

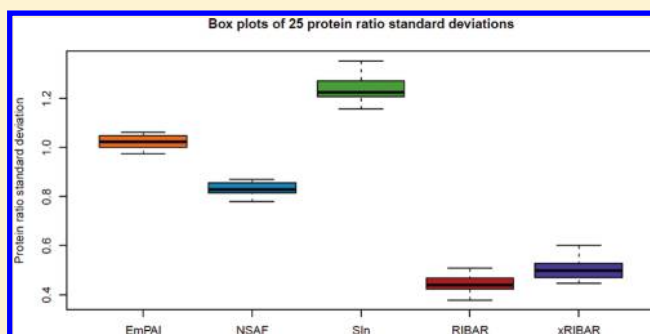
RIBAR and xRIBAR: Methods for Reproducible Relative MS/MS-based Label-Free Protein Quantification

Niklaas Colaert,^{†,‡} Kris Gevaert,^{†,‡} and Lennart Martens^{*,†,‡}[†]Department of Medical Protein Research, VIB, B-9000 Ghent, Belgium[‡]Department of Biochemistry, Ghent University, B-9000 Ghent, Belgium Supporting Information

ABSTRACT: Mass spectrometry-driven proteomics is increasingly relying on quantitative analyses for biological discoveries. As a result, different methods and algorithms have been developed to perform relative or absolute quantification based on mass spectrometry data. One of the most popular quantification methods are the so-called label-free approaches, which require no special sample processing, and can even be applied retroactively to existing data sets. Of these label-free methods, the MS/MS-based approaches are most often applied, mainly because of their inherent simplicity as compared to MS-based methods. The main application of these approaches is the determination of relative protein amounts between different samples, expressed as protein ratios. However, as we demonstrate here, there are some issues with the reproducibility across replicates of these protein ratio sets obtained from the various MS/MS-based label-free methods, indicating that the existing methods are not optimally robust. We therefore present two new methods (called RIBAR and xRIBAR) that use the available MS/MS data more effectively, achieving increased robustness. Both the accuracy and the precision of our novel methods are analyzed and compared to the existing methods to illustrate the increased robustness of our new methods over existing ones.

As a result, different methods and algorithms have been developed to perform relative or absolute quantification based on mass spectrometry data. One of the most popular quantification methods are the so-called label-free approaches, which require no special sample processing, and can even be applied retroactively to existing data sets. Of these label-free methods, the MS/MS-based approaches are most often applied, mainly because of their inherent simplicity as compared to MS-based methods. The main application of these approaches is the determination of relative protein amounts between different samples, expressed as protein ratios. However, as we demonstrate here, there are some issues with the reproducibility across replicates of these protein ratio sets obtained from the various MS/MS-based label-free methods, indicating that the existing methods are not optimally robust. We therefore present two new methods (called RIBAR and xRIBAR) that use the available MS/MS data more effectively, achieving increased robustness. Both the accuracy and the precision of our novel methods are analyzed and compared to the existing methods to illustrate the increased robustness of our new methods over existing ones.

KEYWORDS: quantitative proteomics, label-free, quality control



1. INTRODUCTION

In modern high-throughput proteomics studies, shotgun techniques and high-end instruments allow thousands of peptides and proteins to be identified with relative ease. These peptide identifications and their inferred proteins provide a static snapshot of (a part of) the proteome of the studied biological system.¹ Such snapshots can be made dynamic and more informative by relying on quantitative proteomic techniques that can establish (relative) amounts of proteins in similar samples. Quantitative approaches for shotgun experiments broadly fall into two categories: label-based and label-free approaches.² In the former, peptides are labeled with an isobaric or isotopic label prior to MS and MS/MS analysis. The latter, on the other hand, do not require these additional, sometimes time-consuming, and often expensive protocol steps. Additionally, the absence of specialized sample treatment also allows the label-free methods to be applied retroactively to previously acquired data. The different label-free methods can provide both absolute protein concentrations as well as relative protein concentration changes, based on information from either the MS or the MS/MS spectra acquired. Of these, MS/MS-based label-free methods are typically much easier to perform because they rely on information already present for the peptide identification procedure, whereas MS-based label-free methods require complex and CPU

intensive algorithms for chromatogram alignment, accurate MS peak picking, and peptide intensity calculation.²

Several different label-free MS/MS-based methods have been developed over the past few years to perform MS/MS-based label-free quantification. These methods rely on different aspects (or a combination of aspects) of the MS/MS information including peptide identification count, identified spectrum count, and the sum of MS/MS fragment ion intensity. Typical examples of the various methods in use are provided by the NSAF (Normalized Spectral Abundance Factor)³ method for spectrum counting, the emPAI (exponentially modified Protein Abundance Index)⁴ method for peptide counting, and the SIN (normalized Spectral Intensity)⁵ method for MS/MS fragment ion intensity summing. It is of note that these three algorithms were initially developed to estimate the absolute protein abundance in an experiment. However, if two experiments are performed, the corresponding protein abundance change between the two experiments can be calculated by simply taking the ratio of both absolute protein abundance estimations. As we will show here, however, there are important caveats relating to relative label-free quantification for these three approaches,

Received: March 9, 2011

Published: May 12, 2011

which include the generation of discontinuous ratio distributions, and a lack of robustness in obtaining protein ratios between samples when the proteins were identified by different peptide sets in the different experiments. To address these issues, we have developed two new methods for relative label-free quantification, which are presented and carefully benchmarked here. Our results show that our novel methods exhibit better robustness than existing methods, and we further validate our approaches by examining the precision and accuracy of their results within and across sets. Throughout, we compare our algorithms with the three existing algorithms outlined above.

2. EXPERIMENTAL SECTION

2.1. Data Sets

Two publicly available data sets were used in this study. The first data set is a study of the NCI funded CPTAC (Clinical Proteomic Technology Assessment for Cancer) Network. CPTAC benchmarks and develops standardized technologies and methodologies to enable researchers to conduct and compare protein research.^{6,7} In the sixth study of the CPTAC network, identical samples were analyzed by different laboratories on different instruments to benchmark and analyze the interlaboratory variation and to enable proteomic quality assessment.^{6,7} Here, only raw data results from one laboratory (LTQ-Orbitrap@86, CPTAC experiment identification number 104, Tranche download hash: NGX3cBUAZXSWvc+6XF-NIdVhpLPJT087lzAxUQmwwR2KHUwWDrFwV1dso3bvxf-7HeXZ4C/juqwEUIz4boC9H3HcLrxEAAAAAAAAAmDw==) were used to do the analyses.^{6,7} In this study, five yeast lysates were spiked with different concentrations of the Sigma UPS-1 mixture of 48 human proteins. These samples were analyzed in triplicate on an LTQ-Orbitrap XL generating three data sets for each spiked-in concentration of the Sigma UPS-1 mixture. Only the two highest spiked-in UPS-1 mixture concentrations (yeast digests (60 ng/ μ L) spiked in with either 6.7 fmol/ μ L or 20 fmol/ μ L of the Sigma UPS-1 mixture) were used here as these generated the most UPS-1 protein identifications yielding nine binary comparisons between the two triplicates. Detailed information concerning the yeast lysate, digestion and mass spectrometric analysis can be found in.^{6,7} Mgf files were generated from the raw Thermo results files with DtaSuperCharger.⁸

Identification of the mgf spectra was performed according to the original protocol as described in^{6,7} using a local installation of Mascot (Version 2.2.04) and the yeast UniprotKB/Swiss-Prot protein database (version 15.14), supplemented with the 48 protein sequences of the UPS-1 mixture. Only doubly and triply charged tryptic peptides with at most one allowed missed cleavage were considered for identification. Precursor mass tolerance was set to 10 ppm and fragment ion mass tolerance to 0.5 Da. The identification confidence was set to 0.05 and variable modifications were set to: acetylation of the protein N-terminus, oxidation of methionine, pyro-glutamate formation for N-terminal glutamine, carbamidomethyl cysteine formation and pyro-carbamidomethyl cysteine formation for N-terminal cysteine. No fixed modifications were set.

The second data set is one generated by The Proteome Informatics Research Group (iPRG) of the Association of Biomolecular Resource Facilities (ABRF) for their 2009 study. The goal of this study was to compare and benchmark the reproducibility of the observed differences between two complex samples. Here, an *E. coli* lysate (spiked in with bovine serum albumin)

was separated in two sets of multiple lanes on a 1D SDS-PAGE gel and from each set of lanes a different part of the gel was excised. Following protein band excision, pooling and in-gel trypsin digestion, five technical replicates were analyzed on a LTQ-Orbitrap XL mass spectrometer for both sets. Twenty-five pair wise comparisons can therefore be calculated by making use of the five replicates of the two different samples. Detailed information concerning this study and the protocol can be found on <http://www.abrf.org/index.cfm/group.show/ProteomicsInformaticsResearchGroup.53.htm>. The mgf files were downloaded via Tranche (download hash: m680i/1/SlvignH2tMib+3jUsCzV+hrXBcTKp4OOQsh0Y5G21a+zvzRpWh9BDz9no8OOf/YaDz5xbEc3/yFnLNAeskAAAAAAAAAKKg==). Identification was performed using a local installation of Mascot (Version 2.2.04) and the protein database provided by the iPRG containing UniProtKB/Swiss-Prot *E. coli* proteins, common contaminants and bovine serum albumin, along with the concatenated decoy database consisting of the reversed sequences. Precursor mass tolerance was set to 10 ppm and fragment ion mass tolerance to 0.8 Da. Only doubly and triply charged tryptic peptides with one allowed missed cleavage were considered for identification. The identification confidence was set to 0.05 and carbamidomethyl cysteine formation was set as the fixed modification, while the allowed variable modifications were oxidation of methionine and pyro-glutamate formation for N-terminal glutamine.

2.2. Protein Quantification

Three existing MS/MS-based label-free methods were used to calculate protein ratios for the two data sets. These methods were originally designed for absolute protein abundance estimation, but relative protein abundance differences were estimated by calculating the ratio of the absolute estimations between samples. The Exponentially Modified Protein Abundance Index (emPAI) is a peptide counting method, based on the Protein Abundance Index (PAI) that is calculated by dividing the number of different observed modified peptides by the number of theoretically observable unmodified peptides. A modified peptide is a peptide that carries an amino acid modification (e.g., methionine oxidation, N-terminal acetylation). Peptides with the same peptide sequence but with a different modification status are seen as two different modified peptides. An observable peptide is then defined as any unmodified tryptic peptide with a mass between the smallest and largest mass among the identified peptides. The PAI value is then exponentially modified ($10^{\text{PAI} - 1}$) to obtain the emPAI value. The protein content (mol %) is calculated by normalizing the emPAI by the sum of the emPAI values for all proteins in the set.⁴ The normalized spectral abundance factor (NSAF) is a spectrum counting method. The spectral abundance is calculated by counting the number of identified spectra and then normalizing this count by the protein length in residues. This abundance is then further normalized by the sum of all the protein abundances in the set.³ The spectral index normalized (SIn), finally, uses fragment ion intensity count, along with spectral counting. A spectral index is calculated per protein by taking the sum of the summed intensities of the matched fragment ions across all spectra assigned to that protein. This spectral index is then divided by the sum of all protein spectral indexes.⁵

In addition to the three existing methods, two new methods were used to calculate protein ratios for the two data sets. The RIBAR method (see further) relies only on peptides found in

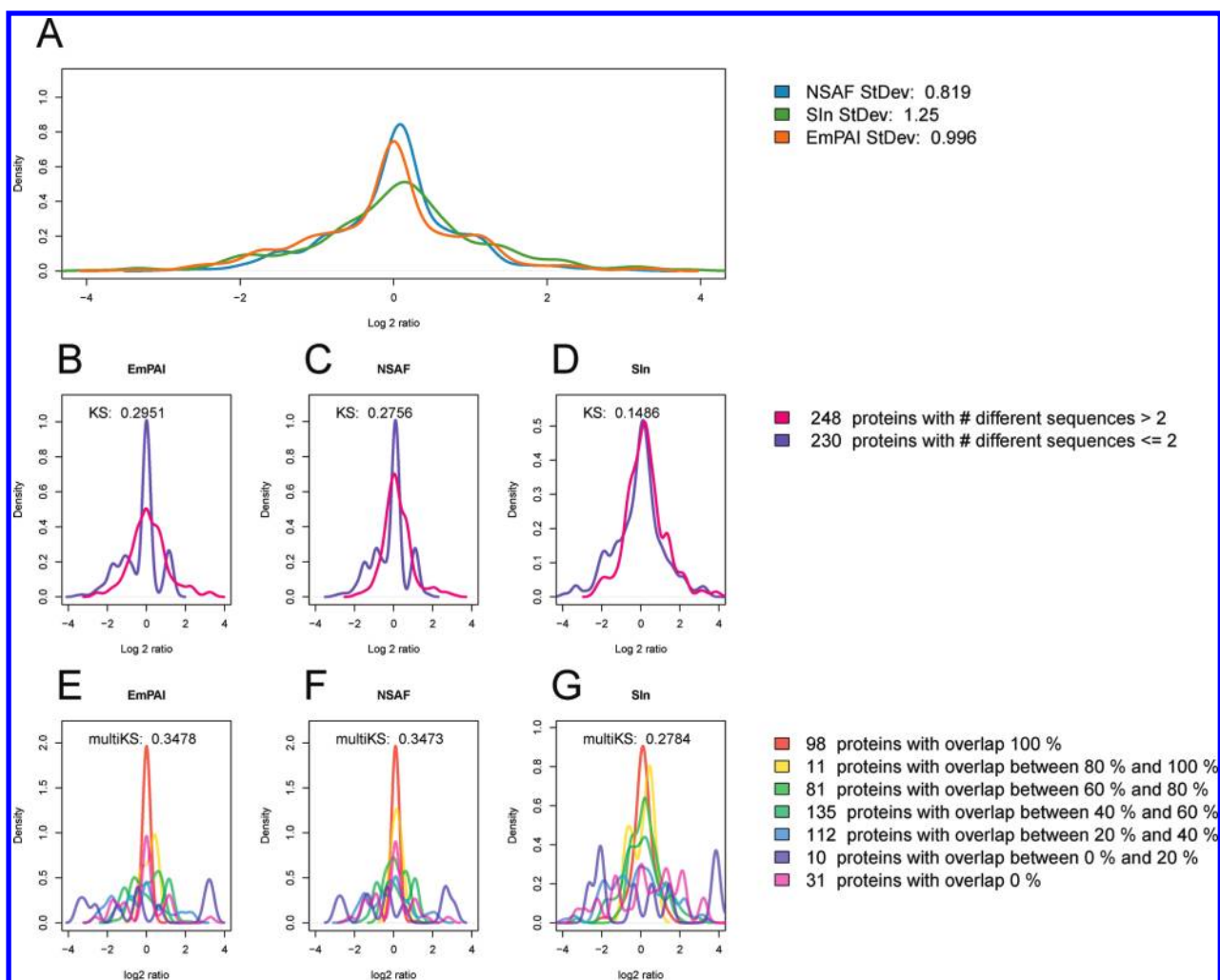


Figure 1. Data used in this plot represents one of the twenty-five pairwise comparisons obtained from the iPRG data set, here replicate one from set one and replicate one from set two are used. Identical plots for the other 24 comparisons are given in Supporting Information. (A) Protein ratios calculated by the different quantification methods as a density plot. The protein ratio standard deviation is indicated in the legend. (B, C and D) Protein ratio estimations are divided in two groups with different number of identified peptides. The first group consists of ratios from proteins with only one or two distinct identified peptide sequences, and the second group from proteins with three or more distinct identified peptide sequences. The Kolmogorov–Smirnov distance between these two distributions is indicated at the top of each chart. (E, F and G) Distributions of ratios divided in seven groups based on different percentages of overlapping peptides. This percentage is calculated by dividing the number of shared unique modified peptides (found in both sets) by the sum of the total number of unique modified peptides in the first and second set. The average pairwise Kolmogorov–Smirnov (multiKS) distances of the different protein ratio distributions are indicated above the chart.

both samples. For each sample, the MS2 intensities of all peaks of each unique modified peptide for all of its peptide-to-spectrum matches are summed. A peptide ratio for each unique modified peptide is then calculated by using these abundance estimations. By averaging the calculated log₂ modified peptide ratios for a protein, a log₂ ratio for that protein is obtained. In the xRIBAR method (see further) an extra ratio for each protein is obtained by calculating the ratio of the average summed MS2 intensities in all spectra linked to the protein. This extra ratio is either added to the list of modified peptide ratios obtained from overlapping peptides, or it is used as the sole metric for proteins that have no overlapping peptides between samples.

3. RESULTS

We evaluate the operational characteristics of three common MS/MS-based relative label-free method by comparing three

simple metrics of their performance as applied to the iPRG data sets. These metrics are summarized in Figure 1 and explained below. The first metric shows the variation of protein ratios within one result set. Since only a small fraction of proteins are typically regulated in quantitative proteomic experiments, most proteins are unchanged and will have ratios similar to one. The standard deviation of the protein ratios distribution within an experiment therefore can be used to indicate the precision of the method. Obviously, this method relies on the assumption that the majority of proteins are detected at a fixed ratio (e.g., 1/1). The protein ratio distribution and standard deviation for the three methods are given in Figure 1A for one of the 25 iPRG comparisons (the additional 24 plots are given in Supporting Information). A density plot and a box plot for the 25 standard deviations can be found in Figure 2A,B. For this metric, SIn shows a larger standard deviation than the other two methods, indicating lower overall precision.

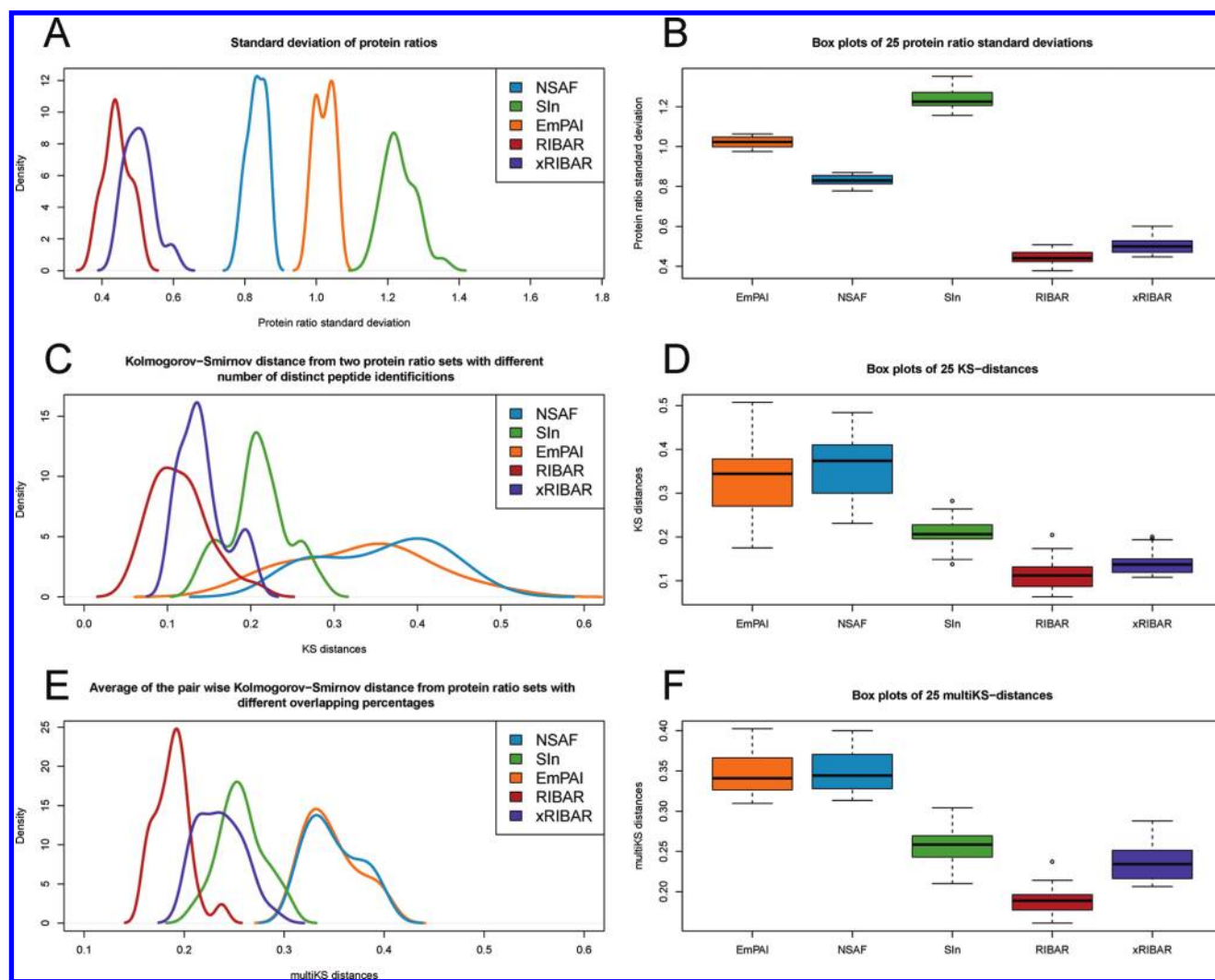


Figure 2. Protein ratio standard deviations calculated for the twenty-five pairwise comparisons for the iPRG data set are plotted as a (A) density distribution and (B) box plot. The 25 Kolmogorov–Smirnov distances between the two groups with different numbers of identified peptides are plotted as a (C) density distribution and (D) box plot. The 25 averaged pairwise Kolmogorov–Smirnov distances calculated for the 7 ratio groups with different protein overlapping percentages are plotted as a (E) density plot and (F) box plot. See main text for details.

The second metric tries to estimate the effect of working with a small set of discrete values in protein abundance estimation. This is important since spectral or peptide counting methods tend to work with a small number of discrete integer values (ranging from 1 to 2 spectra to a few tens of spectra). These discrete values are commonly normalized either by the length of the proteins or by the number of possible identifiable peptides. A second normalization procedure normalizes across the data set, usually by dividing the individual normalized protein values by the sum of all normalized protein values in the data set. The end result of these steps is typically a discontinuous protein ratio distribution (e.g., 0.5, 1.0, 1.5, 2.0, ...). To test the effects of this idiosyncrasy of the ratio distribution, the protein ratios calculated for the iPRG study were divided in two groups based on the number of distinct matching peptides. One group consisted of ratios coming from proteins with one or two distinct identified peptides, while the second group consisted of ratios coming from proteins with three or more distinct identified peptides. The emPAI, NSAF and SIn protein ratio distributions for 1 of the 25 iPRG comparisons are given in Figure 1B, C and D, respectively (the additional 24 plots

are given in Supporting Information). The spectrum and peptide counting methods (NSAF and emPAI) clearly have discontinuous ratio distributions for the group of proteins with only one or two distinct identified peptides. In order to quantify the difference between the two protein ratio distributions for each method, a Kolmogorov–Smirnov distance is calculated (indicated as the KS metric in Figure 1B, C and D). A density plot and a box plot for the 25 KS distances can be found in Figure 2C,D. Interestingly, the protein ratio distribution of the SIn method is continuous because it uses the fragment ion intensities within the MS2 spectra rather than their count, yielding a quasi-continuous distribution of protein ratios. By comparing the KS distance for the NSAF and emPAI data with the SIn data, it is clear that protein ratios calculated by NSAF and emPAI are differently affected when compared to SIn because of their reliance on discrete values derived from the number of matching spectra or peptides.

The third and last metric investigates the effect of identifying a protein with different sets of peptides in each sample. The protein abundance in each of the compared samples is then

derived from different peptides that can all possess different MS properties. In such a situation, the basic assumption that the protein quantification method is independent of the sample studied can be violated. Indeed, if different peptide sequences are used to identify a protein in different samples, a different number of MS/MS spectra can be recorded for that protein in each sample even though the protein is present in equal amounts, for instance because the peptides identified in one sample tend to ionize better than the ones identified in the other sample. Of course, the same logic applies to summed fragment ion intensities for the various peptides. To investigate the impact of this unwanted variation, the protein ratios calculated for the iPRG study were divided in seven different groups based on how strongly their identifying peptides overlap between samples. A percentage is calculated by dividing the number of shared unique modified peptides (identifying the protein in both samples) by the sum of the total number of unique modified peptides in the first and second set. A high percentage overlap thus represents a larger degree of similarity between the protein abundance estimation in each sample. Seven different groups are obtained by splitting the data in decreasing steps of 20% overlap, and maintaining 100% and 0% as separate categories. The protein ratio distributions for these seven groups are plotted for NSAF, emPAI and SIn in Figure 1E, F and G, respectively for one of the 25 iPRG comparisons (the other 24 are provided in Supporting Information). The average pairwise Kolmogorov–Smirnov (multiKS) distance between the different protein ratio distributions is indicated in Figure 1E, F and G. A density plot and a box plot for the 25 multiKS distances can be found in Figure 2E,F. Taken together, the plots and metrics indicate that proteins with a different percentage of overlap between the two samples have different protein ratio distributions. Indeed, the plots show broader distributions for proteins with less overlap. On the basis of the multiKS distance, however, it can be seen that the SIn method has a more consistent performance than NSAF and emPAI.

The three metrics taken together indicate that no single method clearly outperforms the others. SIn is affected by a large protein ratio variation whereas NSAF and emPAI are influenced by discrete protein abundance values and show a decreased performance for proteins with low overlap.

To optimize the performance of label-free proteomics across these three metrics, we propose two new relative label-free protein quantification methods that robustly combine the various strengths of the existing algorithms.

The first method, called robust intensity based averaged ratio (RIBAR), only relies on peptides found in both samples to calculate a protein ratio. First, the abundance of each unique modified peptide is calculated for each sample by adding up the summed MS2 intensities of all peaks in all of its peptide-to-spectrum matches. These abundances then allow a peptide ratio to be calculated for each modified peptide. By subsequently averaging the calculated log₂ modified peptide ratios for a protein, a log₂ ratio for that protein is obtained. Since this method will rely exclusively on peptides found in both samples, it cannot calculate ratios for proteins without overlapping peptides. To compensate for this loss of quantifiable proteins, a second, derived method was developed, called extended robust intensity based averaged ratio (xRIBAR). Here, an extra ratio for each protein is obtained by simply calculating the ratio of the average summed MS2 intensities in all spectra matching to the protein. This extra ratio is either added to the list of modified peptide

ratios obtained from overlapping peptides, or it is used as the sole metric for proteins that have no overlapping peptides between samples. As such, this derived method ensures that a protein ratio can be calculated for every protein that can be identified in both samples, even if there is no peptide overlap. For proteins with at least some overlapping peptides, we can correlate the RIBAR with the xRIBAR result. As shown in Supplementary Figure 25 (Supporting Information), this correlation is nearly perfect (Pearson correlation of 0.99).

The three metrics previously used to benchmark the robustness of the existing methods have also been used to test the robustness of our two proposed methods. The iPRG data set is again used for benchmarking, but the results across the twenty-five possible pairwise comparisons are now summarized in a single plot rather than presented individually. Figures 2A and B plot the protein ratio standard deviations calculated by the three existing and two new methods for the twenty-five comparisons. Both the RIBAR and xRIBAR methods outperform the other methods, displaying overall much smaller standard deviations. Figures 2C and D plot the Kolmogorov–Smirnov distance between the distribution of ratios of proteins with only one or two distinct modified peptides, and the distribution of proteins with three or more distinct modified peptides. RIBAR has the lowest Kolmogorov–Smirnov distances and thus outperforms existing methods for this metric as well. xRIBAR is slightly worse than RIBAR, but it still outperforms SIn, that in turn leaves NSAF and emPAI behind. Figures 2E and F then show the average pairwise Kolmogorov–Smirnov (multiKS) distance of the ratio distributions of the seven different groups with different peptide overlapping percentage (see above for grouping details). RIBAR and xRIBAR again outperform the existing methods.

The two methods proposed here thus demonstrate their overall robustness by scoring consistently better than the existing methods. The crucial innovations in these two methods are the removal of effects derived from low peptide overlap between samples, and the use of the practically continuous distribution of abundance values offered by the total MS2 spectral intensity. The standard deviations for the individual protein ratios (Figure 2A and B) are not negatively affected by making these methods robust; rather the contrary: the standard deviation is lower for the new methods. The RIBAR method consistently scores better on the three metrics than its sibling xRIBAR, but is hampered by its inability to calculate a protein ratio for those proteins that do not have overlapping peptides across the samples. Overall however, the performance difference between RIBAR and xRIBAR is much smaller than the difference in performance between these two methods and the existing approaches. The slightly reduced performance of the xRIBAR method is thus compensated by its ability to provide a robust protein ratio for every protein that was identified across the samples, regardless of its peptide-level overlap.

The robustness of existing and new methods was analyzed, clearly indicating that NSAF and emPAI are the least robust methods. SIn outperformed the two existing methods when the robustness is analyzed; however, the standard deviations for the protein ratios were larger for SIn than for NSAF and emPAI. The new methods presented here show a clear improvement for the three analyzed metrics, indicating that both RIBAR and xRIBAR perform better than the existing methods. RIBAR performs slightly better than xRIBAR but the RIBAR method is unable to calculate a protein ratio for 7.8% of the proteins on average for the iPRG data set when compared to xRIBAR and the existing methods.

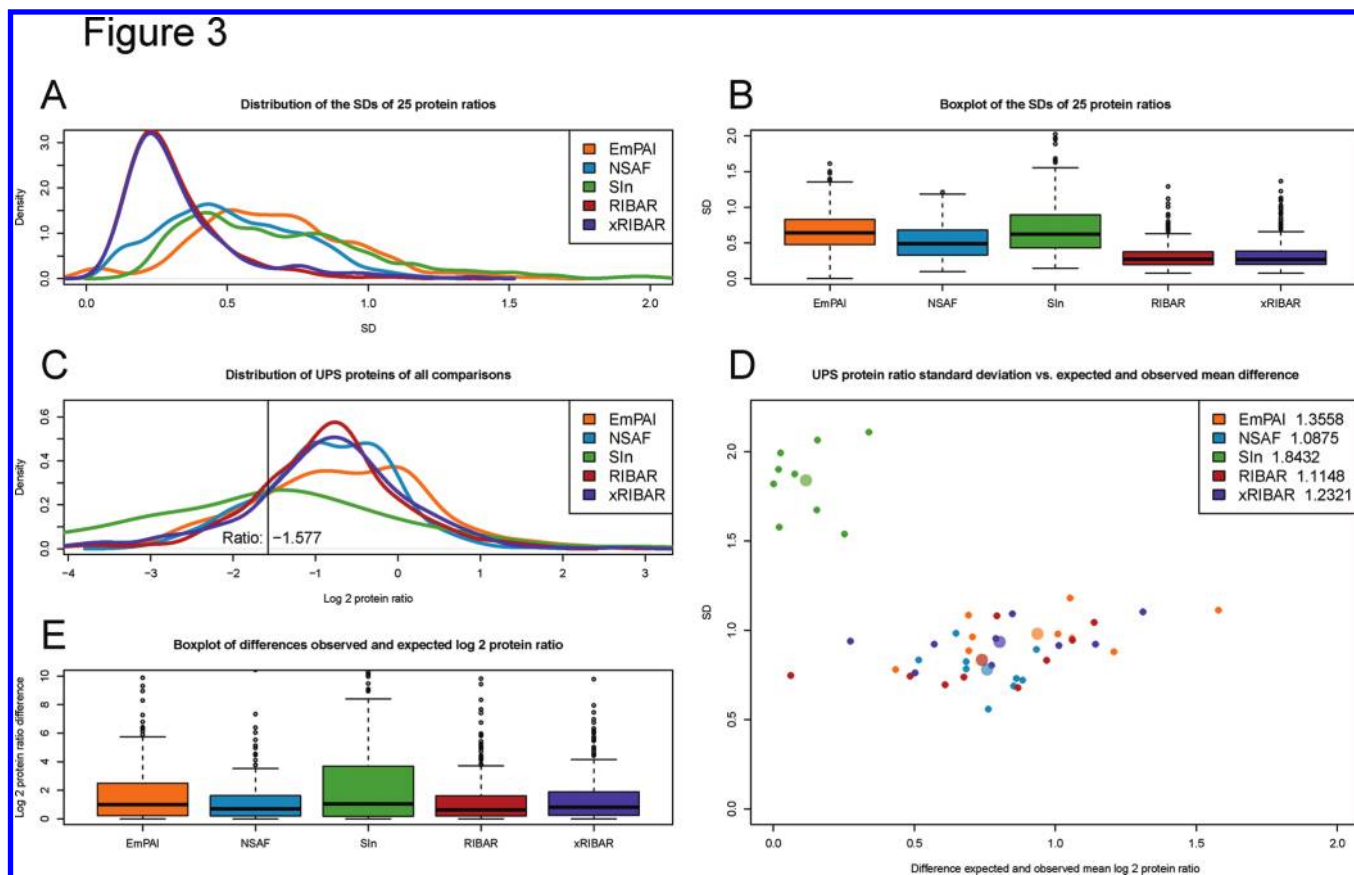


Figure 3. (A and B) Plot standard deviations for the 380 protein ratios that were found in all 25 pairwise comparisons from the iPRG study as a density plot and a box plot, respectively. (C) Three-hundred forty-seven protein ratios (from nine pairwise comparisons) of the UPS-1 proteins with an expected log₂ ratio of -1.577 in the CPTAC study as a density plot. (D) Standard deviation of the UPS-1 protein ratios and the difference from the expected and observed mean of these proteins for the 9 pairwise comparisons are plotted as a scatter plot. The bigger dot represents the average of the standard deviation and the mean for the 9 comparisons. The distance of this average to the origin is indicated in the legend. (E) Squared difference of the expected and observed UPS-1 protein ratios as a box plot.

4. DISCUSSION

Label-free methods provide an efficient approach to relative quantitative proteomics, as they base their calculations on readily available MS/MS data. The most commonly used methods are emPAI (exponentially modified protein abundance index),⁴ NSAF (normalized spectral abundance factor)³ and SIn (spectral index normalized),⁵ building on algorithms for peptide counting, spectral counting and a mix of peptide, spectral and ion counting, respectively. We have here analyzed these existing methods using three distinct and complementary metrics to uncover their relative strengths and shortcomings. The results indicate that NSAF and emPAI are the least robust methods, but while SIn outperformed these two methods overall, the standard deviations for the protein ratios were larger for SIn than for NSAF and emPAI, thus showing no clear optimal algorithm overall. We therefore proposed two related new methods, RIBAR and xRIBAR, that calculate protein ratios based on peptide spectral intensities from peptides identified across the compared samples. These methods show a clear improvement for the three key metrics analyzed here.

In a recent study, we compared the quantification quality of the three existing methods.⁹ Four different metrics were proposed and tested in that study to allow the benchmarking of algorithm performance. We have also tested these 4 metrics for the new methods and compare the results against the existing

methods. First of all, the reproducibility across different experiments was evaluated for the 5 methods. A standard deviation was calculated for each protein that had been quantified in all 25 pairwise iPRG comparisons. This standard deviation is a measurement for the (technical) reproducibility of the protein ratio calculation. Indeed, a lower standard deviation indicates that repeated protein ratio estimations are more reproducible. The standard deviations from the 380 proteins that were quantified across all 25 pairwise comparisons in the iPRG data set are plotted in Figure 3A and B. Both the RIBAR and the xRIBAR methods estimate the protein ratios more consistently than the existing methods, as indicated by the smaller standard deviations.

Second, the accuracy of the protein ratio estimations was tested. This was done by looking at the protein ratios for the spiked in Sigma UPS-1 proteins in the CPTAC data set. Nine pairwise comparisons can be made between the two triplicate analyses of a yeast digest with either 6.7 fmol/ μ L or 20 fmol/ μ L of spiked in Sigma UPS-1. The Sigma UPS-1 proteins should have a protein ratio of 0.335 (6.7/20), translating into a log₂ protein ratio of -1.58 . The distributions of the UPS-1 protein ratios for all nine comparisons are plotted in Figure 3C. On average, SIn seems to perform best here, since the mean of the Sigma UPS-1 protein ratios is centered on the expected protein ratio (indicated by a black vertical line in the plot). However, this

second metric only reflects the overall accuracy of the experiment, and therefore ignores the accuracy of individual protein quantification. This accuracy of individual proteins is reflected by a third metric: the standard deviation across the estimated ratios for the Sigma UPS-1 proteins. A low standard deviation implies a consistent ratio for each Sigma UPS-1 protein, whereas a large standard deviation indicates large differences in estimated ratio for each protein. Obviously, the Sigma UPS-1 spike is constant, and the standard deviation across the proteins should correspondingly be small. Figure 3D plots the nine standard deviations (one for each pairwise comparison) versus the difference between the mean of the protein ratios and the known spike-in value in a scatter plot. Note that we are thus plotting within-experiment precision against within-experiment accuracy. A bigger dot is furthermore plotted to indicate the average for each cluster of nine different comparisons. The distance from this dot to the origin (smaller is better) is given in the legend. The scatter plot clearly shows that although SIn has very good accuracy, it has very poor precision when compared to the other methods. The RIBAR and xRIBAR methods outperform SIn in terms of precision, and are nearly indistinguishable from emPAI and NSAF in terms of accuracy and precision.

Lastly, the accuracy on the protein level can be studied by looking at the squared difference between the expected ratio value and the estimated protein ratio. The squared differences are plotted for the Sigma UPS-1 proteins as a box plot in Figure 3E. The five methods all perform similarly, but RIBAR and xRIBAR have slightly narrower (and thus more precise) distributions.

In this paper, we outline some issues with the robustness of protein quantification by MS2-based label-free relative protein quantification methods, and we proposed two new methods that are not affected by these issues. We show that our methods are indeed more robust than existing methods, and that their overall quantification precision and accuracy is identical to, or better than existing methods on both the individual protein as well as the whole experiment-level. Furthermore, our methods achieve better reproducibility across multiple experiments the RIBAR and xRIBAR methods presented here thus provide the most robust, yet accurate and precise protein ratio estimations, while ensuring optimal across-experiment reproducibility. RIBAR and xRIBAR are implemented as freely available open source software under the permissive Apache2 license and can be downloaded from <http://code.google.com/p/compomics-ribar/>.

■ ASSOCIATED CONTENT

5 Supporting Information

Supplementary Figures 1–25. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*Professor Lennart Martens, Department of Medical Protein Research and Biochemistry, VIB and Faculty of Medicine and Health Sciences, Ghent University, A. Baertsoenkaai 3, B-9000 Ghent, Belgium. Email: lennart.martens@ugent.be. Fax: +32-92649484.

■ ACKNOWLEDGMENT

We thank Joël Vandekerckhove for his support and the NCI CPTAC project and the ABRF iPRG for making their data sets

publicly available. We acknowledge the support of Ghent University (Multidisciplinary Research Partnership “Bioinformatics: from nucleotides to networks”). We declare no conflict of interest.

■ ABBREVIATIONS:

NSAF, normalized spectral abundance factor; emPAI, exponentially modified protein abundance index; SIn, normalized spectral intensity.

■ REFERENCES

- (1) Domon, B.; Aebersold, R. Mass spectrometry and protein analysis. *Science* **2006**, *312* (5771), 212–217.
- (2) Vaudel, M.; Sickmann, A.; Martens, L. Peptide and protein quantification: a map of the minefield. *Proteomics* **2010**, *10* (4), 650–670.
- (3) Paoletti, A. C.; Parmely, T. J.; Tomomori-Sato, C.; Sato, S.; Zhu, D.; Conaway, R. C.; Conaway, J. W.; Florens, L.; Washburn, M. P. Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (50), 18928–18933.
- (4) 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.
- (5) 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* (1), 83–89.
- (6) Paulovich, A. G.; Billheimer, D.; Ham, A. J.; Vega-Montoto, L.; Rudnick, P. A.; Tabb, D. L.; Wang, P.; Blackman, R. K.; Bunk, D. M.; Cardasis, H. L.; Clauser, K. R.; Kinsinger, C. R.; Schilling, B.; Tegeler, T. J.; Variyath, A. M.; Wang, M.; Whiteaker, J. R.; Zimmerman, L. J.; Fenyo, D.; Carr, S. A.; Fisher, S. J.; Gibson, B. W.; Mesri, M.; Neubert, T. A.; Regnier, F. E.; Rodriguez, H.; Spiegelman, C.; Stein, S. E.; Tempst, P.; Liebler, D. C. Interlaboratory study characterizing a yeast performance standard for benchmarking LC-MS platform performance. *Mol. Cell. Proteomics* **2010**, *9* (2), 242–254.
- (7) 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* (2), 225–241.
- (8) Mortensen, P.; Gouw, J. W.; Olsen, J. V.; Ong, S. E.; Rigbolt, K. T.; Bunkenborg, J.; Cox, J.; Foster, L. J.; Heck, A. J.; Blagoev, B.; Andersen, J. S.; Mann, M. MSQuant, an open source platform for mass spectrometry-based quantitative proteomics. *J. Proteome Res.* **2010**, *9* (1), 393–403.
- (9) Colaert, N.; Vandekerckhove, J.; Gevaert, K.; Martens, L. A comparison of MS2-based label-free quantitative proteomic techniques with regards to accuracy and precision. *Proteomics* **2011**, *11* (6), 1110–1113.