation of moderate-abundance proteins (MAPs), the

plasma sample (~200 μ L) from a healthy young woman was fractionated with IgY14-SuperMix tandem column (Sigma-Aldrich), as previously reported.²⁸ Three buffers (dilution/washing buffer: 10 mM Tris-HCl, 150 mM NaCl, pH7.4 (TBS); stripping buffer: 100 mM glycine, pH2.5; neutralization buffer: 100 mM Tris-HCl, pH8.0) were, respectively, used for loading/washing, eluting, and neutralization. The resulting flow-through fraction (low-abundance proteins) and the bound/eluted fractions from IgY-14 (high-abundance proteins) and from SuperMix (MAPs) were collected separately. All fractions were then individually concentrated in Amicon centrifugal filter with 3-kDa molecular mass cutoff (EMD Millipore), followed by buffer exchange to 50 mM NH_4HCO_3 according to the manufacturer's instruction. Protein concentration was measured using BCA Protein Assay (Pierce, Rockford, IL). The amounts of 100 and 90 μ g *E. coli* extracts were, respectively, spiked with bovine serum albumin (BSA) at four different levels (0.025, 0.05, 0.075, and 0.1% of total proteins) and MAPs at two different levels (5 μ g and 7.5 μ g). All samples (each containing ~100 μ g of total protein) were reduced with TCEP (3 mM) for 10 min and then alkylated with 20 mM IAM for 30 min in darkness. A precipitation/precipitation-digestion procedure was applied to performed precipitation and tryptic digestion as previously described.^{9,29}

NanoLC–MS/MS Analysis

Peptide samples were analyzed using an ultrahigh pressure Eksigent (Dublin, CA) nano-2D Ultracapillary/nano-LC system coupled to a LTQ/Orbitrap XL hybrid mass spectrometer (Thermo Fisher Scientific, San Jose, CA). The mobile phase consisted of 0.1% formic acid in 2% acetonitrile (A) and 0.1% formic acid in 88% acetonitrile (B). Samples were loaded onto a reversed-phase trap (300 μ m ID \times 1 cm), with 1% mobile phase B at a flow rate of 10 μ L/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 μ m ID \times 75 cm, packed with 3 μ m particles) for separation. The nano-LC column was heated to 52 °C to greatly improve both chromatographic resolution and reproducibility. To stabilize ionization efficiency, the spray tip was cleaned by dripping 50% methanol by gravity after every three runs. The parameters for MS were demonstrated in our previous publications.^{9,12}

In this study, for the spiked-in BSA experiment, each group at different BSA concentration was analyzed four times; for the spiked-in MAP experiments, two groups at different MAPs concentration were alternatively analyzed five times. Five consecutive runs of the rat brain sample and six runs with different load amount of PC3-LN4 cell (1 and 2 μ g, three replicates per group) were further analyzed to assess different normalization methods. In addition, to assess the performance of biomarker discovery by ICan and iBAQ, we employed the “Study 6 LTQ Orbitrap XL @P65” data set generated by the program of Clinical Proteomic Technology Assessment for Cancer (CPTAC).³⁰ 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 and 0.74 fmol/ μ L into yeast lysate for sets A and B, 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 Validation

Proteome Discoverer version 1.4.1.14 (Thermo-Scientific) was used to perform database searching against Swiss-Prot protein database (version 06/13/2012) for the BSA spiked-in experiment, five consecutive LC–MS runs experiment, and six LC–MS runs with different load amount experiment. MaxQuant³¹ v1.4.1.2, incorporated with the Andromeda search engine,³² was used for the MAP spiked-in experiment and CPTAC study 6 data. A total of 7766 protein entries, 20 238 entries, 4431 entries, and 7801 entries were presented in respective rat, human, *E. coli*, and yeast database. The databases were augmented with sequence of BSA, the UPS1 48 proteins (Sigma-Aldrich), and 118 MAPs (achieved from three replicate LC–MS/MS runs of MAP sample) 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.³³ The sequence database contains each sequence in both forward and reversed orientations, enabling FDR estimation. For resulted files from Proteome discoverer, Scaffold v4.2.0 (Proteome Software, Portland, OR) was used to validate MS2-based peptide and protein identification based on cutoffs of cross-correlation (Xcorr) and Delta Cn values. The FDR was set to 0.01 and 0.05, respectively, for peptide and protein identifications. For MaxQuant, the FDR was set to 0.01 for peptide and protein identifications, respectively. The identifications from the reverse database and common contaminants were eliminated.

Protein Quantification

The protein quantitative values based on MS2-TIC, NASF and emPAI for each data set were obtained using Scaffold v4.2.0 under the same peptide/protein identification criteria. The iBAQ intensities, the sum of intensities of all peptides divided by the number of theoretically observable peptides, were achieved from the MaxQuant using standard settings with the option “match between” runs selected. The iBAQ values for each protein are normalized against sum of quantitative values in individual runs. The quantitative analysis by ICan was performed as shown in the pipeline (Figure 1). The peak detection and chromatographic alignment based on retention time, m/z , and charge states were analyzed by SIEVE v2.1 (Thermo Scientific, San Jose, CA). Quantitative frames/features 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. Subsequently, using tools in-house the MS2 fragmentation scans associated with each frame were assigned to the peptide/protein identifications from Proteome Discoverer or MaxQuant as previously described. Frames assigned to multiple peptides were excluded in ICan. The LOESS normalization³⁴ was performed to reduce the systematic bias. In the case of missing data, a value of 1000 as the baseline quantitative value was assigned.¹³ After further excluding frames shared with multiple proteins, intensities for frames with the same sequence were combined to be the unique peptide intensity and then intensities for unique peptides of the same protein were further combined to be the protein intensity with Grubbs' test analysis in both steps.

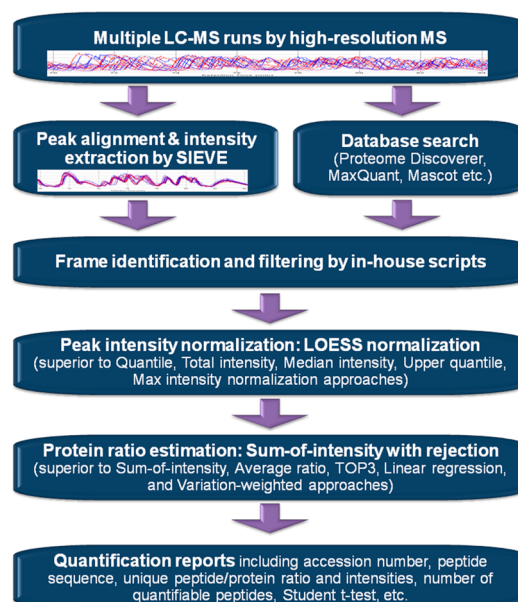


Figure 1. Flowchart of ICan. The ICan supports identification results from Proteome Discoverer, MaxQuant, and Mascot. The optimal normalization and protein ratio estimation approaches were also integrated.

Grubbs' test was performed by the ListPOR (v Version 2.2.2104) program (panomics.pnnl.gov). Minimum data set presence 3 and 2, p value cutoff of 0.01 and 0.05, were, respectively, set at frame level and unique peptide level. The relative protein ratio was calculated by comparing the summed abundance values of the protein in each group. Student's t -test statistics was applied to analyze log-transformed values of protein intensities for all of these methods. Abundance change ≥ 1.3 -fold and p value ≤ 0.05 were used as the thresholds to define altered proteins. The p -value adjustments for multiple testing were evaluated according to sequential Bonferroni correction (SB),³⁵ Benjamini–Hochberg FDR control (BH),³⁶ and sequential Fisher's combined probability test (SFisher).³⁷

RESULTS AND DISCUSSION

For label-free proteomic quantification, accurate and precise quantification of low-abundance proteins remains challenging. As demonstrated by various laboratories including ours, spectral count-based approaches resulted in suboptimal quantification of low-abundance proteins due to the inherent biases and variations in data-dependent sampling of fragment ions (MS2).^{9,10,38} By comparison, ion-current-based approaches have been shown to afford markedly improved quantification for low-abundance proteins when efficient and reproducible liquid-chromatography (LC) separation and high-resolution MS are employed.^{9,10,13,39}

To date, owing to the prevalent use of high-resolution MS, ion-current-based methods have become the most promising label-free approaches.^{10,13} However, comprehensive evaluation and optimization of data analysis approaches for ion-current-based quantification have not been adequately reported. Here, based on extensive evaluation and optimization, we developed an optimal ion-current-based procedure (Figure 1) termed ICan (ion-current-based analysis) and assessed its capacity for proteomic quantification and discovery of significantly altered proteins, even for these with small-fold changes (1.5-fold). The ICan is designed for data generated from high-resolution MS;

in the current work, we chose to interface SIEVE (Thermo Scientific, San Jose, CA) with this pipeline, which performs peak detection and chromatographic alignment based on retention time, m/z , and charge states. Each aligned quantitative feature (i.e., a frame, the set of peak areas of a specific peptide) was correlated with the peptide/protein ID information from popular software such as Proteome Discoverer, MaxQuant, and Mascot with scripts developed in-house. Streamlined processes for frame filtering, LOESS normalization,³⁴ and outlier detection by Grubbs test⁴⁰ on both frame and peptide levels were integrated in ICan. These processes were comprehensively optimized and proved to significantly improve the quantitative accuracy and sensitivity and performance in discovering altered proteins over existing strategies.

Frame Identification and Filtering

In this study, the frame identification is derived from the spectrum identification results from popular database search algorithms such as Proteome Discoverer and MaxQuant. We used in-house scripts to assign these peptide identifications to the distinguished frames. On the basis of our previous studies,^{9,13} it was observed that some frames contained multiple unique peptides and thus may lead to unreliable quantification. Here we examined the shared frame issue using the analysis of a series of *E. coli* extracts spiked with BSA at four different levels (0.025, 0.05, 0.075, and 0.1% of total proteins; four replicates per group). A total of 818 proteins including BSA were identified with a peptide FDR of 0.1% (Supplemental Table 1 in the Supporting Information). Among the total of 13 801 quantitative frames (Supplemental Table 2 in the Supporting Information), 654 (4.7%) assigned to multiple peptide IDs were observed. Of these shared frames, 617 (94.3%) frames only have two unique peptides (Supplemental Figure 1 in the Supporting Information). The peptides with shared frames likely have indistinguishable m/z and retention time, or some of them were derived from misassigned peptide during database search. We evaluated different cutoff thresholds for peptide/protein identification and found that more stringent cutoffs for identification (e.g., lower identification FDR threshold) reduced the percentage of shared frames, and thus stringent criteria for identification is advisable. Some representative data are shown in Table 1. Moreover, the peptide FDR in shared

frame-associated spectra is much higher (~9-fold) than the determined global peptide FDR (Table 1), indicating increased incorrect identifications in shared frames. Thus, in this study, those share frames containing multiple unique peptides were eliminated.

Evaluation of Normalization Approaches

An optimal normalization method is indispensable to reduce systematic biases and variations and thus to ensure the accuracy and precision of relative quantification in multiple samples. Previously, many normalization approaches have been evaluated for label-free quantification on relatively high-abundance peptides that are commonly identified in all or the majority of LC–MS runs in an experimental set.^{14,19} Here we evaluated all of the identified peptides with a wide range of abundance levels by six different normalization methods, including LOESS, quantile, upper-quantile, maximum intensity, median intensity, and total intensity normalization (Supporting Information). The LOESS and quantile normalization achieved best performances in the spiked-in BSA experiment, which decreased the median coefficient variations (CVs) of *E. coli* peptide intensities by an average of 29% compared with the original data (Figure 2). We also evaluated these methods on

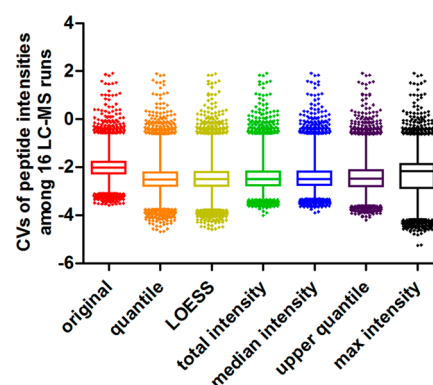


Figure 2. Evaluation of different normalization approaches using the spiked-in BSA data. BSA was spiked into *E. coli* extracts at four different levels (0.025, 0.05, 0.075, and 0.1% of total proteins; four replicates/group). Box and Whiskers (1–99 percentile) plot was used to analyze the coefficient variations (CVs) of *E. coli* peptide intensities among these 16 LC–MS runs using different normalization methods.

Table 1. Evaluation of the Percentage of Shared Frames in Multiple Data Sets with Different Peptide FDR^a

	five repli- cates of yeast		20 repli- cates of rat brain		spiked-in BSA ex- periment	
peptide FDR	0.51%	0.10%	0.52%	0.10%	0.45%	0.11%
protein FDR	7.69%	1.55%	9.92%	2.26%	8.35%	2.08%
identified proteins	1119	970	1199	1063	886	818
total frames	15508	12486	34264	29832	14557	13801
shared frames	502	236	2533	1553	958	654
percentage of shared frames	3.24%	1.89%	7.39%	5.21%	6.58%	4.74%
peptide FDR in shared frames	4.09%	0.61%	4.77%	0.92%	5.79%	1.10%

^aSpiked-in BSA experiment, 5 replicates of yeast and 20 replicates of rat brain were analyzed in this study. Shared frames: frames assigned to multiple unique peptides.

two other data sets, five LC–MS runs of the same rat brain digest and six runs of the same digest with different load amounts, respectively, representing data sets with minimal and substantial variations of sample preparation and loading. The LOESS approach showed the most effective normalization (Supplemental Figure 2 in the Supporting Information) and thus was employed in ICan and subsequent studies.

After normalization, the level of missing data and reproducibility of the quantitative features by ICan was further evaluated based on replicate LC–MS runs. One of the most prominent advantages of ion-current-based approach over spectral counting or fragment-ion intensities (MS2-TIC) is the reliable quantification of low-abundance peptides, even though a peptide was only identified for once in the entire sample set,^{9,10,13} thus substantially reducing the frequency of missing values and improving the analytical reproducibility. As shown in Supplemental Figure 3 in the Supporting Information, although several methods were employed to improve the reproducibility of LC–MS analysis as previously described,⁹

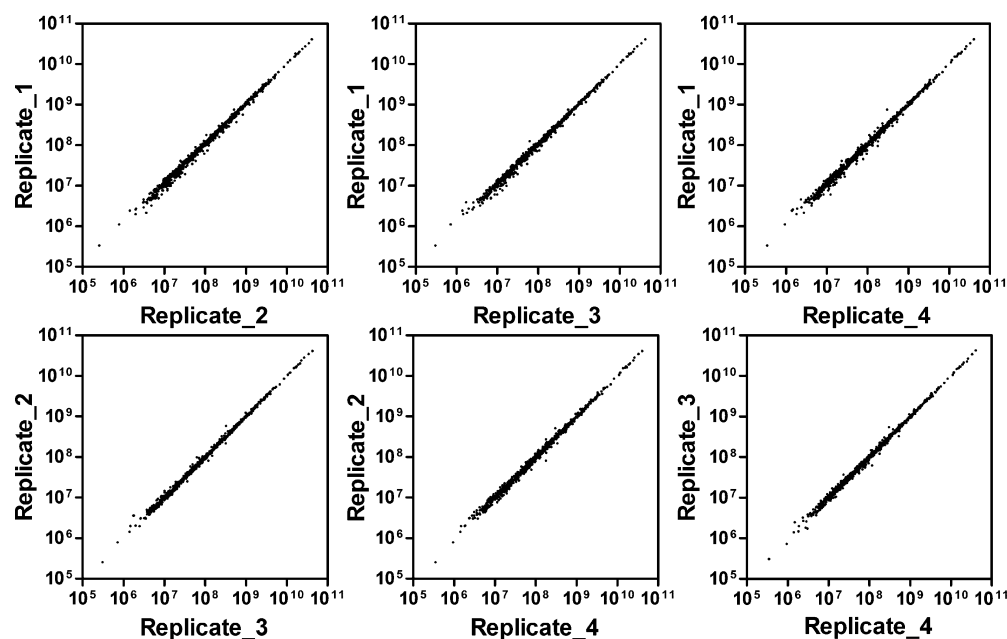


Figure 3. Scatter plot of quantitative feature pairs to evaluate the analytical reproducibility. The excellent correlation of protein intensities between different replicates of a spiked-in BSA (0.075%) group was observed. The two axes represent the quantitative abundance values of the same proteins, respectively, by the two duplicate runs.

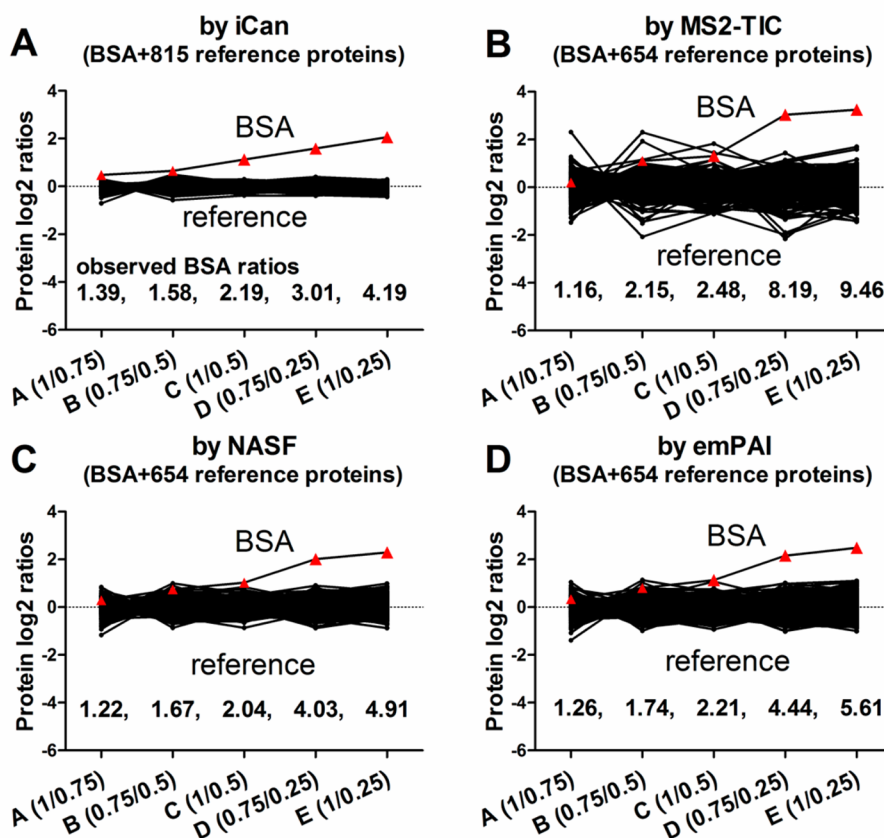


Figure 4. Evaluation of accuracy and precision of relative quantification analysis by iCan, MS2-TIC, NASF, and emPAI using the spiked-in BSA data. BSA was spiked into *E. coli* extracts at four different levels (0.025, 0.05, 0.075, and 0.1% of total proteins). The expected ratio of 1 for reference proteins and five theoretical fold changes of 1.33 (0.1/0.075), 1.50 (0.075/0.05), 2.00 (0.1/0.05), 3.00 (0.075/0.025), and 4.00 (0.1/0.025) for BSA were investigated.

only 655 (80.1% of all identified) proteins were identified in all 16 LC–MS runs and thereby quantifiable by spectral counting or MS2-TIC without missing data. Per contra, the iCan was

able to quantify 816 (99.8%) proteins without any missing value across the 16 LC–MS runs. Two proteins (0.2%) were filtered out because all frames assigned to these two proteins

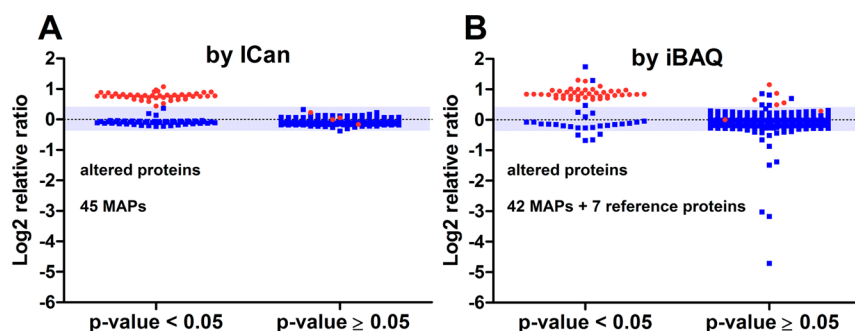


Figure 5. Relative ratios obtained by (A) ICan and (B) iBAQ for a quantitative experiment of *E. coli* extracts spiked with human plasma moderate-abundance proteins (MAPs) ($N = 5/\text{group}$). In total, 49 MAP proteins (red dots, expected ratio is 1.5 between two groups) and 726 *E. coli* proteins (blue dots, expected ratio of 1.0) were quantified. Gray shade denotes ≤ 1.3 -fold change (i.e., the cutoff threshold).

were shared frames. We evaluated the quantitative reproducibility of ICan by correlating the protein intensities between any two of the four LC–MS analyses in the spiked-in BSA (0.075%) group. Here the protein intensity was obtained by summing the areas of all peptide peaks assigned to the specific protein. Linear regression of the correlation between two replicate runs was performed, and the R -squared values for paired correlations are all above 0.99, indicating the excellent quantitative reproducibility (Figure 3). Moreover, a high quantitative reproducibility was also achieved for both high- (the upper segment of each line) and low-abundance proteins (the lower segment of each line). The reproducibility of spectral counting or MS2-TIC methods was far inferior, which is particularly sound for low-abundance proteins (Supplemental Figure 4 in the Supporting Information). These results are in agreement with previously reported.^{9,13}

Accuracy and Precision of Relative Quantification by ICan

The preliminary assessment of quantitative accuracy and precision by ICan was performed using the BSA spiked-in *E. coli* data. The expected ratio of the reference proteins (*E. coli*) was 1.00, and the five possible changes of BSA have expected ratios of 1.33 (0.1%/0.075% BSA in *E. coli*), 1.50 (0.075%/0.05%), 2.00 (0.1%/0.05%), 3.00 (0.075%/0.025%), and 4.00 (0.1%/0.025%), respectively. As shown in Figure 4, ICan quantified nearly all identified proteins without missing data (as previously described), and the measured BSA ratios agreed very well with the expected values with small relative deviations (0.3–9.5%). Excellent linearity between the nominal and observed ratios was achieved (Supplemental Figure 5 in the Supporting Information). The ratios of reference proteins determined by ICan were tightly centered around the theoretical value. The means and standard deviations of the ratios of reference proteins, were, respectively, 0.99 ± 0.06 , 1.01 ± 0.07 , 1.00 ± 0.05 , 1.00 ± 0.08 , and 0.98 ± 0.05 for the five comparisons previously mentioned ($N = 4/\text{group}$), reflecting the high accuracy and precision achieved by ICan in calculating protein expression ratios. The popular MS2-based methods such as MS2-TIC, the normalized spectral abundance factor (NASF),⁴¹ and exponentially modified protein abundance index (emPAI)⁴² were also evaluated (Figure 4B–D). To achieve optimal analysis for these methods, we employed only the 655 proteins that had no missing data in any of the replicates when performing quantification with these approaches. Even ICan calculated the lowest 20% proteins in abundance while other MS2-based methods did not, it still performed significantly better in terms of quantitative accuracy and precision, as shown

in Figure 4. On the basis of these results, the following sections are focused on the comparison of ion-current-based strategies.

Sensitivity and False-Positive Rate for Discovering Altering Proteins

For proteomics quantification, one of the major aims is to completely discover the true altered proteins to the extent possible, while minimizing false-positives that can otherwise lead to misleading biological clues and waste of resources in informatics analysis and validation. To evaluate the sensitivity and false positive rate (FPR) of biomarker discovery by ion-current-based quantification methods, we spiked a mixture of MAPs obtained from human plasma into *E. coli* extracts at two different levels (MAP-A: 90 μg of *E. coli* and 5 μg of MAP; MAP-B: 90 μg of *E. coli* and 7.5 μg of MAP). In this set, the expected ratio of reference proteins (*E. coli*) and MAPs were 1.00 and 1.50 (MAP-B/MAP-A), respectively. A total of 775 proteins including 49 MAPs were identified with a peptide and protein FDR of 1%, respectively, using MaxQuant³¹ (version 1.4.1.2). The list of peptide and protein identifications was shown in Supplemental Table 3 in the Supporting Information. When using 1.3-fold change (the lowest quantifiable fold-change by our ion-current-based quantitative method based on our previous investigation⁹) and $p \leq 0.05$ (t test) as the cutoff thresholds, 45 of 49 (91.8%) MAPs and none of reference proteins were determined as altered proteins (FPR = 0%) by ICan, as shown in Figure 5A (details in Supplemental Table 4 in the Supporting Information). The reference proteins and MAP were, respectively, indicated by blue and red dots. The spots in gray shade have ratios below 1.3-fold change. The mean and standard deviation of the ratios of 45 altered MAPs quantified by ICan was 1.71 ± 0.13 , demonstrating excellent sensitivity and accuracy in discovery of changed proteins. The outstanding ability of ICan for identifying altered proteins was further proved by the area under curve value of 0.97 using receiver-operating characteristic (ROC) analysis (Supplemental Figure 6 in the Supporting Information). In this study, it was showed that ICan could quantify nearly all identified proteins and assign significance with high sensitivity and low FPR to small changes of 1.5-fold, providing a competitive ability in the field of quantitative proteomics.

The iBAQ method, which divides the sum of intensities of all peptides by the number of theoretically observable peptides, was shown to be the most accurate among different absolute quantification methods in a previous work.⁴³ The intensities or iBAQ values for proteins were also achieved from the MaxQuant using standard settings with the option of “match between runs” selected. Thus, here the same list of peptide/

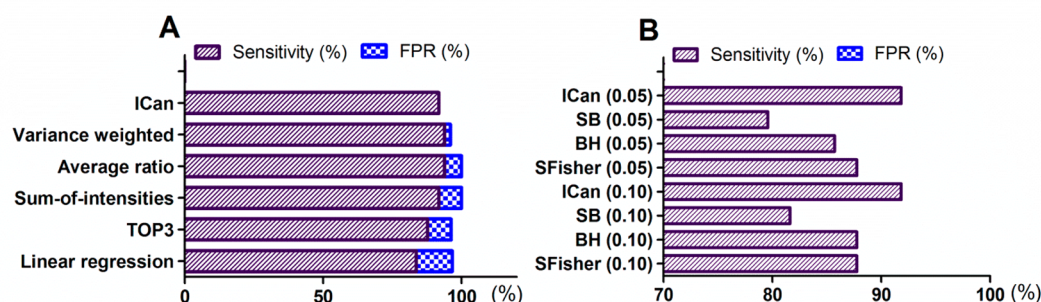


Figure 6. Evaluation of (A) different methods for aggregating quantitative data from peptide-level to protein levels and (B) multiple testing approaches for ICan-based quantification. The false-positive rate (FPR) and sensitivity for discovering altered proteins were investigated with the combination of statistical analysis and a fold-change filter (1.3-fold). A p value of 0.05 was adopted. For investigation of multiple testing, critical significance levels of both 0.05 and 0.10 were evaluated. SB, Sequential Bonferroni test; BH, Benjamini and Hochberg test; SFisher, Sequential Fisher combined probability test.

protein identifications from MaxQuant was shared and analyzed by ICan and iBAQ. We calculated the relative ratios of proteins by the iBAQ values (Supplemental Table 4 in the Supporting Information) against ICan. Using the same threshold, 42 of 49 (85.7%) MAPs and 7 reference proteins (Figure 5B) were determined as altered proteins (FPR = 14.3%) by iBAQ, with a significantly lower sensitivity and higher FPR than ICan. The mean and standard deviation of the ratios of 42 altered MAPs quantified by iBAQ was 1.8 ± 0.2 , also indicating good accuracy and precision in discovery of changed proteins.

A third-party, publicly available data set, the Clinical Proteomic Technology Assessment for Cancer (CPTAC) study 6 data,³⁰ was employed for further investigation of these two methods on the relative quantification. Here Study 6B (0.74 fmol/ μ L UPS1 spiked into yeast lysate) versus 6A (0.25 fmol/ μ L UPS1 spiked into yeast lysate A) samples, which contain relatively low abundance of UPS proteins, were selected to analyze. The expected ratios for yeast proteins and UPS were, respectively, 1.0 and 3.0. After database searching by MaxQuant, a total of 777 proteins including 15 UPS proteins were identified in this study. Using the same threshold (≥ 1.3 -fold and $p \leq 0.05$), all (100%) UPS with a median ratio of 3.25 and 1 yeast protein were determined by ICan as significantly altered proteins, while 12 (80%) UPS with a median ratio of 4.78 and 5 yeast proteins were determined by iBAQ (Supplemental Figure 7 and Supplemental Table 5 in the Supporting Information). Again, ICan was demonstrated to be superior in that it identified more true-positives (UPS proteins) with higher quantitative accuracy and lower FPR than iBAQ method.

Evaluation of Protein Ratio Determination and Multiple Hypothesis Testing

Wrong peptide identification or incorrect assignment of peptide ID to quantitative frames may severely compromise the quantification of the affected proteins; in quantitative analysis, these incorrectly identified/assigned peptides often take the form of outliers, which must be removed to ensure reliable quantification. Here we used Grubbs' test⁴⁰ to identify and then eliminate outliers arising from wrong peptide assignment or large biological/technical variations before the calculation of the quantitative values of unique peptides and proteins.

In this study, we further evaluated the protein ratio determination method by comparing a sum-of-intensities method with outlier removal versus other popular approaches.

Using the abundance values obtained by ICan, approaches for aggregating quantitative data from peptide-level to protein levels such as TOP3, sum-of-intensity, average ratios, variance-weighted (on coefficient variation of peptide), and linear regression (Supporting Information) were evaluated versus ICan. As previously described, these approaches have been widely used in quantitative proteomics. As shown in Figure 6A, similar sensitivity for biomarker discovery was achieved by ICan, variance-weighted, average ratio, and sum-of-intensity approaches using the spiked-in MAP data. The ICan and variance-weighted approaches showed the lowest and second lowest FPR for identifying altered protein. Without outlier analysis, variance-weighted approach achieved the comparable sensitivity with ICan in discovering altered proteins in this study, while the sensitivity of other approaches are inferior (Figure 6A). In addition, it is clear that Grubbs' test outlier analysis greatly reduced the false-positives (ICan vs sum-of-intensity) (Figure 6A). For instance, *E. coli* protein glutamine-fructose-6-phosphate transaminase (Glms, expected ratio is 1.0), determined as an altered protein (1.56-fold and p value = 0.02) by sum-of-intensity method, was quantified by 11 unique peptides, but 10 (90.9%) of them have ratios around the expected ones, as shown in Supplemental Figure 8 in the Supporting Information. The ICan analysis (sum-of-intensity with rejection) removes the outlier (red spots in Supplemental Figure 8 in the Supporting Information) and gives the protein ratio (0.97-fold and p value = 0.43) that agrees well with most of the peptide ratio. Therefore, here we utilized the sum-of-intensity with rejection for protein ratio estimation in the ion-current-based quantification procedure to replace the sum-of-intensity approach we described in previous studies.^{9,13}

We also evaluated multiple hypothesis testing such as Sequential Bonferroni correction (SB),³⁵ Benjamini-Hochberg FDR control (BH),³⁶ and Sequential Fisher's combined probability test (SFisher)³⁷ to adjust the p value of t test (Supporting Information). For investigation of multiple testing, we used 0.05 and 0.10, respectively, as the critical significance level. With the combination of fold-change (1.3-fold threshold) and statistical testing (0.05 or 0.10), the superior performance of biomarker discovery was observed by SFisher compared with the other two methods (Figure 6B and Supplemental Table 6 in the Supporting Information) in ion-current-based quantification. Forty-three (87.8%) MAPs and none of reference proteins were identified as altered proteins with the thresholds of ≥ 1.3 -fold and $p \leq 0.05$ by SFisher. The CPTAC data described

above were also tested using multiple testing, and a similar result was shown in Supplemental Figure 9 and Supplemental Table 7 in the Supporting Information, indicating the superior of SFisher.

CONCLUSIONS

Ion-current-based quantitative approach has emerged as an attractive alternative to both spectral counting and labeling methods, which can analyze many biological samples for large-scale studies such as clinical and pharmaceutical investigations. Recently, the wide prevalence of high-resolution MS has greatly boosted the quality of ion-current-based analysis. Moreover, the substantial advancements in MS instrumentation (e.g., analysis of ~4000 unique yeast proteins in 1 h of LC-MS/MS run using a hybrid Orbitrap MS instrument⁴⁴), will markedly enhance the coverage of ion-current-based analysis. For ion-current-based strategy, a data-processing procedure enabling accurate, precise, and sensitive quantification is critical. Here we demonstrated that the ICan procedure is optimal for ion-current-based quantitative analysis, which provides superior quantitative accuracy and higher sensitivity for biomarker discovery with a lower FDR than these popular methods we've tested. Furthermore, the comparative investigations of various quantitative features in this study provide highly valuable information for the development and evaluation of algorithms for both labeling and label-free methods.

DATA SHARING

All raw files associated with this paper are available at <https://chorusproject.org/pages/dashboard.html> for free downloads.

ASSOCIATED CONTENT

Supporting Information

Supplementary Figure 1. The distribution of shared frames with different unique peptides in the spiked-in BSA experiment. Supplementary Figure 2. Evaluation of different normalization methods. Supplementary Figure 3. Frequency of quantitative abundance values of proteins among 16 replicate LC-MS runs of the spiked-in BSA experiment. Supplementary Figure 4. Quantitative reproducibility of spectral count and MS2-TIC. Supplementary Figure 5. The excellent linear relationship between the expected and observed ratios by ICan. Supplementary Figure 6. The excellent discriminatory ability of ICan for identifying altered proteins. Supplementary Figure 7. The CPTAC study 6B versus 6A set analyzed by ICan and iBAQ. Supplementary Figure 8. The distribution of peptide ratios of *E. coli* protein Glms. Supplementary Figure 9. Evaluation of multiple testing approaches in ion-current based quantification. Supplementary Table 1. The protein report file from the spiked-in BSA experiment. Supplementary Table 2. The frame report file from the spiked-in BSA experiment. Supplementary Table 3. Final list including 775 proteins and 6194 peptides were identified from the spiked-in MAPs experiment data by MaxQuant. Supplementary Table 4. Protein abundance values by ICan and MaxQuant for each protein group identified in the spiked-in MAPs experiment. Supplementary Table 5. Protein abundance values by ICan, and MaxQuant for each protein group identified in the CPTAC study 6A versus 6B data set. Supplementary Table 6. Evaluation of multiple testing using the spiked-in MAPs data. Supplementary Table 7. Evaluation of multiple testing using CPTAC

study 6B versus 6A set. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Authors

*C.T.: Phone: (716) 888-4729. Fax: (716) 645-3693. E-mail: ctu3@buffalo.edu.

*J.Q.: Phone: (716) 645-2844 x283. Fax: (716) 645-3693. E-mail: junqu@buffalo.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported, in part, by American Heart Association (AHA) award 12SDG9450036 (J.Q.), by NIH grants U54HD071594 (J.Q.), AG048388 (J.Q.), HL103411, and AI060260 (J.Q.) by a Center of Protein Therapeutics Industrial Award (J.Q.).

ABBREVIATIONS

CPTAC, the clinical proteomic technology assessment for cancer; CV, coefficient of variation; FDR, false discovery rate; FPR, false-positive rate; ICan, ion-current-based analysis; LTQ/Orbitrap, a linear ion trap combined with an Orbitrap analyzer mass spectrometer; MAP, moderate-abundance proteins; SpC, spectral count; TIC, total ion current; UPS, universal proteomics standard

REFERENCES

- (1) 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.; Bartlett-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* (12), 1154–1169.
- (2) 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* (8), 1895–1904.
- (3) 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* (5), 376–386.
- (4) 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* (4), 332–334.
- (5) 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* (8), 1231–1238.
- (6) Liu, H.; Sadygov, R. G.; Yates, J. R., 3rd. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* **2004**, *76* (14), 4193–4201.
- (7) 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* (20), 6085–6096.
- (8) 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* (18), 4741–4749.
- (9) Tu, C.; Li, J.; Jiang, X.; Sheflin, L. G.; Pfeiffer, B. A.; Behringer, M.; Fliesler, S. J.; Qu, J. Ion-current-based proteomic profiling of the

retina in a rat model of Smith-Lemli-Opitz syndrome. *Mol. Cell. Proteomics* **2013**, *12* (12), 3583–3598.

(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* (1), 51–61.

(11) 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* (4), 535–553.

(12) Qu, J.; Young, R.; Page, B. J.; Shen, X.; Tata, N.; Li, J.; Duan, X.; Fallavollita, J. A.; Canty, J. M., Jr. Reproducible Ion-Current-Based Approach for 24-Plex Comparison of the Tissue Proteomes of Hibernating versus Normal Myocardium in Swine Models. *J. Proteome Res.* **2014**, *13* (5), 2571–2584.

(13) Tu, C.; Li, J.; Sheng, Q.; Zhang, M.; Qu, J. Systematic Assessment of Survey Scan and MS2-Based Abundance Strategies for Label-Free Quantitative Proteomics Using High-Resolution MS Data. *J. Proteome Res.* **2014**, *13* (4), 2069–2079.

(14) Callister, S. J.; Barry, R. C.; Adkins, J. N.; Johnson, E. T.; Qian, W. J.; Webb-Robertson, B. J.; Smith, R. D.; Lipton, M. S. Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics. *J. Proteome Res.* **2006**, *5* (2), 277–286.

(15) Rudnick, P. A.; Wang, X.; Yan, X.; Sedransk, N.; Stein, S. E. Improved normalization of systematic biases affecting ion current measurements in label-free proteomics data. *Mol. Cell. Proteomics* **2014**, *13* (5), 1341–1351.

(16) Yang, Y. H.; Dudoit, S.; Luu, P.; Lin, D. M.; Peng, V.; Ngai, J.; Speed, T. P. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* **2002**, *30* (4), e15.

(17) Bolstad, B. M.; Irizarry, R. A.; Astrand, M.; Speed, T. P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **2003**, *19* (2), 185–193.

(18) Karpievitch, Y. V.; Dabney, A. R.; Smith, R. D. Normalization and missing value imputation for label-free LC–MS analysis. *BMC Bioinf.* **2012**, *13* (Suppl 16), S5.

(19) Kulima, K.; Nilsson, A.; Scholz, B.; Rossbach, U. L.; Falth, M.; Andren, P. E. Development and evaluation of normalization methods for label-free relative quantification of endogenous peptides. *Mol. Cell. Proteomics* **2009**, *8* (10), 2285–2295.

(20) Cox, J.; Hein, M. Y.; Luber, C. A.; Paron, I.; Nagaraj, N.; Mann, M. MaxLFQ allows accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction. *Mol. Cell. Proteomics* **2014**, *13*, 2513–2526.

(21) Carrillo, B.; Yanofsky, C.; Laboissiere, S.; Nadon, R.; Kearney, R. E. Methods for combining peptide intensities to estimate relative protein abundance. *Bioinformatics* **2010**, *26* (1), 98–103.

(22) Wilhelm, M.; Schlegl, J.; Hahne, H.; Moghaddas Gholami, A.; Lieberenz, M.; Savitski, M. M.; Ziegler, E.; Butzmann, L.; Gessulat, S.; Marx, H.; Mathieson, T.; Lemeier, S.; Schnatbaum, K.; Reimer, U.; Wenschuh, H.; Mollenhauer, M.; Slotta-Huspenina, J.; Boese, J. H.; Bantscheff, M.; Gerstmair, A.; Faerber, F.; Kuster, B. Mass-spectrometry-based draft of the human proteome. *Nature* **2014**, *509* (7502), 582–587.

(23) L, I. J.; Stoop, M. P.; Stingl, C.; Sillevius Smitt, P. A.; Luidert, T. M.; Dekker, L. J. Comparative study of targeted and label-free mass spectrometry methods for protein quantification. *J. Proteome Res.* **2013**, *12* (4), 2005–2011.

(24) Park, S. K.; Venable, J. D.; Xu, T.; Yates, J. R., 3rd. A quantitative analysis software tool for mass spectrometry-based proteomics. *Nat. Methods* **2008**, *5* (4), 319–322.

(25) 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.

(26) Grossmann, J.; Roschitzki, B.; Panse, C.; Fortes, C.; Barkow-Oesterreicher, S.; Rutishauser, D.; Schlapbach, R. Implementation and evaluation of relative and absolute quantification in shotgun proteomics with label-free methods. *J. Proteomics* **2010**, *73* (9), 1740–1746.

(27) 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* (5), 1214–1223.

(28) Qian, W. J.; Kaleta, D. T.; Petritis, B. O.; Jiang, H.; Liu, T.; Zhang, X.; Mottaz, H. M.; Varnum, S. M.; Camp, D. G., 2nd; Huang, L.; Fang, X.; Zhang, W. W.; Smith, R. D. Enhanced detection of low abundance human plasma proteins using a tandem IgY12-SuperMix immunoaffinity separation strategy. *Mol. Cell. Proteomics* **2008**, *7* (10), 1963–1973.

(29) 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* (6), 2838–2850.

(30) 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* (2), 761–776.

(31) 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* (12), 1367–1372.

(32) Cox, J.; Neuhauser, N.; Michalski, A.; Scheltema, R. A.; Olsen, J. V.; Mann, M. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **2011**, *10* (4), 1794–1805.

(33) 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* (9), 667–675.

(34) Smyth, G. K.; Speed, T. Normalization of cDNA microarray data. *Methods* **2003**, *31* (4), 265–273.

(35) Holm, S. A simple sequentially rejective multiple test procedure. *Scand J. Stat.* **1979**, *6*, 65–70.

(36) Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **1995**, *57*, 289–300.

(37) Fisher, R. A. *Statistical Methods for Research Workers*; Oliver and Boyd: Edinburgh, 1932.

(38) Carvalho, P. C.; Han, X.; Xu, T.; Cociorva, D.; Carvalho Mda, G.; Barbosa, V. C.; Yates, J. R., 3rd. XDIA: improving on the label-free data-independent analysis. *Bioinformatics* **2010**, *26* (6), 847–848.

(39) Gautier, V.; Mouton-Barbosa, E.; Bouyssie, D.; Delcourt, N.; Beau, M.; Girard, J. P.; Cayrol, C.; Bulet-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* (8), 527–539.

(40) Grubbs, F. Procedures for Detecting outlying Observations in Samples. *Technometrics* **1969**, *11* (1), 1–21.

(41) Zybailov, B.; Mosley, A. L.; Sardi, M. E.; Coleman, M. K.; Florens, L.; Washburn, M. P. Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*. *J. Proteome Res.* **2006**, *5* (9), 2339–2347.

(42) 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* (9), 1265–1272.

(43) Arike, L.; Valgepea, K.; Peil, L.; Nahku, R.; Adamberg, K.; Vilu, R. Comparison and applications of label-free absolute proteome quantification methods on *Escherichia coli*. *J. Proteomics* **2012**, *75* (17), 5437–5448.

(44) Hebert, A. S.; Richards, A. L.; Bailey, D. J.; Ulbrich, A.; Coughlin, E. E.; Westphall, M. S.; Coon, J. J. The one hour yeast proteome. *Mol. Cell. Proteomics* **2014**, *13* (1), 339–347.